A social contract model of 'disintegrity' within the dual-process paradigm of moral psychology: Reducing the scope of the 'belief-behavior incongruity'

Bradford Barnhardt

A thesis submitted in fulfillment of the requirements for the degree of

Doctor of Philosophy

Faculty of Education and Social Work

University of Sydney

December 2014

The University of Sydney

Faculty of Education and Social Work

Office of Doctoral Studies

AUTHOR'S DECLARATION

This is to certify that:

- This thesis comprises only my original work towards the PhD degree.
- Due acknowledgement has been made in the text to all other material used.
- The thesis does not exceed the word length for this degree.
- No part of this work has been used for the award of another degree.
- This thesis meets the University of Sydney's Human Research Ethics Committee (HREC) requirements for the conduct of research.

Signature:

Name: Bradford Barnhardt

Date: 2 December 2014

Acknowledgements

I am deeply grateful to the many people who assisted and influenced my research. My family, first and foremost, was a source of endless patience and support without whom this work would have been impossible. My parents and children provided encouragement and inspiration that was more important to me than they will ever know. My wife, in particular, was an indispensable friend and ideal companion through the many challenges of a raising young family during this degree program, while on the move from Saudi Arabia to Australia to Japan. I am also deeply indebted to my Doctoral Supervisor, Dr. Paul Ginns, for embracing my ideas and ambitions, and challenging my assumptions. I could not have hoped for a more devoted, intelligent, or highly skilled supervisor. A number of other academics additionally went out of their way to assist and encourage my research. Among these individuals I would especially like to thank Associate Professor Richard Walker, Professor Andrew Martin, and Professor Emeritus Ray Debus (1931 – 2014). My doctoral research was also only possible thanks to the many students who made time in their busy lives to participate, and through the generous and highly professional support of teachers and administrators at fifteen international schools around the world.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Explaining why students cheat when it violates their moral beliefs, also called the 'belief-behavior incongruity' (BBI), is a difficult challenge most often overcome by referring to *neutralization techniques*, first described by Sykes and Matza (1957), whereby individuals deceive themselves with specious justifications for ignoring the moral imperative to follow rules. An underlying assumption of the neutralization view, that individuals' abstract moral beliefs apply automatically to all contexts, is critiqued in the present work. The account of academic dishonesty developed herein is centered on the hypothesis that adolescent students' felt moral obligation is informed by an intuitive sense of reciprocity between themselves and their learning contexts, which resembles a social contract, or 'psychological teaching-learning contract' (PTLC). Students who regard a class context or teacher more negatively are thus expected to feel less moral obligation to follow rules, and to cheat more as a result.

The hypothesized PTLC model, which included key variables related to (A) self-concept, (B) achievement goal structure, (C) learning strategies, (D) moral obligation, and (E) social comparison theory, was tested with data from a diverse sample of secondary students in fifteen international schools across Asia, Europe, and Africa. A pilot study (N = 96) of the construct validity of psychometric measures was conducted prior to the Main Study, which included a Time 1 sample of N = 493, a Time 2 sample of N = 297 (spaced by approximately one year), and a longitudinal matched sample of N = 225. Structural equation modeling techniques were used to test the validity and invariance of the measurement model, as well as the structural relations hypothesized between variables. A small degree of gender non-invariance prompted separate analyses of gender-specific models. Results supported the PTLC hypothesis. Moral obligation overwhelmingly mediated the effects of perceived class quality on academic integrity, indicating that students felt morally obliged to be honest in a given class, as a function of their regard for its quality.

To my children.

Think for yourselves.

CHAPTER 1

OVERVIEW: HARD PROBLEMS

In 1900, renowned mathematician David Hilbert famously challenged his field to solve a set of so-called 'hard problems', which he speculated would take most of the coming century. The majority of Hilbert's hard problems have since been solved, and, in the process, have motivated the field of mathematics to challenge its limits (Sampson, 2013). Hard problems are not unique to mathematics. They have also been identified in fields such as law, criminology, and physics. Such challenges press scholars to reassess and often reformulate existing paradigms in order to address the shortcomings of current understanding.

The preeminent 'hard problem' of scholarship on academic cheating is to explain why so many students who view cheating as immoral, still cheat. This discrepancy between moral beliefs and moral action, which pervades research on moral cognition generally (Bergman, 2002; Blasi, 1980; 1983), has been described in cheating literature as the 'judgment-action gap' (Olafson, Schraw, Nadelson, Nadelson, & Kehrwald, 2013) and the 'belief-behavior incongruity' (hereafter 'BBI') (Stephens & Nicholson, 2008). The BBI should not, theoretically, endure in individuals over time. The rational-cognitive paradigm of moral psychology holds that acting against one's moral beliefs should cause cognitive dissonance, which individuals should then strive to eliminate by changing either their beliefs or their behaviors (Aronson, 1968; Blasi, 1980; Jurdi, Hage, & Chow, 2011a). Scholarship on cheating clearly indicates that this does not happen in the way rational-cognitive theories predict. Quantitative studies of cheating, such as those conducted biannually by the Josephson Institute (e.g. 2012), indicate that the BBI is widespread and persistent among secondary students. Approximately half of the students surveyed in 2012 (Josephson Institute, 2012) admitted, for instance, to cheating on a test during the previous year (52%) despite agreeing that 'People should play by the rules even if it means they lose' (92%), and indicating that having good moral character was important to them (98%).

A rational-cognitive explanation for the BBI problem that has been invoked by numerous scholars is that students 'neutralize' the internal discomforts of cognitive dissonance by blaming their cheating behavior on external factors (e.g. Beasley, 2014; Blasi, 1983; Olafson et al., 2013). *Techniques of neutralization* (Sykes & Matza, 1957) have been used in many studies to catalogue the justifications and excuses offered by students for why they cheat. Students who cheat are held, by this view, to understand that cheating is absolutely immoral, and to actively exploit opportunities to blame situations for their behavior.

The inability of rational-cognitive frameworks to explain BBI has, in addition to popularizing the neutralization framework, more recently helped engender support for a dual-process paradigm of moral psychology, according to which rational-cognitive processes operate alongside, and interact with, non-rational processes related to emotion and intuition (Ajzen & Sexton, 1999; Mallon & Nichols, 2010). Dual-process perspectives, which are currently nascent in the literature of cheating (McTernan, Love, & Rettinger, 2014; Murdock, Beauchamp, & Hinton, 2008), may open the way to new explanations of the BBI that exceed the normal limits of rational-cognitive theories of moral judgment, such as that embodied by Kohlberg's (1968) notion of 'children as moral philosophers'. Children might be better portrayed, from the dual-process perspective, as 'philosophizing moral intuits.'

This thesis develops a dual-process framework for cheating in which moral obligation is held to fluctuate as a function of contractarian reciprocal fairness. Within this framework, when students perceive class contexts to be unfair or of low quality, the nature of the rules that forbid cheating may shift in their view from expressing moral imperatives to expressing social conventions (Turiel, 1983, 2002, 2006), due to non-rational processes such as emotion and intuition.

A contractarian theory of adolescent moral judgment does not fit within the rational-cognitive frameworks associated with Jean Piaget (1896 – 1980) and Lawrence Kohlberg (1927 – 1987), wherein moral judgment based on social contract heuristics is a 'postconventional' mental operation, held to be beyond the developmental limits of most adolescents (Colby & Kohlberg, 1987; Colby, Kohlberg, Gibbs, et al., 1983; Rest, 1986). However, studies in the field of evolutionary psychology have shown that contractarian judgment is performed more commonly than these frameworks allow (Cosmides, 1989; Rettinger, 2007). Reciprocal fairness, the most fundamental element of social contracts, is seen to be understood by children as young as three years-old (Cosmides & Tooby, 2013; Olson & Spelke, 2008), as well as by a variety of primate species (Brosnan & de Waal, 2002, 2003; de Waal, 1991, 2014). The Kohlbergian perspective on moral judgment appears to be correct insomuch contractarian judgments in young children cannot be accounted for by rational cognition. Such judgments appear, instead, to occur automatically – a hallmark of emotional-intuitive processes (Cushman, Young, & Greene, 2010; Kahneman, 2011). The dual-process paradigm posits that moral judgments may arise from both rational-cognitive processes, such as moral reasoning, and emotional-intuitive processes, such as disgust (Narvaez, 2010; Schnall, Haidt, Clore, & Jordan, 2008), which, as in the case of detecting violations of reciprocal fairness, often occur too quickly to be accounted for by conscious reasoning (Kahneman, 2011).

The social contract framework for academic cheating developed herein holds that a shift in one's view of cheating from the moral domain to the conventional domain is tantamount to a 'felt' reduction of the moral obligation to be honest, or a reduction in what has also been described as 'moral motivation' (Schroeder, Roskies, & Nichols, 2010). The framework is contractarian in nature because it proposes that within specific contexts, such as classrooms, moral obligation is a "two-way street". Students are hypothesized to feel less morally obliged within academic contexts when they perceive moral failures on the parts of teachers and schools, and *vice versa*. Feeling less moral obligation in a given context should alleviate, or at the extreme preclude, the moral dilemma of BBI for students who cheat. Instead of neutralizing the immorality of cheating, therefore, students may genuinely feel unrestrained by moral imperatives against it – a feeling rooted in non-rational processes associated with social contract-based judgment that may not respond to honor codes, well-reasoned exhortations, or even threats of external punishment.

It is important to note that no judgment is passed in the present work on whether circumstances ever reduce the immorality of cheating. Feeling that circumstances nullify the moral imperative to be honest, does not make it true. To the extent that such feelings are genuine, however, potentially due to non-rational processes rooted in brain architecture (Greene, Nystrom, Engell, et al., 2004; Haidt, 2007; Knoch, Pascual-Leone, Meyer, et al., 2006), they may represent truth to the actor. This perspective suggests a more complex picture of cheating psychology than currently prevails, and advocates for new approaches to addressing the problem.

CHAPTER 2

REVIEW OF THE LITERATURE

> Just as a man cannot be a good cosmopolitan and humanitarian until he has first been a good nationalist, so he cannot be devoted to abstract social ethics until he has served his apprenticeship in personal ethics. To prematurely act on general ethical grounds is to destroy the very foundations of the moral nature. And so we must be patient with children, and university students, and with ourselves until we grow up to social manhood and womanhood.
>
> —Barnes, 1904, p. 488

The present review of literature begins with a discussion of the incidence and definition of cheating. Subsequently, because the broader research project is international in scope, findings on how culture and group-level identity affect cheating are summarized in section 2.2. Sections 2.3 – 2.4 cover the personological and situational predictors of cheating that stand out most prominently in the literature of the past 110 years. Section 2.5 then reviews multivariate person/situation models of cheating that have been developed within the rational-cognitive paradigm of moral psychology and that illustrate the contributions and problems, such as BBI, associated with that paradigm. The inability of such models to predict behaviors based on cognitive factors has frequently been explained as the result of neutralization techniques, whereby students find or invent justifications for cheating behavior in order to evade negative self-perceptions. The neutralization framework is critically reviewed in section 2.6. Finally, section 2.7 reviews contractarian perspectives on cheating that have been raised metaphorically in several major studies, but that have never been studied empirically.

More recent studies will be privileged, generally, over older ones, as will those conducted in secondary contexts, or that involve secondary school-aged subjects, over those conducted in tertiary or post-tertiary contexts. While research into academic integrity at the secondary level has increased considerably over the last decade, studies conducted in tertiary settings are still far more common and contain many valuable insights about cheating that generalize to the secondary level. Tertiary students report cheating for many of the same reasons, and based on many of the same attitudes, as high school students. It has been commonplace, therefore, to capitalize on the greater depth and range of work provided by tertiary literature in studies of secondary cheating (Miller, Murdock, Anderman, & Poindexter, 2007).

2.1 Definition and incidence of cheating

Many studies have found that, for several decades, cheating incidence among American students has been approaching 'epidemic proportions' (e.g. Desruisseaux, 1999). While such findings often accompany concern over the apparent crash in American morality that an 'epidemic' of cheating suggests, the incidence and moral implications of cheating depend largely on how it is defined. Students and teachers have been found to differ in how they define cheating (Broeckelman-Post, 2008; Higbee, Schultz, & Stanford, 2011), and in how they believe various acts of cheating should be judged, and punished (Feinberg, 2009). This is an important consideration for researchers whose choice of definition influences how they measure cheating, and what implications they draw from their results.

The incidence of cheating is widely perceived to have increased dramatically for several decades, in both American secondary schools (e.g. Zito & McQuillan, 2011; Galloway,

2012) and colleges (e.g. Bernardi, Baca, Landers & Witek, 2008), to levels approaching 'epidemic' (Alschuler & Blimling, 1995; Desruisseaux, 1999; Haines, Diekhoff, LaBeff, & Clark, 1986; Miller et al., 2007; Schraw, Olafson, Kuch, et al., 2007; Seider, Novick & Gomez, 2013; Stephens & Nicholson, 2008; Wellborn, 1980). In a recent study by Galloway (2012), 93% of 4,136 American high school students in grades 9-12 reported having cheated at least one time during their high school careers. Cheating was defined by Galloway's (2013) study as engagement in any of thirteen cheating behaviors listed on an inventory developed by McCabe and Treviño (1993) (e.g. turning in work done by another; using cheat sheets; working together when the instructor asked for individual work; getting an extension using a false excuse). Behavioral inventories such as this have been widely used in research on cheating for the last half-century, at both the college and high school levels (e.g. Evans, Craig & Mietzel, 1993; McCabe & Treviño, 1997; Miller, Shoptaugh & Wooldridge, 2011; Schab, 1969, 1980, 1991).

The high incidences of cheating frequently found on wide-ranging inventories of cheating behavior tend to support the 'epidemic' narrative, especially when summarized as the percentage of students who report having cheated 'in any form' (e.g. O'Rourke, Barnes, Deaton et al., 2010). It is not always clear, however, whether respondents to such measures realize that researchers will interpret all of the listed behaviors unequivocally as cheating. Informing students of how researchers define cheating does appear to have a significant effect on how much cheating is reported. Burrus, McGoldrick, & Schuhmann (2007) examined the effect of providing respondents with a clear definition of cheating on its incidence in a self-reported format, by asking 300 American undergraduates to respond to an inventory of cheating behaviors twice: once before being provided with the definition, and once after. The incidence of reported cheating increased significantly after the definition was provided. This means that, on the one hand, students did not initially understand what the researchers considered cheating to be. On the other hand, after being provided with a definition, respondents reported acts of cheating they had apparently not understood to be cheating when they committed them. It is not immediately clear, therefore, which incidence is the more valid; the pre-definition incidence, which included only those acts that students themselves defined as cheating, or the post-definition incidence, which additionally included acts of cheating that were unintentional. Bisping, Patron, and Roskelley (2008) addressed this conceptual issue by asking students to indicate both whether they had 'done it', with respect to each of 31 cheating behaviors gleaned from across the literature, and whether they 'knew it was wrong'. In the case of 22 behaviors, half or more of the respondents who admitted to engaging in them also indicated not having known it was wrong.

All forms of academic misconduct, accidental or otherwise, are harmful to equity, undermine the mission of schools to foster intellectual growth, and should be taken seriously. The urgency expressed in the literature frequently emphasizes above all, however, the specter of moral decline, which is apparently portended by the ever-increasing prevalence of cheating (e.g. Callahan, 2004; Haines et al., 1986; Harding, Carpenter, & Finelli, 2012; McCabe, 1999; Schraw et al., 2007). Inasmuch as moral behavior is a function of internal moral judgment (Blasi, 1980), a trend of moral decline implies that students are becoming more likely to commit acts they judge to be immoral. This perspective champions the pre-definition incidence of cheating in Burrus et al. (2007), discussed above, in that it excludes unintentional cheating. The alternative perspective, that cheating is defined externally, champions the post-definition incidence of cheating in Burrus et al. (2007), which ignores whether the acts of cheating are recognized as such by students when they commit them. The latter of these perspectives is implied by measures of cheating 'in any form' that do not differentiate between intentional and unintentional acts. While such measures may be legitimate from an administrative perspective, in terms of the equity and integrity of educational programs, they provide a weak basis for raising moral concerns because they ignore intentionality.

Students' judgments of the immorality of cheating have also been found to vary predictably by what type of cheating was committed, to what extent, and why. Two multidimensional scaling analyses of students' perceptions of cheating, conducted at the secondary and tertiary levels by Liora Schmelkin and colleagues (Schmelkin, Gilbert, Spencer, et al., 2008; Schmelkin, Gilbert & Silva, 2010) found that students judged cheating along two dimensions: (1) papers *vs*. exams, and (2) serious *vs*. trivial. Exam cheating behaviors such as using crib notes, copying during exams, and passing along questions and answers to peers in later sections were among the infractions that students viewed as being most serious. This implies, moreover, that students would be more likely to recognize these acts as cheating, without needing a definition, than acts they consider 'trivial'.

Acts of cheating on exams were also found to form a coherent factor in a study conducted among tertiary students in the UK by Franklyn-Stokes and Newstead (1995), whereas Rakovski and Levy (2007) identified two coherent factors related to severity (greater *vs.* lesser) in an analysis of American tertiary students' perceptions of fifteen acts of cheating. Respondents in the latter study indicated believing that more serious forms of cheating deserved "higher penalties", and that all forms of exam cheating were "more serious" (pp. 476-477).

These studies indicate, in sum, that students' judgments of the morality of various acts of cheating have a consistent latent structure, wherein acts of cheating on exams are generally considered among the worst. For this reason, self-report measures that lump all forms of cheating together as a homogenous moral abstraction known as 'cheating' may lead to unnecessarily dire concerns about moral decline. Ignoring the variation in how students judge the seriousness of different cheating behaviors overlooks differences in the intrinsic moral content of such behaviors. Figures 2.1 – 2.3 present data from studies published over the last eighty years, of high school cheating incidence on tests *vs*. 'in any form'.

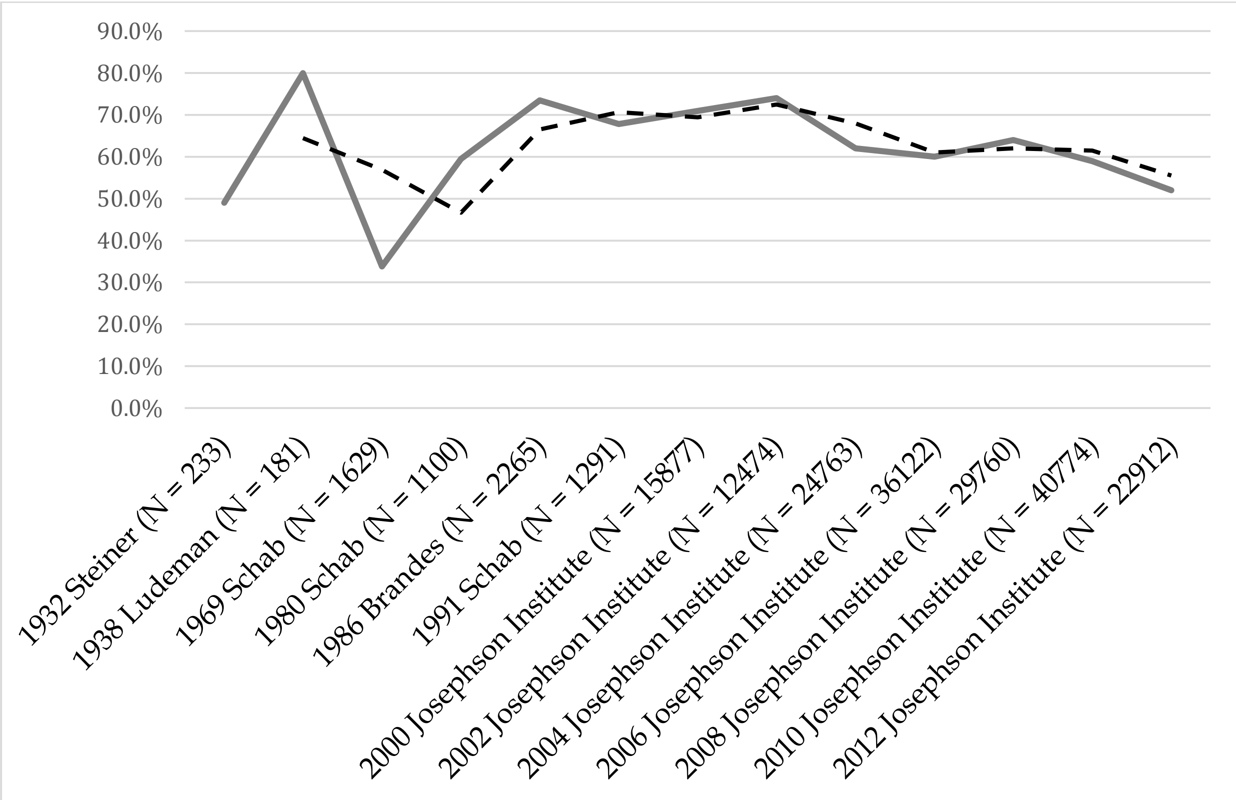

Figure 2.1. Incidence of high school cheating – on tests (1932-2012); solid line = individual data points; dashed line = moving average

Figure 2.1 presents the results of thirteen measurements of cheating on tests between 1932 and 2012. Some of the variability may reflect the use of different methods. For instance, Steiner (1932) used an experimental technique, whereas Ludeman (1938) asked college students the question "Did you cheat in high-school tests?" The measures used by Schab (1969, 1980, 1991) and Brandes (1986) both asked specifically whether students had used crib notes to cheat on tests. Brandes (1986) also asked whether students had copied from other students during tests, and found similar results for both measures (73.5% and 75%, respectively). The Josephson Institute took a broader approach on all of its biannual *Ethics of*

American Youth questionnaires, by simply asking respondents whether they had cheated on tests during the past year. The moving average for cheating on tests over the last eighty years rises from approximately 47% in the 1960s to a sustained level of approximately 70% between Schab (1991) and Josephson Institute (2002), after which it declines, overall, until Josephson Institute (2012).

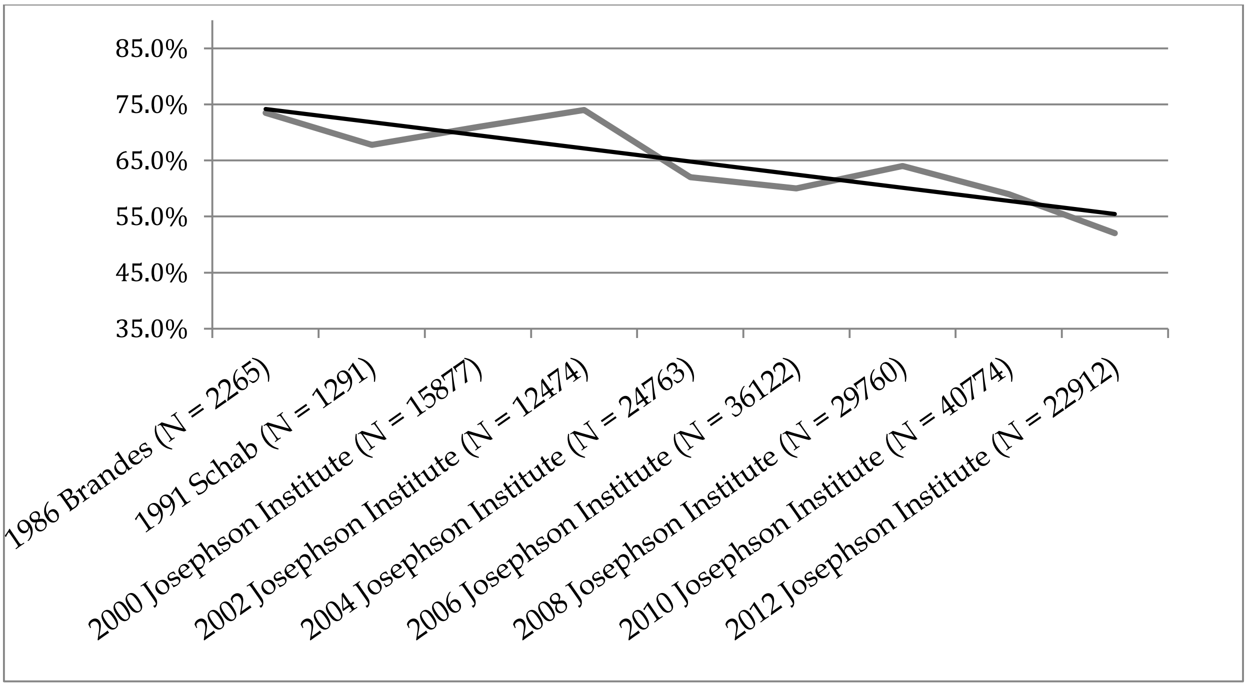

Figure 2.2. Incidence of high school cheating – on tests (1986-2012).

Figure 2.2 shows a twenty-six year decline in the rate of high school cheating on tests, which Schmelkin et al.'s (2010) multidimensional analysis found to be the most 'serious' form of cheating in the view of most high school students. Brent and Atkisson (2011) found similarly that, among 420 undergraduates at an American university, cheating on tests was viewed as more 'serious' than cheating on homework. "They offer fewer justifications for it," the authors observe, "further acknowledging the illegitimacy of such cheating" (p. 655). The observation that during 2012 approximately 50% of American high school students cheated on tests has distressing moral and administrative implications. That the problem appears to have been improving for a quarter-century suggests, however, that a growing body of research on cheating over the same period may have had benefits. Figure 2.2 does not, at any rate, portray an increase in immoral behavior.

The data in Figure 2.3 tell a very different story. According to these data, the incidence of high school cheating 'in any form' has increased dramatically over the past two decades. Ignoring differences in the seriousness of the forms of cheating lumped together by these studies, one is inclined to interpret this trend as a crash in student morality. While it is possible that these data do reflect an epidemic rise of cheating 'in any form', however, Burrus et al.'s (2007) finding that students report more cheating when they understand how it is defined, suggests an alternative explanation. The increasing prevalence of self-reported cheating 'in any form' may, especially in studies where measured behaviors are not identified explicitly as cheating, reflect an increasingly sophisticated awareness among students of the full range of behaviors considered to be cheating, which may lead, in turn, to increased rates of reporting.

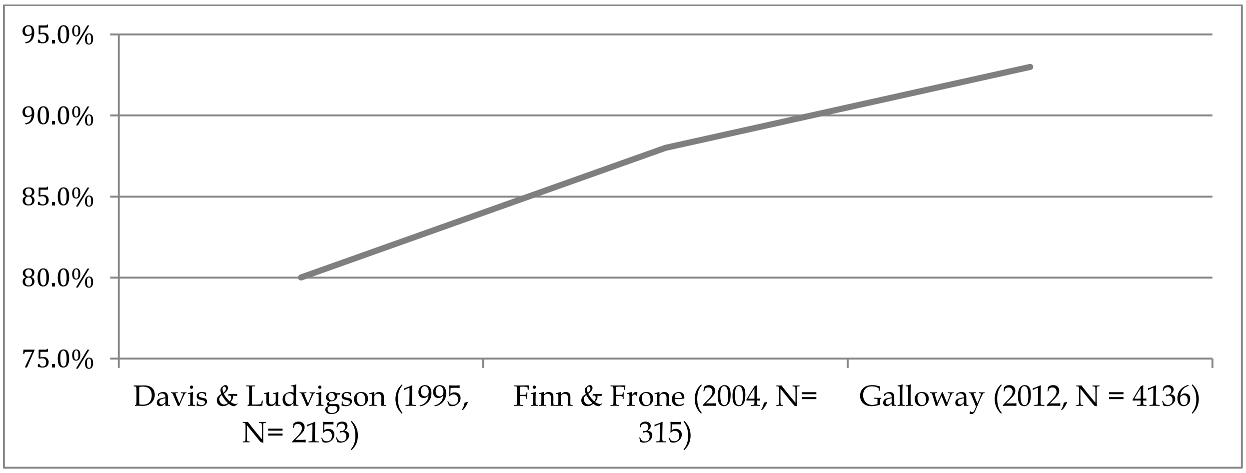

Figure 2.3. Incidence of high school cheating – 'in any form' (1995-2012)

While neither of the trend lines in the preceding two figures (2.2 and 2.3) is proof of an underlying pattern of moral change, the negative trend line for cheating on tests in Figure 2.2 makes a more compelling moral case than does the positive trend line for cheating 'in any form' in Figure 2.3, because of the seriousness that students ascribe to cheating on tests. From a rules-based perspective on why cheating is wrong, however, a student's judgment of the seriousness of different forms of academic dishonesty is not necessarily 'moral' judgment. Perceptions of seriousness may also reflect perceived risk. Eisenberg (2004) used vignettes with a sample of 161 Israeli middle school students to examine how attitudes toward cheating might differ between those who view it as a moral issue and those who view it as a conventional, or rules-based, issue. Forty-five percent of the sample was found to take a conventional, or a-moral, view of cheating, *versus* fifty-five percent who viewed it in moral terms. As Eisenberg (2004) expected, students classified as "morals" were found to have less favorable attitudes towards cheating than those classified as "a-morals". While all students appeared to recognize rules-based definitions of cheating, only a fraction of that number understood why cheating was morally wrong.

Standing in contrast to behavioral inventories are conceptions of cheating that emphasize the abstract, unifying definitional properties of cheating behavior, such as intentionality (Colnerud & Rosander, 2009; Feinberg, 2009), shirking responsibility (Evans et al., 1993), gaining unfair advantage (Fang & Casadevall, 2013), and breaking school rules (Eisenberg, 2004). Garavalia, Olson, Russell, and Christensen (2007) object to abstract definitions of cheating that fail to include unintentional acts because such definitions, they argue, "allow clever individuals to readily defend academic dishonesty" (p. 34). They adopt, instead, Cizek's (2003) tripartite definition of cheating (1. unauthorized information exchange, 2. use of prohibited materials, 3. otherwise gaining unfair advantage), which amounts to shorthand for the lengthy inventories of cheating behavior discussed above. While an inventory approach to cheating behavior may be more cut-and-dry than an abstract definition for administrative purposes, the prevention of excuses is also more of a policing issue than a scholarly one. By ignoring intentionality, many inventories also ignore whether individual acts of cheating have moral content in the minds of the actors. Definitions of cheating that focus on its abstract properties provide, by contrast, more fundamentally accurate expressions of the consciously immoral and/or rule-violating character of what students mean when they admit to having done something called 'cheating'.

Several scholars have recently expanded the notion of cheating to include behaviors that, while not involving cheating in the mainstream sense of the word, share with it the crucial abstract quality of obtaining grades without learning. These behaviors, such as rote memorization, 'plug-and-chug' applications of mathematical formulae, and otherwise attending exclusively to the superficial aspects of learning tasks in order to 'earn' grades with minimal effort, have been grouped together with cheating under the term 'disintegrity' (Miller et al., 2011). Acts of disintegrity "may lack integrity or subvert the goals of education, [though] we may or may not refer to them as cheating" (p. 170). Kohn (2007a) described a similarly broad conception of cheating in terms of privileging product over process, citing Renard's (1999) lament that the countermeasures deployed against cheating behaviors generally ignore the fundamental problem of systemic forces, such as high stakes assessment regimes, that teach students that "the final product takes precedence over learning" (p. 11). In his foreword to the monograph *Psychology of Academic Cheating*, Kohn (2007b) used the term "legal cheating" to refer to strategies that involve "teaching to the test… when real instruction gives way to extensive exam preparation [such that] scores can be raised without improving learning at all" (p. XII). It follows from this that all activities, whether teacher- or student-driven, that focus on the superficial aspects of learning, purely for the sake of obtaining grade-credentials, share important abstract definitional properties with cheating, such as intentionality and shirking responsibility. This novel conception, referred to hereafter as 'disintegrity', has never been investigated in an empirical study.

Measures of cheating as an abstraction, which tend to comprise fewer items, have been used in a number of studies at the secondary level (Anderman, Cupp & Lane, 2010; Anderman

& Midgley, 2004; Brown-Wright, Tyler, Stevens-Watkins et al., 2012; Murdock, Hale & Weber, 2001). Two measures of self-reported cheating as an abstraction are the three-item scale from the *Manual for the Patterns of Adaptive Learning Scales* (PALS), developed by Midgley, Maehr, Haruda, et al. (2000), and a five-item scale, developed by Anderman, Griesinger, and Westerfield (1998). These scales tend to include a mix of items that refer either to acts that are readily recognized by students as cheating, such as cribbing during exams, or to the concept of cheating writ large, as a realm of behavior unified by definitional properties that both students and researchers would agree on. The PALS scale asks, for instance, whether respondents have (1) copied during tests, (2) copied from other students during class work, and (3) cheated on class work. The first and second of these items refer to behaviors that students are held to recognize readily as cheating, whereas the third item refers to cheating by name, thus appealing to students' knowledge of what 'cheating' on class work entails, in a moral and/or regulatory sense. Phrasing the third item this way escapes the ambiguity around what acts students do and do not regard as 'true' academic transgressions, by asking them, very simply, whether they think they have transgressed.

Cheating is a multidimensional construct that students, educators, and researchers often interpret differently. The amount of attention that the definition of cheating has received in recent literature highlights a substantial amount of disagreement over whether all acts of academic misconduct should be viewed unequivocally as cheating, or whether cheating should be treated as a broad realm of behavior unified by abstract definitional properties. The demonstrated complexity of students' judgments of cheating suggests that lengthy inventory measures of cheating do not capture moral conceptions of cheating. This is because they (1) fail to account for variations in the degree of immorality that students ascribe to different acts, and (2) include non-intentional cheating. The most effective measures of cheating as an intentionally immoral behavior include items that query respondents' engagement in dishonest acts that students tend to recognize as 'more serious' (Rakovski & Levy, 2007; Schmelkin et al., 2008; Schmelkin et al., 2010), and/or ask students directly whether they have cheated.

2.2 Socio-cultural systems and group-level identity

Individual identity is shaped largely by affiliation with *socio-cultural systems*, which have been described as encompassing the "patterns-of-life-of-communities" (Keesing, 1974, p. 82). Socio-cultural systems have traditionally been seen to characterize national, regional, religious, ethnic, and artistic groups (Tylor, 1871), as well as industries and organizations such as corporations (e.g. Sørensen, 2002), and schools (e.g. Hargreaves, 1995). Among the most important defining features of socio-cultural systems are behavioral norms.

Individual cheating behavior has been found in experimental studies to be influenced by perceptions of cheating among members of groups with which individuals identify. Gino, Ayal, and Ariely (2009) investigated, for instance, the degree to which undergraduate students' cheating behavior fluctuated in response to overt cheating by a 'confederate'. Confederate cheaters were either readily identifiable as members of the subjects' in-group (a college-age individual wearing a sweatshirt from the subjects' university) or as members of out-groups (older and wearing regalia from a different university). Subjects who observed the 'in-group confederate' cheat successfully and receive the maximum reward subsequently cheated more than subjects who observed the same behavior by 'out-group confederates'. The purpose of the confederate was to challenge social norms for honesty that subjects might have assumed to prevail in the experimental setting. The fact that in-group confederates were successful at challenging norms related to honesty, while out-group confederates were not, is evidence that socio-cultural identification has the power to shape context-bound behaviors such as cheating (see also Callahan, 2004; Crittenden, Hanna & Peterson, 2009b; Kidwell, Wozniak & Laurel, 2003). Membership in Greek fraternities and sororities has, for instance, been positively correlated with cheating behavior in numerous studies since the early Twentieth century (Burrus et al., 2007; Harding et al.; 2012; Parr, 1936; Whitley, 1998), as have political and religious affiliations in the United States. Shipley (2009) found that college students who identified with the liberal Left had, on average, stricter views on cheating than those who identified with the conservative Right, whereas Hartshorne and May (1928) found that Baptists cheated more than Lutherans.

Cheating has also been found to relate to socio-cultural aspects of secondary schools. School level features associated with cheating include motivational goal structure (Anderman et al., 1998), climate (Stephens & Nicholson, 2008; Zito & McQuillan, 2011), and moral tone (Steiner, 1930, 1932; Stephens & Nicholson, 2008). In a seminal study of the relationship between cheating and achievement goal motivation, Anderman et al. (1998) found that cheating was predicted by perceptions of school-level performance goal structure ($\beta = .49$), which implies a school-level socio-cultural system marked by heightened competition and peer comparison.

Socio-cultural influences on cheating appear to operate also at the national level. In a study of 6,226 tertiary-level business students in 36 nations, Crittenden, Hanna and Peterson (2009a) found significant correlations between participants' views on business ethics aggregated by nation, and the rankings of those nations according to Transparency International's *Corruption Perceptions Index*. Students in countries ranked as less corrupt were significantly less likely to agree with the statements "In order to succeed in business, it is often necessary to compromise one's ethics" ($r = -.158$, $p < .001$), and "Business behavior that is legal is ethical" ($r = -.173$, $p < .001$) (p. 7). A study of 7,213 undergraduates in 21 countries by Teixeira and Rocha (2010) found, additionally, that students in countries rated as the least corrupt by Transparency International, such as Denmark and Sweden, reported the lowest incidence of academic cheating (0-10%).

Two multi-national studies of cheating have indicated that national identity also influences cheating among high school students. Firstly, in a study involving 322 high school students from West Germany, Costa Rica, and the United States, Evans et al. (1993) found that cheating was perceived to be a significantly less serious problem among German students than among American or Costa Rican students. This finding may reflect socio-cultural differences at the national level or, in fact, at the school level. The authors point out that German participants in this study were drawn from the segment of the German education system known as ‘gymnasium’. Gymnasium is the university track in Germany, wherein students may experience a less competitive school environment than in preceding years when they were competing for entry into more desirable educational tracks (Miller et al., 2007).

Secondly, Magnus, Polterovich, Danilov, and Savvateev (2002) conducted a study of attitudes towards cheating among 885 students from four countries (Russia, Israel, the United States, and the Netherlands), that included 92 high school students from Russia (N = 73) and the Netherlands (N = 19). As in the two exclusively tertiary studies mentioned above, differences in attitude toward cheating were found, at the national level, to be consistent with national rankings on Transparency International’s *Corruption Perceptions Index*. High school students were significantly more tolerant of cheating in Russia than in the Netherlands, which were ranked, respectively, as the most and least corrupt countries in the study. Russian participants also indicated ‘hating’ students who report cheating to authorities, which Magnus et al. (2002) interpret as a reflection of anti-government sentiment in post-Soviet Russia, embodied by the saying “First whip the informer” (p. 128). Latova and Latov (2008), who investigated the use of crib notes among Russian secondary and tertiary students, describe in greater depth the effects of Russian society and culture on academic cheating. They cite the persistence of an illicit shadow economy of massive proportions that exerts a corruptive influence on the mentality of many Russians. In this shadow economy, low-level crime such as video and software piracy is ubiquitous. “In the minds of the majority of Russians”, the authors write, “there is a common opinion that cheating is one of the obligatory components of school” (p. 26).

The behavioral norms that one perceives among the members of the socio-cultural groups with which he or she identifies appear to exert strong influences on his or her behavior. Perceiving that in-group members, at the school, community, and national levels, believe cheating is justifiable appears to increase, on average, one’s tolerance of academic cheating. Despite the potentially crucial role of school culture in shaping attitudes that both reflect and perpetuate broader socio-cultural norms, very few international studies of cheating have been conducted at the secondary school level.

2.3 Personological Variables

Research into the causes of academic integrity can be organized into two broad categories: studies that investigate students’ personal characteristics, and studies that examine situational or contextual factors (McCabe & Treviño, 1997). These two broad veins of research overlap in many studies, and have converged more recently into what could be described as a third line of research characterized by multivariate models that include both personological and contextual factors related to cheating.

Dozens of personological variables have been studied in relation to cheating, among which the most prevalent are demographics. While scholars have recently argued that covariates of cheating such as age, gender and socio-economic status are of limited use to educational researchers or practitioners (Miller et al., 2007), these variables continue to receive attention in most empirical studies. Personological variables emphasized in the literature also include personality constructs such as conscientiousness and locus of control, as well as self-belief factors such as self-efficacy and self-concept.

2.3.1 Gender

While findings related to gender are mixed, the literature suggests that, overall, cheating is more common among males who are younger than females who are older, at both the secondary (Finn & Frone, 2004) and tertiary levels (Newstead, Franklyn-Stokes, & Armstead, 1996). Of twenty-seven studies of secondary cheating to have reported on gender since Hartshorne and May (1928), all but two found evidence that males tend to cheat more than females. Feldman and Feldman (1967) found, using an experimental method, that Grade Seven females cheated more than their male counterparts, whereas David (1973, cited by Bushway & Nash, 1977) found that, among a single group of American students, females cheated more on a math test, whereas males cheated more on a vocabulary test. Gender differences at the tertiary level are somewhat less consistent, tending to disappear in experimental studies and field observations (Whitley, 1998). A meta-analysis of gender effects related to college cheating conducted by Whitley, Nelson, and Jones (1999) found a small effect size (d = .17, Z = 25.98, p < .001) that suggested college males might, indeed, cheat more.

Gender differences in cheating behavior have been explained in at least two ways. Firstly, Whitley (1998) suggests that gender differences on self-report measures may reflect a tendency among females to under-report their actual cheating due to higher guilt-proneness. At least three experimental studies at the tertiary level report finding that females cheat more than males (DePalma, Madsey, & Bornschein, 1995; Jacobson, Berger, & Millham, 1970; Leming, 1980), whereas Canning (1956) reported finding that, under experimental conditions, females were more likely to lie about cheating.

Secondly, it has been suggested that gender differences with respect to cheating reflect gender socialization (Ward & Beck, 1990). In a study involving 229 American high school students, for instance, Jensen, Arnett, Feldman, and Cauffman (2002) found that cheating was rated as less acceptable among females than among males. Inasmuch as cheating behavior is less accepted by one's in-group members, who among adolescents are frequently of the same gender, it may be a greater cause for shame to cheat, or to admit to having cheated. A study of 2,197 Taiwanese high school students conducted by Tsai (2012) also found that peer influence played a significant role in gender differences. Females indicated being more susceptible to the influence of other females than to the influence of males, or than males were to that of either gender. Taken together, these studies suggest that societal gender roles may exert considerable pressure on individual-level cheating behaviors via peer-to-peer influences.

2.3.2 Age and grade-level

While the practice of basing grade-level groupings on student age produces a large correlation between these two variables at the secondary level, they do imply distinct sources of variance. Age is generally associated with internal processes of cognitive and physiological development, whereas grade-level refers to one's external learning context. This distinction is important from the practitioner's point of view, when attempting to understand how changes in internal factors, external factors, or interactions between the two affect cheating.

Secondary students appear to cheat more as they get older and progress to higher grade-levels. Following approximately the end of high school, however, the prevalence of cheating is found to decline (for reviews see Cizek, 1999; Miller et al., 2007). An exception to this general observation was reported by Steiner (1932), who found, using an experimental technique on students from eight separate secondary schools, that the incidence of cheating actually decreased from Grade Seven to Grade Ten. A survey-based study conducted shortly thereafter by Ludeman (1938) found, by contrast, that more college students reported having cheated in high school (80%) than in grade school (43%), and indeed, most of the subsequent literature affirms this general pattern (Davis & Ludvigson, 1995; Evans & Craig, 1990a; Galloway, 2012; Jensen et al., 2002; Miller et al., 2007; Schab, 1969, 1991).

In college, statistical associations between cheating, age, and year show a distinctly different pattern. The prominent correlation between grade-level and cheating in secondary school disappears in college, and the direction of association between age and cheating becomes negative. Whitley's (1998) review of 107 tertiary-level studies of academic integrity concluded that "although cheating is negatively correlated with age, it is essentially uncorrelated with year in college, $d = -.038$" (p. 242).

Kohlberg's (1958) theory of moral development posits that moral reasoning should progress with age, according to Piaget's stages of cognitive development. The Kohlbergian framework holds that younger individuals should be less morally developed, on average, than older individuals, and therefore more likely to cheat (Briggs, Workman, & York, 2013; Kohlberg, 1971). Experimental studies of the relationship between moral reasoning and cheating among primary- and secondary-age children conducted during the latter half of the Twentieth-century were reviewed by Blasi (1980), who found, however, that there was little evidence in favor of the purported link between moral reasoning and moral behavior. The general hypothesis that less engagement in immoral acts will occur as moral cognition matures with age is also contradicted by the above-mentioned findings that cheating tends to increase as students matriculate to higher levels in secondary schooling.

An alternative source of variance in how cheating behaviors change over time is educational context. As students matriculate to higher grade-levels, they tend to experience greater degrees of challenge and intensifying pressure to achieve good grades. In American education, for instance, middle school grades are not usually included on the transcripts that students submit with their college applications, whereas from beginning of Grade Nine, the grades that students receive directly influence their college prospects. Many students may, therefore, feel markedly more pressure to make good grades from the moment they transition to high school. In the only longitudinal study of cheating over the transition from Grade Eight to Grade Nine, Anderman and Midgley (2004) found that it increased principally for students who perceived greater competitive pressures in Grade Nine. This finding is consistent with a growing literature on the effects of motivational structure that suggests that more competitive, grade-focused educational contexts foster more cheating (e.g. Anderman et al., 1998; Bong, 2008; Murdock et al., 2001; Tas & Tekkaya, 2010).

Educational experience may also affect cheating in a cumulative manner over time, through self-belief variables such as self-esteem, self-efficacy, and self-concept. Self-belief variables are held both to reflect past self-experience and to shape future behavior (Bandura, 1997; Bong, Cho, Ahn, & Kim, 2012; Pajares, 1996). Students who, due to myriad circumstances, cheat at lower grade-levels may come, therefore, to see themselves as less honest, and may cheat more as a result. Self-beliefs are discussed in the next section.

2.3.3 Personality and self-beliefs

The intra-psychic factors that have been related most consistently to cheating in empirical studies are self-belief and personality constructs that feature aspects of morality and control, such as conscientiousness, self-control, and self-efficacy. According to social cognitive theory (Bandura, 1997; Pajares, 1996), self-beliefs reflect one's past experiences of self that are both central to individual identity, and that shape future behavior accordingly (Bong et al., 2012). Self-beliefs are, as such, subject to change over time, especially during adolescence (Schwartz, Klimstra, Luyckx et al., 2012). Viewed as key channels for the reciprocal influence of experience and behavior, self-belief variables appear to play important roles in how individuals' attitudes toward cheating evolve over time. It is not clear, however, to what extent personality factors, generally conceptualized as stable and trait-like, shape self-beliefs. Inasmuch as personality factors affect behavior, the self-perceptions upon which self-beliefs are theoretically based may, in fact, be shaped by personality. While personality factors and self-beliefs are generally thought of as fundamentally different phenomena, they may often share a great deal of common variance. Incorrectly assuming that measures with different names represent fundamentally different phenomena, or, conversely, that measures of the same name represent the same phenomenon, are long-standing problems in social sciences that have, together, been referred to as the *jingle-jangle fallacy* (Kelley, 1927, cited in Marsh, 1994; Marsh, Craven, Hinkley & Debus, 2003).

Self-esteem, a relatively weak correlate of cheating (Aronson & Mettee, 1968; Van Gundy, Morton, Lui, & Kline, 2006; Whitley, 1998), offers an excellent example of the jingle-jangle fallacy. Self-esteem, which refers to the overall positive/negative evaluation one makes of oneself, has been widely studied in social sciences, and, as a result, has been measured in a variety of ways (Bong et al., 2012; Rosenberg, 1965; Rosenberg, Schooler, Schoenbach, & Rosenberg, 1995; Van Gundy et al., 2006). A review of self-esteem literature conducted by Scheff (2011) identified no less than 200 distinct measures, representing multiple interpretations of the construct. This stands as an example of the 'jingle' fallacy, whereby disparate measures are grouped under the same name and treated as if they represent the same thing. Moreover, the antecedents of self-esteem, identified by Van Gundy et al. (2006, p. 374) as "processes of reflected appraisal…, comparison…, and self-attribution", happen to be strikingly similar to the antecedents of 'global self-concept' (Campbell, 1990). Global self-concept, which comprises the affective and cognitive aspects of how individuals view themselves (Kornilova, Kornilo, & Chumakova, 2009), has, like self-esteem, small statistical associations with cheating (Rost & Wild, 1994), and has been found to be multicollinear with measures of self-esteem in empirical studies (Harter, 1999). This appears, therefore, to stand as an example of the 'jangle' fallacy, whereby measures with different names, such as 'self-esteem' and 'global self-concept', are incorrectly assumed to measure fundamentally different constructs.

Measures of self-concept that are domain-specific, such as *Subject self-concept* and *Honesty-trustworthiness self-concept*, have a strong and growing research base in secondary educational literature (Leung, Marsh, Yeung, & Abduljabbar, 2015; Marsh, 1989; Marsh, 1992; Marsh et al., 2005). None of these, however, has been used to investigate secondary cheating. At the tertiary level, Antion and Michael (1983) found relatively small correlations between five self-concept dimensions (levels of aspiration, anxiety, academic interest and satisfaction, leadership and initiative, and identification *vs.* alienation) and both the amount and incidence of cheating ($r = |.07 - .20|$). Arvidson (2004) employed a measure of self-concept, also at the tertiary level, that included fifteen subscales. Significant correlations between cheating and intellectual, scholastic, self-worth, and moral self-concept were uniformly small and negative across an inventory of 22 cheating acts (*r*'s ranging from -.10 to -.28). Among these, the four largest correlation coefficients all involved moral self-concept.

Constructs that are suggestive of moral self-concept have also produced larger effects in several other studies. Whitley (1998) identified, for instance, three studies of the moral obligation not to cheat, with an overall effect size of $d = -.79$. 'Moral obligation not to cheat' remains of interest in current literature, particularly as a key construct in Ajzen's (2002) theory of planned behavior (Mayhew et al., 2009; Simkin & McLeod, 2010; Harding et al., 2012). A similar construct, dubbed 'self-rated honesty', has also been found to negatively predict cheating in several tertiary studies (Burrus et al., 2007; Rakovski & Levy, 2007; Whitley, 1998).

A self-belief-related construct that has been related recently to ethical behavior is 'moral self-concept maintenance'. In a series of six experiments, Mazar, Amir, and Ariely (2008) found that participants allowed themselves to benefit from limited amounts of dishonesty by re-categorizing dishonest acts in ways that allowed perpetrators to maintain a positive moral self-concept. Participants cheated to acquire more reward money when, for instance, tokens were used as an intermediary. Even though the tokens would be converted to money later, cheating for tokens was less damaging to the subjects' moral self-concept than cheating directly for money. Two laboratory experiments in an Israeli context, reported by Shalvi, Dana, Handgraaf, and De Dreu (2011), found, similarly, that participants actively balanced the benefits of lying against the harm it posed to their self-concept. When opportunities were made available for participants to plausibly justify acts of deception, and earn more money as a result, the balance tipped and they lied more. The authors argue that the justifications allowed participants to engage in more dishonesty without correspondingly greater harm to their self-concepts, a phenomenon referred to as 'ethical maneuvering'. These results suggest that positive self-concept has a prominent moral dimension that makes it vulnerable to immoral behavior.

Moral self-concept has also been used to reduce cheating under experimental conditions. Bryan, Adams, and Monin (2013) found, in an experimental field study of individuals chosen at random on Stanford University campus, that being asked to "not be a cheater" resulted in less cheating in a guessing game than being asked not "to cheat". The implication of this finding was that asking someone not to *be* something negative, i.e. a cheater, was more powerful than asking someone not to *do* something negative, i.e. to cheat. The reference to 'being' appeared to implicate participants' moral self-concepts, and motivate them to be honest.

The strength of the moral component of individual self-concept appears, however, to vary with self-control. Gino, Schweitzer, Mead, and Ariely (2011) found that participants in a set of experiments, whose cognitive resources were depleted by demanding mental tasks, demonstrated less self-control and a higher likelihood of unethical behavior. An exception to this trend was found, however, among participants who indicated having strong moral identity on a post-treatment questionnaire. Strong moral identity appeared to override the impulse to cheat for material gain by moderating the relationship between self-control and behavior. Individuals with stronger moral identity exhibited better self-control, even after other mental resources had been depleted.

Control is a theme that emerges consistently in research on intra-psychic factors. Conscientiousness is a personality construct related, for instance, to both morality and self-control. Conscientious individuals tend to be organized, responsible, and methodical (Day, Hudson, Dobies, & Waris, 2011; Kisamore et al., 2007; Lee, Ashton, & Shin, 2005). Tertiary-level studies of conscientiousness and cheating generally find small to moderate inverse associations (de Bruin & Rudnick, 2007; Nathanson et al., 2006). In a vignette study of attitudes toward cheating that involved 44 undergraduate business students, Day et al. (2011) found, for instance, that conscientiousness predicted judgments of the morality of cheating (β = -.33), the likelihood of cheating (β = -.44), and the justifiability of cheating (β = -.44), as well as the degree to which either the teacher was at fault (β = -.26) or the student was at fault (β = .33) for cheating under various circumstances.

Impulsivity is a personality construct of recent interest that, in contrast to conscientiousness, involves a tendency to act without thinking through the consequences. Impulsivity has been positively associated with cheating at both the tertiary level (Kelly & Worrell, 1978; Nagin & Pogarsky, 2003) and the secondary level (Anderman et al., 2010). Impulsive individuals tend to exhibit less self-control, in that they pay less heed to the risks and costs associated with the behavioral choices they make (Anderman et al., 2010). While at the tertiary level, Kelly and Worrell (1978) found that impulsivity was related to cheating in females exclusively, Anderman et al. (2010) found that the odds of cheating among secondary students increased by 3.74 times with each one-unit increase on impulsivity, irrespective of gender.

'Locus of control' has been an especially consistent personality correlate of cheating behavior (Crown & Spiller, 1998). Locus of control refers to whether an individual tends to identify the causes of event outcomes as being internal or external to him/herself (Rotter, 1966; Srull & Karabenick, 1975). Whitley (1998) reports that while the relationship of locus of control with cheating appears, overall, to be small (d = .27), the relationship is found to be stronger in laboratory and experimental studies than in self-report studies. An internal locus of control tends to lead students to cheat when they perceive that success on a particular task is a matter of skill, whereas students who externalize control tend to cheat more when they perceive that success is a matter of chance (Whitley, 1998). While locus of control was a widely studied phenomenon in the 1970s and 80s, nearly all locus of control studies were conducted at the tertiary level. A single non-tertiary study of locus of control and cheating by Johnson and Gormly (1972) found that, among 113 American Grade Five students, greater externalizing was related to cheating, but only among female participants.

Self-efficacy is a control-related self-belief with consistently inverse associations to cheating at the secondary level (e.g. Finn & Frone, 2004; Lee, Bong, & Kim, 2014; Nora & Zhang, 2010), as well as at the tertiary level (e.g. Bing, Davison, Vitell et al., 2012; Elias, 2009). Bandura (1986) defines self-efficacy as "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances… concerned not with the skills one has but with judgments of what one can do with whatever skills one possesses" (p. 94), or, in other words, one's sense of being able to 'exercise control' over outcomes (Bandura, 1997). Cheating behavior implies, by contrast, a recognition that one cannot achieve the outcomes one desires, unless one breaks the rules. Zwagerman (2008) argues that when students feel apprehension and uncertainty over grades, in other words when they experience low self-efficacy, cheating "mitigates the randomness of the outcome—it eliminates the personal factor and puts the student more firmly in control" (p. 684). It stands to reason, therefore, that what students feel they cannot control legitimately, they may be enticed to cheat their way around.

Academic self-efficacy was found, in a study of 495 Grade Seven and Grade Eight students by Murdock et al. (2001), to be the most salient motivational predictor of cheating ($\beta = -.54$, $p < .01$). Two more recent studies of secondary cheating in South Korea (Bong, 2008; Lee et al., 2014) and one in Turkey (Tas & Tekkaya, 2010) found, similarly, that self-efficacy was a reliable predictor of cheating in regression analyses. Studies at the tertiary level suggest, however, that the influence of self-efficacy on cheating is complex. Ogilvie and Stewart (2010) divided 536 Australian university students into three groups based on their self-efficacy scores (low, moderate, and high). Low self-efficacy students' intention to plagiarize (ITP) was predicted exclusively by prior cheating; for moderate self-efficacy students, ITP was predicted exclusively by the perceived benefits of plagiarism; and for high self-efficacy students, ITP was predicted exclusively by the perceived shame associated with plagiarism. Jurdi et al. (2011a) found that, among 321 Canadian university students, academic self-efficacy interacted with 'instrumental motivations' for studying such as acquiring a job or degree. Instrumentally motivated students were more likely to report cheating, unless they also exhibited high self-efficacy. Self-efficacy served, the authors argue, like a "'protective factor' that interacted with instrumental motives to reduce the likelihood of academic dishonesty, by keeping instrumentally oriented students motivated to try hard" (p. 24).

Whether due to personality, self-beliefs, or interactions of the two, the degree to which cheating is likely under given circumstances clearly varies among individuals. Two elements of individual difference that emerge repeatedly in research on cheating are morality and control. Morality has been measured in terms of moral obligation to be honest, self-rated honesty, and moral identity. While recent experimental studies have found that individuals tend to curtail immoral behavior in order to maintain a positive moral self-concept, the ability to act morally is also largely a matter of self-control. More conscientious, less impulsive individuals generally exhibit higher levels of self-control, which may be depleted by cognitively demanding activities, resulting in a higher likelihood for immoral behavior in all but those with a strong sense of moral identity.

2.3.4 Learner characteristics

Academic motivational and behavioral tendencies, such as achievement goal orientations and approaches to learning, are important learner characteristics that students develop during the course of their educational careers, and that may, together, constitute an integrated "stance towards academic tasks" (Murphy & Alexander, 2000, p. 41). Similar to self-beliefs, learner characteristics both reflect past experience and shape future behavior. Pintrich (2000) argues that an individual's approaches to, and purposes for, learning might come to "reflect an organized system, theory, or schema" for learning (p. 94), such that over time they tend increasingly to be activated together. While such characteristics may be relatively stable across various contexts, they may also be modifiable by countervailing influences in "'strong' classroom contexts or experimental manipulations" (p. 102).

Achievement goal orientations. Achievement goal theory (AGT) holds that students' learning processes and outcomes can be influenced by whether their learning activities are motivated by mastery or performance goals (Meece, Anderman, & Anderman, 2006; Phan, 2009b). A mastery goal orientation is characterized by intrinsic aspirations such as "developing one's abilities, mastering a new skill, trying to accomplish something challenging, and trying to understand learning materials. Success is evaluated in terms of self-improvement, and students derive satisfaction from the inherent qualities of the task, such as its interest and challenge" (Meece et al., 2006, p. 490). A performance orientation involves, by contrast, extrinsic aspirations such as "demonstrating high ability relative to others, striving to be better than others, and using social comparison standards to make judgments of ability and performance. A sense of accomplishment is derived from doing better than others and surpassing normative performance standards" (Meece et al., 2006, p. 490). The intrinsic *vs.* extrinsic contrast referred to here is, in fact, one of several dyadic conceptions of motivational orientation, like task *vs.* ego-orientation (Maehr, 1983), that coalesced under the mastery *vs.* performance framework of modern achievement goal theory beginning in the late 1980s (Ames, 1992; Ames & Archer, 1988; Elliot, 2005; Marsh et al., 2003). While prior work had previously demonstrated that cheating is more likely among students who are motivated by performance-related variables such as the need for approval (e.g. Lobel & Levanon, 1988), achievement anxiety (Shelton & Hill, 1969), high *vs.* low achievement motivation (Johnson & Gormly, 1972), and grade-pressure (e.g. Smith, Ryan, & Diggins, 1972; Evans & Craig, 1990a; Diekhoff, LaBeff, Clark, et al., 1996), the first study of cheating from the perspective of AGT, by Anderman et al. (1998), proved seminal, inspiring a body of research that continues in current literature (e.g. Galloway, 2012).

Initial findings by Anderman et al. (1998) suggested that students with personal extrinsic orientations were significantly more likely to cheat than those with personal intrinsic orientations. Intrinsically oriented students were found, in fact, to be 1.54 times less likely to cheat than their extrinsically oriented counterparts. While research has generally supported this pattern of association between personal goal orientation and cheating, results have been somewhat inconsistent. At the secondary level, Murdock et al. (2001) found that while personal extrinsic goal orientation predicted cheating in logistic regression (β = .38, p < .05), the effect of a mastery goal orientation on cheating was nil (see also Rettinger & Cramer, 2009). A study of secondary students conducted by Stephens and Gehlbach (2007) found, by contrast, that a mastery orientation negatively predicted cheating, whereas a performance orientation exerted no significant effect. These inconsistencies also show up in relation to students' views on cheating. In response to an open-ended section on a questionnaire used by

Olafson et al. (2013), for instance, just one-third of non-cheaters identified mastery goals as a key reason why they did not cheat. A study by Koul, Clariana, Jitgarun, and Songsriwittaya (2009) found, by contrast, that stricter views on cheating tended to be held by performance-oriented students.

Performance and mastery goal orientations have more recently been subdivided into approach and avoidance orientation constructs, thus creating a four-factor structure, or '2 x 2' achievement goal model (Elliot & McGregor, 2001; Elliot & Murayama, 2008). Elliot, Murayama, and Pekrun (2011) have additionally formulated a '3 x 2' achievement goal model, by applying a trichotomous conception of the reference points used for measuring competence (task, self, and other) to the approach-avoidance dimension of the 2 x 2 model. Approach constructs involve proactively seeking either understanding in the case of mastery-approach, or favorable peer comparisons in the case of performance-approach. Avoidance constructs involve defensive measures to either avoid failing to learn, as in the case of mastery-avoidance, or to avoid negative peer comparisons, as in the case of performance-avoidance.

Findings as to how the four-factor structure relates to cheating remain, like the original two-factor structure, inconsistent. Bong (2008), who included measures for performance-approach, performance-avoidance, and mastery orientations (without distinguishing between mastery approach/avoidance) found, in a study of 753 South Korean high school students, that only performance-avoidance orientation predicted cheating (β = .25, p < .05). An experimental study of cheating that used the same measures among 70 American university students conducted by Niiya et al. (2008), found similar results: performance-avoidance was the only orientation that predicted cheating at a significant level, albeit exclusively among male participants. A multi-level study conducted at the secondary level by Tas and Tekkaya (2010) found, by contrast, that while a mastery orientation negatively predicted cheating at both the student and classroom levels (γ = -.366 and γ = -.397, respectively), a performance-approach orientation exerted a relatively weak effect at the student level (γ = .118), and none at the classroom level, whereas performance-avoidance goal orientation was not a significant predictor at either level.

The results of research into associations between cheating and achievement goal orientation suggest, in sum, that students are generally more likely to cheat when they pursue performance goals, and especially when they pursue performance-avoidance goals, instead of mastery goals. The inconsistencies noted above appear, however, to suggest that the influence exerted by achievement goal orientation on cheating behavior involves unrecognized complexity, as may arise from intra-psychic and/or contextual factors that mediate or modify the relationship.

Preferred learning strategies. Another major contribution made by Anderman et al. (1998) was the inclusion, for the first time in a study of cheating, of a measure of deep learning strategy usage, which turned out to be the strongest predictor of cheating in the study (β = -.86). While research into AGT proliferated in academic integrity literature following Anderman et al. (1998), research on learning strategy has since been conducted in just three works at the tertiary level (Jurdi et al., 2011a, 2011b; Norton, Tilley, Newstead, & Franklyn-Stokes, 2001), and one work at the secondary level (Bong, 2008).

Learning strategies are widely held to be behavioral complements to corresponding types of motivation (Tait & Entwistle, 1996). The measure that Anderman et al. (1998) included, for instance, entitled *Deep learning strategies*, served, in effect, as a complement to the measure for personal mastery orientation (r = .65). The motivational complement to deep strategies has been measured in other works as a 'deep motive' (Biggs, Kember, & Leung, 2001), which is conceptually and empirically similar to a mastery goal motivation. An orientation to mastery goals has been shown, additionally, to predict deep learning strategies in several structural equation modeling studies at the tertiary level (Fellonar, Román, & Cuestas, 2007; Phan, 2008, 2009a, 2009b).

When combined, deep learning motivation and deep learning strategy are often referred to as a 'deep approach to learning' (Biggs et al., 2001). Student learning theory, a prolific source of research on approaches to learning (Biggs & Tang, 2011), has generally focused on two approaches to learning: deep and surface. Students who take a deep approach to learning tend to be motivated by intrinsic interest, and generally seek to make personal meaning of what is to be learned. Students who adopt a surface approach tend, by contrast, to fear failure, and to prioritize the end results of learning, such as passing marks, over the actual substance of intellectual challenge and achievement. Surface strategies, such as rote learning and generally focusing on the superficial aspects of learning tasks (Biggs, 1987; Wilson & Fowler, 2005), are characteristic of acts described by Miller et al. (2011) as 'disintegrity', in that they violate not the rule, but the spirit, of academic integrity. Surface approaches to learning aim to counterfeit intellectual achievement with the production of end results, which effectively invalidates assessment results intended to reflect levels of understanding.

While learning strategies are, like achievement goal orientations, responsive to environmental variables, students may also develop relatively stable preferences for certain strategies due to past experience (e.g. Marton & Säljö, 1997; Wilson & Fowler, 2005). A semester-long field experiment involving a class of 180 Belgian university students conducted by Gijbels, Segers, and Struyf (2008) found, for instance, that participants' approaches to learning at the beginning of the semester tended to carry through to the end, despite the use of constructivist, deep-level assessment methods throughout the semester. While students did indicate recognizing the assessment methods as being geared for deep learning, those who preferred a surface approach initially, did not, on average, shift to deeper strategies by the end.

The only secondary-level study of cheating since Anderman et al. (1998) to incorporate measures of learning strategy was a structural equation modeling study conducted by Bong (2008). Both 'cognitive strategy use', and 'self-regulatory strategy use' were found to have significant negative correlations with self-reported cheating behavior (r = -.14 and -.29, respectively). Neither measure was included in the final model, however, due to multicollinearity between them, as well between each, respectively, and mastery goal orientation.

The association between learning strategy and cheating appears to be weaker at the tertiary level than at the secondary level. In a study of 267 university students in the UK, Norton et al. (2001) found, for instance, no correlation between self-reported cheating and approaches to learning. A more recent study by Jurdi et al. (2011a) found small but significant correlations between cheating and surface learning approaches ($r = .12$, $p < .05$) and deep learning approaches ($r = -.22$, $p < .01$). While these results appear to merit little interest in approaches to learning as variables related to cheating at the tertiary level, the aforementioned studies by Anderman et al. (1998) and Bong (2008) suggest stronger relationships among secondary students. The observed difference between the tertiary and secondary school students in how approaches to learning tend to relate to cheating might reflect either the use of different measures, or differences in the consistency and intensity of the classroom experience between the tertiary and secondary school level. Measures used by Norton et al. (2001) and Jurdi et al. (2011a) came from the *Approaches to Study Inventory* (ASI) (Richardson, 1990) and the *Study Process Questionnaire* (SPQ) (Biggs, 1987), respectively, which are both traditional 'approaches to learning' measures developed within the mainly tertiary-level literature of student learning theory, whereas Anderman et al. (1998) and Bong (2008) used measures designed for the secondary level that focused specifically on the use of learning strategies that are 'deep' in character.

Alternatively, students' academic behaviors may be more responsive to environmental factors at the secondary level because secondary classrooms tend to play more a salient role in their academic lives. In a qualitative study of learning strategies in Scotland, Selmes (1986) found strong evidence that secondary students frequently approach learning in a surface manner, and that surface approaches to learning tend to arise for the same reasons at the secondary level as they do at the tertiary level, while "possibly having stronger effects in secondary than in higher education" citing "formal assessment and teaching methods, [and] dependence on the teacher" as key contextual factors to inducing a surface approach (p. 25). The scarcity of research on learning strategies and their correlates at the secondary level is a notable lacuna in the academic integrity literature.

Achievement goal orientations and learning strategy preferences may, over time, develop into relatively stable tendencies that serve as students' default learning characteristics when countervailing contextual variables are weak. Research generally indicates that the pursuit of performance goals tends to increase cheating, whereas the pursuit of mastery goals tends to decrease cheating. The inconsistency with which goal orientations appear to predict cheating suggests, however, that the relationship may be complicated by additional factors. Learning strategies, which were also found by Anderman et al. (1998) to have strong associations to cheating at the secondary level, have received comparatively little attention in published research.

2.4 Situational variables

The second broad vein of research on academic integrity, related to situational variables, has traditionally been focused in two areas. The first of these areas emphasizes conditions that make cheating more or less risky, such as the arrangement of seating (e.g. Genereux & McLeod, 1995; Houston, 1986), use of multiple test copies (e.g. Hollinger & Lanza-Kaduce, 1996; Houston, 1983a), whether honor policies exist and are understood by students, the influence of peers, and the perceived risk of detection and severity of penalties (Covey, Saladin, & Killen, 1989; Dix, Emery, & Le, 2014; McCabe, Feghali, & Abdallah, 2008; McCabe & Treviño, 1993, 1997; Miller et al., 2011). Research in this area indicates that students are generally more likely to cheat when conditions make it easy to get away with (Houston, 1977; Tittle & Rowe, 1973; Whitley, 1998). Most of these variables were studied, however, principally at the tertiary level during the 1970s-1990s, and have low-to-moderate effect sizes (for a review, see Whitley, 1998). With the exception of peer influence, which continues to be actively investigated (e.g. Briggs et al., 2013; Nora & Zhang, 2010), contemporary interest in this line of research has shifted to assignment design (e.g. Briggs et al., 2013; Heckler et al., 2013), the use of anti-plagiarism software (e.g. Batane, 2010; Gannon-Leary et al., 2009; Heckler et al., 2013), and the use of devices to jam and disrupt illicit communication during examinations (Latova & Latov, 2008).

The second area of research on situational variables has focused on how situations influence students' motivational goals, and the evaluations of fairness and quality that students make of academic contexts. This literature indicates, overall, that students cheat when they are oriented to extrinsic achievement goals; and/or they feel *unprepared* to succeed honestly; and/or they feel alienated and *do not want to try* (e.g. Anderman et al., 1998; Baird, 1980; Brandes, 1986; Evans et al., 1993; Galloway, 2012; Latova & Latov, 2008; Schraw et al., 2007; Sheard, Markham, Dick, et al., 2003; Sisti, 2007; Whitley, 1998). The above reasons for cheating have been related to nearly identical sets of situational variables. Factors that hinder students' efforts to prepare for success include heavy workloads, difficult tasks, ineffective teachers, boring subject matter, and, if success is defined on a relative basis, the extent to which peers gain advantages by cheating. These same factors, in addition to perceiving a teacher as disrespectful or uncaring, are also implicated in student feelings of dissatisfaction and alienation (Anderman et al., 2010; Ashworth, Banister, & Thorne, 1997; Barnhardt & Ginns, 2014; Evans & Craig, 1990a; McCabe, 1999; Murdock, Miller, & Kohlhardt, 2004; Murdock et al., 2008; Zito & McQuillan, 2011). There seems to be a point, therefore, at which some students make a transition from feeling that they are simply unable to prepare adequately for a given class, to feeling alienated and victimized by aspects of the class that they blame for their lack of success. Students who feel alienated from a given class often still feel a pressing need for good grades, and may readily sacrifice the substance of learning by cheating for the grade-credentials that they see as necessary for longer-term success (Anderman et al., 1998).

A metaphor for the apparent balance between students' evaluations of the fairness and quality of a learning context, and the obligation they feel to be honest, is the 'teaching-learning contract' (Brent & Atkisson, 2011; Murdock et al., 2001; 2004). Contractarian judgment of the acceptability of cheating implies situated moral flexibility consistent with the notion of 'situation ethics' (Brent & Atkisson, 2011; McCabe & Katz, 2009). According to the situation ethics perspective, acts that are usually seen as being wrong become morally acceptable when the particular circumstances of a given situation create necessities that alter the rules (Crittenden et al., 2009a; Fletcher, 1966). Fluctuations in cheating behavior may, by this view, reflect the influence of situational variables on whether students perceive the rules that forbid cheating to be valid (Brent & Atkisson, 2011).

While some scholars have argued that the 'moral flexibility' inherent to situation ethics is new (McCabe & Katz, 2009), studies of adolescent ethical judgment suggest that situation ethics may be an enduring norm. A study of 1,800, 11-16 year-olds in the UK conducted by Thomson and Holland (2002) found that participants were grappling with multiple moral frameworks referenced to disparate cultural and moral touchstones such as peers, family, religion, and global youth culture. An instance of ethical judgment was found to depend largely upon which framework a participant referenced. These findings suggest that adolescent moral judgment is not composed of strict categories of right and wrong, but is varied, complex, and situation-sensitive. Keltikangas-Järvinen and Lindeman (1997) found, similarly, that among 2,594 Finnish secondary students, participants judged the immorality of acts such as fighting, lying, and theft on the basis of contextual criteria such as what the act was, how it was carried out, or the actors' rationales. Secondary students also appear to judge cheating in a context-specific light (Murdock et al., 2008). Schab (1980) reported, for instance, that scholastic cheaters tended to be judged leniently by their peers, because cheating was viewed as context-specific.

2.4.1 The nature and quality of teaching

Research on pedagogical factors related to cheating has generally focused on either motivational orientation variables, such as whether teachers are perceived to encourage mastery or performance goals, or on relational variables, such as the degree to which a teacher is perceived as good, or a subject is perceived as interesting. These two sets of variables differ fundamentally in that motivational goals reflect the objectives toward which learning efforts are directed, whereas relational variables reflect students' evaluations of the quality of learning contexts. Relational variables, also referred to as social motivation variables (Murdock et al., 2001), include the evaluations that students make of teachers, in terms of both whether they are skilled, and the quality of interpersonal connection they form with students. Students who think their teacher does a poor job may feel let down or disadvantaged, even if the teacher otherwise displays high levels of interpersonal caring. Likewise, low interpersonal caring may be damaging to teacher-learner relationships, even when the teacher is otherwise pedagogically competent (Murdock et al., 2004; 2008). While motivational orientation has been a subject of considerable interest in cheating literature since the seminal work on achievement motivation and cheating by Anderman et al. (1998), relational variables have been studied for longer (e.g. Hartshorne & May, 1928) and findings have generally been more consistent.

Motivational goal structures. Motivational goal structures have been studied in relation to cheating behavior in secondary students since at least Mills (1958), who found that extrinsic structures, such as tangible reward systems, increased the likelihood that Grade 6 students would cheat in experimental settings. In modern achievement goal theory, the concept of 'classroom goal structure' emphasizes the situated nature of motivation, whereby students are held to adopt personal goal orientations in response to the goals that they see emphasized in class (Ames & Archer, 1988; Meece et al., 2006), through at least six dimensions of pedagogical practice: task, authority, recognition, grouping, evaluation and task (Ames, 1992; Meece et al., 2006). A mastery classroom structure emphasizes the importance of intrinsic goals such as aiming to develop genuine, masterful competence of course material, whereas a performance classroom structure emphasizes, by contrast, extrinsic goals such as achieving high grades and favorable peer comparisons (Ames & Archer, 1988; Anderman & Midgley, 1997).

Anderman et al. (1998) administered self-report measures of achievement goal structure pertaining to both school and classroom levels to 285 American Sixth, Seventh, and Eighth Grade students. Their use of hierarchical logistic regression determined that while cheating behavior was predicted by both school-level performance structure and classroom-level extrinsic structure, only the latter predicted the belief that cheating was acceptable. Firstly, these findings suggest that "if the incentive value of the reward is more important to the adolescent than the academic task itself, then the student may see cheating as acceptable" (p. 89). This proposition is supported by several earlier experimental studies that found cheating behavior in children to relate positively to extrinsic motivation (Lobel & Levanon, 1988; Mills, 1958) as well as knowledge of peer performance (Shelton & Hill, 1969; Taylor & Lewit, 1966). Secondly, these findings point to the primacy of classroom contexts over larger school contexts in affecting students' judgments of the acceptability of cheating.

In contrast to Anderman et al. (1998), subsequent studies have indicated that classroom mastery structure has a stronger effect on cheating at the secondary school level than performance goal structure (Anderman & Midgley, 2004; Murdock et al., 2001; Tas & Tekkaya, 2010; Stephens & Gehlbach, 2007). Exceptions to this pattern have, however, also been found. In a structural equation modeling study conducted by Bong (2008), neither mastery nor performance-approach goal structures exerted a direct effect on cheating behavior. Anderman et al. (2010) found, similarly, that among 583 American high school students, cheating behavior was related negatively to both extrinsic and mastery classroom structures, which also shared a significant positive correlation with one another (r = .55). By examining this relationship among students who reported cheating at a 'moderate' level *versus* an 'extensive' level in logistic regression, however, Anderman et al. (2010) found that mastery goal structure was a strong negative predictor of cheating among extensive cheaters (β = -.60, p < .01).

Relational variables. A series of three studies of cheating among secondary students conducted by Murdock and colleagues (Murdock et al., 2001, 2004) has indicated that relational variables such as 'teacher competence' and 'interpersonal caring' exert an influence on cheating behavior that is highly interrelated with that of classroom mastery structure, albeit with greater strength and consistency. This series of studies has also traced, as such, a progression away from achievement goal structure to a renewed emphasis on relational variables across the field (e.g. Anderman et al., 2010; Day et al., 2011; Murdock & Anderman, 2006; Murdock, Miller & Goetzinger, 2007). In the first study to include both classroom goal structure and relational variables, Murdock et al. (2001) found that, among 495 American Seventh and Eighth Grade students, both Teacher commitment and Mastery goal structure were strong predictors of self-reported cheating. These two contextual variables were, however, also strongly correlated with each other (.70), which confounded efforts to distinguish their respective effects.

In the second of these studies, involving 204 American Ninth and Tenth-Grade students, Murdock et al. (2004) used vignettes to isolate the unique effects of relational variables and classroom goal structure. A follow-up questionnaire queried (1) participants' personal goal orientations and academic self-efficacy, (2) whether they thought the student or the teacher was more to blame for cheating, and (3) the degree to which they judged cheating to be (A) likely and (B) acceptable in the vignette scenario. The 'acceptability of cheating' scale used in this study was adapted from Anderman et al. (1998). The measure separated, however, into two statistically coherent measures that Murdock et al. (2004) named 'morality' (three items) and 'justifiability' (four items). Items on the morality scale were described as pertaining to the "absolute acceptability of cheating", whereas items on the justifiability scale were described as pertaining to whether cheating is acceptable under certain circumstances. The study found that when pedagogy was perceived as poor, the likelihood of cheating was high regardless of classroom goal structure. Blame was a function of both pedagogical competence and goal structure. Participants assigned equal amounts of blame to students and teachers in learning situations that were portrayed as either poor in pedagogy or having a performance goal structure. When pedagogy was portrayed as competent or a learning situation was portrayed as mastery-oriented, however, participants tended to shift blame from the teacher back to the student.

In the third study (Murdock et al., 2004), measures of perceived classroom goal structure were replaced with a measure of pedagogical caring. Vignettes were modified accordingly. The resulting regression model, which included blame as a mediator, explained 18% of variance in morality, and 39% of the variance in justifiability. In both studies reported by Murdock et al. (2004), pedagogical competence was the most consistent predictor of the morality, justifiability, and likelihood of cheating. The situated nature of cheating was also suggested by a group of large correlations in both studies between the likelihood of cheating, justifiability of cheating, and blame assigned to the teacher. Across both studies these factors correlated within a range of .69-.73, suggesting that cheating is more likely when it is more justifiable and the teacher is to blame. The vignette methodology used by Murdock et al. (2004) has been replicated at the tertiary level by Day et al. (2011) and Murdock et al. (2007). Both studies suggest that students' reactions to elements of teacher performance are largely uniform across secondary and tertiary contexts (see also Schraw et al., 2007).

When students feel beset by the urge to cheat for grades that will characterize their ability and shape their future, an effective teacher may strengthen their resolve to be honest by helping them feel prepared for success. In a study of 583 American high school students, Anderman et al. (2010) found that for every one-unit improvement in 'teacher credibility', defined as the degree of trustworthiness, competence, and caring, cheating among 'extensive cheaters' became 33% less likely. Cheating appears, by contrast, to become more likely when students feel unsupported by a teacher (Evans & Craig, 1990a, 1990b; Davis et al., 1992; Shipley, 2009; Steininger, Johnson, & Kirts, 1964). An international study of secondary students conducted by Evans et al. (1993) found that 85% of respondents agreed that cheating is more likely when a teacher is disorganized and difficult to understand. When teachers come across as disorganized and unclear, students may cheat as a result of feeling abandoned in their effort to grapple with course material.

Teachers appear to influence cheating both by the type of motivational goals they emphasize, and the levels of interpersonal caring and pedagogical skill they demonstrate to students. Research reviewed below suggests, however, that secondary students tend to hold expectations for fairness and quality that go beyond broad notions of pedagogical skill and interpersonal caring, to a more nuanced array of dimensions of learner experience. These dimensions include assessment (e.g. Sisti, 2007), interestingness (e.g. Rowe & Hill, 1998; Schraw et al., 2007), fairness and consistency of rules (e.g. McCabe et al., 2008; Thornberg, 2008), and appropriateness of workload (e.g. Galloway, 2012; Evans et al., 1993). While the practice and persona of the teacher greatly informs each of these dimensions in a given class, they have also been investigated in the literature of cheating as distinct and generalizable variables.

2.4.2 Interest

Interest, defined by the *New Oxford American Dictionary* as "the state of wanting to know or learn about something or someone", is a core motivation for engagement in meaningful learning (Schraw et al., 2007), in that it connotes genuine appreciation for the value of a particular skill or body of knowledge. An interest in learning is, by this definition, inherently antithetical to pretending to learn for the sake of a grade, i.e. disintegrity (Miller et al., 2011). Results of the only major published study to focus expressly on interest and cheating (Schraw et al., 2007) were consistent with this definitional relationship. Based on a framework developed by Renninger, Hidi, and Krapp (1992) (see also Hidi & Renninger, 2006), Schraw et al. (2007) distinguished between *personal interest*, which refers to topics or subject areas upon which an individual places relatively high personal value over time, and *situational interest*, which refers to the information that an individual values temporarily within a specific context. Schraw et al.'s (2007) findings suggest that while personal interest in a class or topic appears to have a stronger overall influence on cheating behavior than situational interest, the latter was also more strongly related to other situational factors that affect cheating independently, such as the perceived effectiveness of the teacher. Situational interest may, therefore, act as a moderator or mediator for the effects of situational variables on cheating.

While negative associations between cheating and interest, *per se*, have been noted in two additional secondary-level studies (Ma, Lu, Turner, & Wan, 2007; Sisti, 2007), a number of concepts that appear to connote *disinterest*, or the state of *not* wanting to know or learn about something, have also been found to relate to cheating in a positive manner. These include, at the secondary level, "lack of clarity about the reasons or purposes of learning" (Evans & Craig, 1990a, p. 334); the teacher being perceived as boring (Evans et al., 1993); and under-engagement (Stephens & Gehlbach, 2007). In a study involving 463 American secondary students, Evans and Craig (1990a) found, for instance, that students were more likely to cheat in required courses than in elective ones. Studies at the tertiary-level also report positive associations between cheating and variables connoting disinterest, such as "irrelevant and boring course material" (Baird, 1980, p. 517), "a devalued sense of the worth of education" (Diekhoff et al., 1996), "trivial, uninteresting assignments" (McCabe, Treviño & Butterfield, 2002), and "assignments… having little learning value" (McCabe & Katz, 2009, p. 17). While no study of disinterest, *per se*, has been reported in the literature, the empirical consistency between the abovementioned observations that cheating is negatively related to interest and positively related to variables that appear to connote disinterest suggests, in sum, that students tend to cheat less on tasks that they value more.

2.4.3 Assessment

A relatively small amount of research on assessment and cheating at the secondary level indicates that cheating is more likely when assessments are perceived as higher stakes (Nichols & Berliner, 2007), less personally relevant (Sisti, 2007), and less fair (Evans & Craig, 1990a, 1990b). Cheating also tends to be more endemic to certain types of assessment, such as homework, than to others (Galloway, 2012; Jensen et al., 2002). Factors such as stakes, relevance, fairness, and type suggest that the assessment regime in a given class should be viewed as a multidimensional context in its own right (Dorman & Knightley, 2006), wherein factors from the broader learning environment that are related to students' levels of alienation, preparedness, and anxiety over grade performance become sharply focused.

Taking realistic account of the importance that students ascribe to grades under most circumstances, and the levels of worry and competition that grades can engender, it stands to reason that cheating will be more prevalent when the stakes are high and the risk of detection is low (Nichols & Berliner, 2007; Schraw et al., 2007; Vitro & Schoer, 1972). The connection between high stakes and cheating is intuitive; as the stakes increase, so do both the consequences of failure and the benefits of cheating (Lee et al., 2014; Nichols & Berliner, 2007). Jensen et al. (2002) found, for instance, that high school students were more likely to sympathize with cheating in situations where the consequences of failure were high (see also Bracey, 2005; Lavelle, 2008; Levitt & Dunbar, 2005; Sheard et al., 2003).

While the risk of detection that students associate with cheating on a particular assignment may be contingent largely upon their perceptions of the teacher's vigilance, it also depends on type of assignment in question. Cheating appears, for instance, to be more common on assignments that are, by their nature, more difficult to monitor, such as homework and group work (Briggs et al., 2013). Two studies at the secondary level that used self-report inventories of cheating behavior found that the most common form of cheating was homework copying (Galloway, 2012; Jensen et al., 2002). At the tertiary level, Hudd, Apgar, Bronson, and Lee (2009) report finding that some participants held the instructor personally responsible for cheating on out-of-class assignments, on the grounds that he or she had failed to take appropriate steps to address an obvious problem. A study of 456 Canadian university professors by Leonard and LeBrasseur (2008) suggests that many professors are, indeed, aware of this problem, and yet choose to give such assignments anyway. A number of professors who took part in this study realized that cheating on individual homework assignments was widespread, but continued to give such assignments, "to ensure students obtain learning benefits" (p. 37).

While some students undoubtedly benefit from group work and from individual homework assignments, the practice of repeatedly providing low-risk opportunities to cheat in a given class is likely to erode integrity over time. The obvious opportunity to cheat on tests in an experimentally low-risk environment led to a pronounced 'contagion' effect over the course of a semester in a tertiary-level study conducted by Walker, Wiemeler, Procyk, and Knake (1966). The rate of cheating was 23% on the first test, but increased subsequently to 64% on the second test, and 86% on the third test. Students who cheated on the first test appeared to transmit information about the ease of cheating to their classmates, who then also cheated. Qualitative studies at the secondary level have noted, similarly, that students often pass information along to their peers about how easy it is to cheat on specific assessments and/or in particular classes (Schraw et al., 2007), and to cheat more when they perceive lower risks (Ma et al., 2007).

Another dimension of assessment design that has been related to cheating is *authenticity*, or "the extent to which assessment tasks feature real-life situations that are relevant to the learner" (Dorman & Knightley, 2006, p. 56). Sisti (2007) argued that secondary-level teachers could thwart plagiarism by designing writing assignments of a more creative and authentic nature that cannot be readily copied and pasted or purchased on-line. "High school teachers should seek to craft assignments that are not simply rote research tasks but rather encourage and engender a sense of student ownership of the resulting product" (p. 227). At the tertiary level, a number of scholars from United Kingdom universities have formed the Joint Information Systems Committee (JISC), which seeks, among other things, to help educators 'design out' cheating with more authentic types of assessment that require higher levels of individual creativity and analysis, and by integrating assessments with one another across the semester (e.g. Carroll & Appleton, 2001; Evans, 2006; Gallant, Anderson, & Killoran, 2013; Gannon-Leary, Trayhurn, & Home, 2009). Assignments that require students to formulate and explain opinions were, for instance, found to be the least plagiarized of three assessment types examined in a recent study involving 2,826 American university students (Heckler et al., 2013).

Cheating also appears to be more likely on assessments that students believe to be less fair (e.g. Baird, 1980; Genereux & McLeod, 1995; McCabe, 1992). The perceived fairness of assessments is likely to reflect, to some extent, students' perceptions of the fairness of a learning environment. Fairness has, however, been conceptualized specifically in relation to assessment in the cheating literature in terms of general exam difficulty (Welsh, 1993), the harshness of grading practices (Evans & Craig, 1990b; Vowell & Chen, 2004), and whether tested material has been covered in the class (Evans & Craig, 1990a). The perception that a teacher's tests are unfair may lead to cheating both by amplifying students' concerns that they may be under-prepared, and by alienating them due to what appears to be injustice (Thorkildsen, Nolen, & Fournier, 1994; Thornberg, 2008). While research on assessment fairness is sparse, available findings uniformly suggest a negative association with cheating.

Online assessment. Assessment design is of especial interest in relation to online education (Arnold, 2012; Harmon & Lambrinos, 2008), where disagreement has emerged over whether cheating takes on significantly more worrisome dimensions and proportions due to the plethora of resources online that facilitate cooperative cheating and plagiarism (Anderman, Freeman & Meuller, 2007; Batane, 2010; Briggs et al., 2013; McCabe, 2005; Walker, 2010), or whether the Internet is "at most a complication in a long-standing dynamic" (Howard & Davies, 2009, p. 65; also Trushell & Byrne, 2013; Trushell, Byrne & Hassan, 2013; Watson & Sottile, 2010; Williams, 2008). King and Case (2014) collected data on cheating in online educational contexts from 1,817 undergraduate business students at a US university over a five-year period (2009 – 2013), and found that between 34% and 44% reported at least one act of cheating during a given year. Watson and Sottile (2010) report similar incidence of cheating in online classes among a group of 635 American undergraduate and graduate students (33%), which actually compares favorably to rates of self-reported cheating observed in studies of American university students in face-to-face classes during the same period, such as 60% in Stone, Kisamore, Jawahar and Bolin (2014) and 65% in Miller et al. (2011).

A number of scholars see the loss of direct control over assessment processes in online settings as remediable through adapted teaching practice such as personalized writing assignments (Gallant, 2008; Harmon & Lambrinos, 2008; Heckler et al., 2013), online synchronous assessment (Chao, Hung & Chen, 2012; Moton, Fitterer, Brazier, et al., 2013), and student authentication processes such as videos and facial snapshots (McNabb, 2010). Chao et al. (2012) argue for the viability of online synchronous assessments (OSAs), in which multiple students are assessed simultaneously while being monitored by video. The potential for cheating on OSAs may be further reduced by oral assessment methods, or by taking remote control of students' computers for the duration of the assessment (Chao et al. 2012).

The fact that assessment is the venue for cheating in academic environments makes it of special concern to educators and researchers with an interest in academic integrity. Assessment appears to be a multifaceted sub-context across the spectrum of online and face-to-face formal learning environments, wherein grade worries, competitive urges, and concerns over the purposes, quality, and justice of schooling all merge together. The present review of literature suggests that cheating is more likely on assessments that students perceive to be high stakes, inauthentic, unfair, poorly designed, and poorly monitored.

2.4.4 Workload

While heavy workload is a common reason given by students for why they cheat in qualitative studies (e.g. Evans & Craig, 1990a; Galloway, 2012; Sisti, 2007; Zito & McQuillan, 2011), in quantitative studies this factor has demonstrated only moderate statistical associations with cheating behavior (Jurdi et al., 2011a; Smith et al., 1972). At least three possible explanations for how workload affects cheating emerge from the literature. Firstly, a heavy workload may predict cheating in terms of preparation, by making it more difficult for students to finish their work without cutting corners. Secondly, it may lead to a higher likelihood of cheating by depleting psychological resources that individuals need to make accurate moral judgments and to exert self-control. Thirdly, it may predict cheating as a proxy measure for student commitment to a class, where less committed students are more likely to perceive workload as inappropriate, and *vice versa*. Evidence and argument for each of these three possible explanations is reviewed in the given order, below.

Secondary-level students often report that heavy workloads overwhelm their ability to prepare to succeed honestly (Evans & Craig, 1990a; Galloway, 2012; Sisti, 2007; Zito & McQuillan, 2011). The experience of lacking time both to finish homework assignments and to prepare for in-class assessments, which is familiar to most students past and present, is well expressed in a statement from a high school respondent interviewed by Galloway (2012): "It's 1 am; I have just finished 3 hours straight of a calc problem set, Spanish vocab work and history reading and I still have to write an English essay. I can turn in nothing and get a '0' or I can download something from the Internet and take my chances" (p. 392). Students who cannot complete assignments on time are expected, nonetheless, to control the urge to cheat. Depending on teachers' policies regarding late work, this may entail having to accept failing grades.

Students may lack the self-control necessary to accept tough consequences for assignments that, due to sheer overwork or poor planning, they fail to complete on time. Low self-control may reflect immaturity, or cognitive fatigue. Self-control has, for instance, been found to require cognitive resources that can become depleted when individuals are genuinely rushed, overloaded and sleep-deprived (Greene et al., 2008). Strenuous workloads may weaken resistance to the temptation to cheat by exhausting the psychological resources that students need to exert self-control (Gino et al., 2011). In a pair of experimental studies among Israeli undergraduate students, Shalvi, Elder, and Bereby-Meyer (2012) demonstrated, for instance, that increasing the time pressure in a dice-rolling activity triggered dishonest behavior. Cheating is also observed to be significantly more common in people who are sleep-deprived (Barnes, Schaubroek, Huth, & Ghumman, 2011), and cognitively overloaded (Gino et al., 2011; Greene, Morelli, Lowenberg, et al., 2008). Heavy workloads may, therefore, overwhelm students' ability both to recognize the immorality of cheating, and to control the urge to cheat when opportunities arise.

Evaluations of what amount of work is appropriate in a given class may also reflect student commitment. Just as a lack of perseverance and effort have been implicated as causes of cheating behavior in several secondary-level studies (Evans & Craig, 1990a; Hamlen, 2012; Latova & Latov, 2008), the 'appropriateness' of an amount of work in a student's mind is, by definition, a function of the amount of time and effort they believe the work merits. A small amount of work that has been taught poorly or assessed unfairly might feel overwhelming to a student who tends to struggle, or who feels uninterested in a particular subject. Evans and Craig (1990a) found, for instance, that students tended to attribute cheating to workload that was heavy in terms of both amount and difficulty. Work that is difficult and time-consuming may, nevertheless, be perceived as worthwhile to students who feel more interested, capable, and connected to a class, i.e. more committed (Kember, 2004). Student commitment entails, as such, how much work students feel internally obliged to do, or personally responsible for doing, in a given class. What students deem to be their personal responsibility is, indeed, not defined externally so much as internally (Schlenker, 1997). Several recent secondary studies of cheating have reported, for instance, that students described workload in terms of "feeling" too heavy (Sisto, 2007; Zito & McQuillan, 2011). Students who are less committed to a class feel less obligation to work hard at it (Curry, 1984), which may additionally imply low moral obligation to accomplish tasks honestly.

Conceptualizing the appropriateness of workload is complicated by whether it reflects objective reality or subject perceptions. Heavy workload, as objective reality, appears likely to predict cheating as a function of cognitive depletion, which breaks down students' self-control to resist the urge to cheat. As a subjective perception, a student's perception that the workload is too heavy in a given class may reflect his or her self-perceived ability and resulting performance anxiety, or commitment to the class.

2.4.5 Peer influence

In addition to the behavioral norms that individuals perceive as characteristic of the socio-cultural groups with which they identify, distinct behavioral norms also arise in specialized contexts, such as academic classes (Carson, 2013). Context-emergent peer norms have been identified as a major influence on individual cheating behavior in secondary and tertiary settings in America (e.g. Bowers, 1964; Galloway, 2012; Hartshorne & May, 1928), as well as in a variety of other national settings (e.g. Eisenberg, 2004; Latova & Latov, 2008; McCabe et al., 2008; Nora & Zhang, 2010; Teodorescu & Andrei, 2009). The general finding that students cheat more when they perceive that their peers cheat more suggests at least two possibilities: (1) individuals may be influenced to cheat by knowledge of their peers' cheating behavior, and/or (2) individuals and their peers may be simultaneously influenced to cheat by the same environmental factors. In the first instance, believing that one's peers cheat successfully in a given class may lead to perceptions that the risk of detection is lower and competition is more ruthless. Such conditions may suggest that 'good guys finish last', therefore cheating is acceptable, or even necessary. In the second instance, inasmuch as perceptions and opinions are social constructs (Bandura, 1977, 1986), peer norms may mediate the effects of environmental variables on individual cheating behavior. In small-scale class settings, students may co-construct opinions and perceptions of factors such as teacher quality, assessment quality, and usefulness of learning that, together or separately, influence their evaluations of the overall quality of the class. Such evaluations at the individual level would, therefore, be at least partially mediated by group perceptions.

Research shows not only that knowledge of peer cheating is a strong inducement to cheat (Burrus et al., 2007; Carrell, Malmstrom, & West, 2008; McCabe & Treviño, 1997), but also that students tend to over-estimate the amount of cheating their peers engage in (Engler et al., 2008; Jordan, 2001; Rettinger & Kramer, 2009; Shipley, 2009), and tend to condone such behavior by refusing to report it. In what may be the first questionnaire-based study of cheating, Barnes (1904) found that, among a sample of 125 American undergraduate students, most said they would not report a peer for cheating because they regarded 'tale-bearing' as 'contemptible'. This finding has been common in subsequent literature at both levels. Schab (1991) reports that in three studies over the course of thirty years (1969, 1979, 1989), the proportion of American high school respondents willing to report a friend for cheating fell form 12% to 8% to 4%. This finding does not always hold. Simon, Carr, McCullough, et al. (2004) found that 36% of 172 American chemistry undergraduate students were, in fact, willing to report cheating. In a similar university sample, Shipley (2009) found, however, that just 4% of 228 American university students admitted to ever having actually reported cheating to an authority figure. Similar results have also been found in other national contexts. While 20% of a sample of 1,119 undergraduate medical students in Ethiopia indicated willingness to report cheating (Desalegn & Berhan, 2014), only 2% indicated willingness to report cheating in a sample of Singaporean undergraduates (Lim & See, 2001), and 6% in a sample of Lebanese high school students (Bacha et al., 2012). Approval of 'tattletales' in a sample of Russian high school students was reportedly negligible (Latova & Latov, 2008). In view of the strong influence that peer norms exert on individual cheating behavior, the unwillingness to turn in cheaters to school authorities presents a dilemma: the source of harmful influence, namely the knowledge of peer cheating, is protected by those who are harmed.

The harm done by peer cheating may affect more than behavior. In several studies, students who recognized widespread cheating seemed less likely to recognize it as immoral. A study at the tertiary level by Harding, Mayhew, Finelli et al. (2007) found, for instance, that measures of 'moral obligation' and 'perceived social norms' collapsed into a single higher-order factor that turned out to be a strong predictor of the intention to cheat on tests ($\beta = .66$). O'Rourke et al. (2010) observed, also at the college level, that cheating actually became less immoral in the judgment of students who witnessed others cheat. McCabe and Katz (2009) argued, similarly, that peer influence appears to constitute a key source of 'moral flexibility' among students. These results suggest that ethics and peer norms are mutually reinforcing, albeit perhaps less stable in smaller-scale, temporary contexts such as academic classes, than in larger-scale socio-cultural contexts.

The amount by which one perceives his or her peers to cheat may also exacerbate the negative effects of competition in a given learning context. The immorality of cheating has often been traced to the unfair advantage it confers to cheaters, in terms of better grades (e.g. West, Ravenscroft, & Schrader, 2004). Being honest in a performance-oriented class that emphasizes grades, competition, and social comparisons, and where many other students cheat successfully, may strike students as likely to confer an unfair *dis*advantage. Where competitive pressures are already high, the perception that cheating is commonplace may lead to a "cheat or be cheated" mentality (Galloway, 2012, pp. 393-394; also Schwieren, & Weichselbaumer, 2010), even among individuals who simultaneously recognize that cheating is proscribed by norms in the broader socio-cultural groups to which they belong.

Social comparison theory (Festinger, 1954), first introduced as a framework for understanding peer influence on cheating behavior by Broeckelman-Post (2008; also Koul et al., 2009; Nora & Zhang, 2010), holds that individuals formulate and validate their opinions within the context of what their peers think. Individual perceptions of contextual factors, including perceptions of what behaviors are normal, are, by this view, constructed in concert with peer perceptions. Similar positions are supported with regard to peer norms in the literature on independence and conformity (Asch, 1956) and social learning (Bandura, 1977, 1986). Inasmuch as individual-level cheating behavior is influenced by environmental factors such as teachers' pedagogical skill, for example, the fact that cheating occurs at all in a given class may indicate to classmates that cheating is justifiable because the teacher is poor. A 2-level hierarchical linear model tested by Murdock et al. (2008) determined, for instance, that the amount of blame secondary students assign to teachers for the cheating that goes on in their classes was predicted by both individual and classroom aggregate perceptions of the teacher's pedagogy. Individual students' perceptions of their teachers were, in other words, generally consistent with their peers' perceptions, in relation to cheating. Such broad consensus suggests that underlying social mechanisms shape norms related to both cheating behavior and its justifications. Peer norms may, therefore, mediate the effects that context variables have on individual-level cheating behaviors. In a study involving 1,025 Romanian university students, Teodorescu and Andrei (2009) found that the effect of 'quality and relevance of instruction' on the intent to cheat fell from $\beta = -.60$ to $\beta = -.45$ when peer influence was added to the model, suggesting partial mediation. An elaborate theoretical model developed by Whitley (1998), based on Ajzen's (1991) theory of planned behavior, also positioned subjective peer norms as a potential mediator between alienation and the intention to cheat. This particular aspect of Whitley's (1998) model has not, however, been tested in subsequent research.

In determining whether cheating behavior is acceptable in a given class context, students appear to take important cues from the prevalence of cheating among their peers. Such determinations appear able to affect not only cheating behavior, but also judgments of whether cheating is immoral. Such perceptions are, moreover, often exaggerated, which potentially amplifies their influence. Perceiving a high prevalence of cheating may suggest to students that contextual factors justify it. Peer norms that support cheating may sensitize students to arguments that, for example, cheating makes up for a teacher's pedagogical ineptitude, unfair assessment practices, or boring class. A high prevalence of cheating may be felt as a kind of democratic affirmation that a teacher has failed and students are, therefore, relieved of their moral obligation to be honest.

2.5 Rational-cognitive models of deviance

As anticipated by the foregoing review, person and situation variables have been studied both independently and as components of integrated models of cheating. Numerous models for cheating have been introduced since the 1970s that integrate person and situation variables under the assumption that cheating is a fundamentally rational-cognitive phenomenon (e.g. DeVries & Ajzen, 1971; Lau, Yuen, & Park, 2013; Smith, Davy, & Easterling, 2011; Treviño, 1986). These models have been successful at predicting cheating-related cognitions such as the intention to cheat (e.g. Mayhew et al., 2009), but markedly less successful at predicting self-reported cheating behavior (e.g. Bong, 2008; Harding et al., 2012). This apparent shortcoming seems to reflect broadly the disjunction between cognition and action that has long been recognized in research on moral reasoning (Blasi, 1980). Several scholars have, for this reason, recently suggested non-rational explanations for academic cheating, such as automaticity (Harding et al., 2012), emotion and intuition (McTernan et al., 2014; Murdock et al., 2008; O'Rourke et al., 2010), and social contracts (Brent & Atkisson, 2011; McTernan et al., 2014; Murdock et al., 2004; Rettinger, 2007). These perspectives are consistent with more recent experimental evidence that contradicts the belief that moral decision-making is principally a function of rational cognition (e.g. Brüne, Juckel, & Enzil, 2013; Cushman et al., 2010; Gneezy, 2005; Haidt, 2007; Peters, Västfjäll, Gärling, & Slovic, 2006). However, no models for academic cheating have been developed that reflect non-rational processes hypothesized to underlie moral judgment.

The theory of reasoned action (TRA) appears to have been the first framework to integrate personological variables with perceptions of context in a structural model of cheating psychology (DeVries & Ajzen, 1971). TRA was a purely cognitive model that departed from the personality and risk/reward-oriented research characteristic of integrity literature in the 1960s-70s. TRA sought to describe individual differences in terms of beliefs, attitudes and perceived norms that led, in turn, to intentions and thus to behaviors. Moderate success in several early studies of the TRA model for academic cheating (e.g. Pratt & McLaughlin, 1989) prompted the addition of 'perceived behavioral controls', which measured, in essence, the perceived risk of detection and severity of penalties (Ajzen, 1991; Beck & Ajzen, 1991). The modified framework was named the theory of planned behavior (TPB).

A longitudinal study of tertiary student cheating by Beck and Ajzen (1991) found that the TPB model predicted most of the variance in the intention to cheat, but less than half in actual cheating behavior. Attempts to improve the TPB model in subsequent studies by adding variables such as 'moral reasoning level' (Harding et al., 2007), 'professional unethical beliefs toward cheating' (Hsiao & Yang, 2011), and prior cheating behavior (Mayhew et al., 2009), have generated results that remain similar to those of Beck and Ajzen (1991); they predict substantially more variance in cheating cognition (intention to cheat) than in actual cheating behavior. Of these three additions to the TPB framework, however, prior cheating behavior, a non-cognitive factor, has been a substantially stronger predictor of cheating behavior than either moral reasoning or moral obligation not to cheat. This observation led Harding et al. (2012) to speculate that cheating might be better characterized as an automatic habit, than as an outcome of deliberate, rational cognitive processes.

Rational-cognitive models for cheating behavior have also been developed to reflect motivational perspectives. A framework based on motivational variables and Kohlbergian moral cognition was proposed by Newstead et al. (1996), based on a study of the reasons for cheating among tertiary students in the UK. A similar framework has emerged more recently from a synthesis of research on cheating (Murdock & Anderman, 2006), which hypothesizes that the effects of individual and contextual variables on the propensity to cheat, are mediated by three motivational questions: (1) *What is my purpose?* (2) *Can I do this?* and (3) *What are the costs?* Murdock and various colleagues (2001, 2004) have additionally tested integrated models for cheating and cheating-related cognitions, such as the justifiability and morality of cheating, among American secondary students. These were the first models found to incorporate both relational variables, such as teacher quality, and motivational goal constructs, such as classroom goal structure, together with personal variables, such as academic self-efficacy, personal goal orientation, and grade-level. The hierarchical logistic regression model tested by Murdock et al. (2001), which included achievement goal orientation and structure, teacher quality, and social aspects of respondents' experience of school, was able to identify 48% of cheaters and 90% of non-cheaters. A regression model tested by Murdock et al. (2004), which included a similar group of factors, but excluded perceived classroom goal structures, predicted as much as 40 – 42% of the variance in cheating cognitions. No secondary study has yet explained more than 50% of variance in cheating behavior.

The most orthodox rational-cognitive models for cheating are rooted in econometric frameworks based on the work of economic utility theorists such as Becker (1968) and Simon (1982). Such models assume a cost/benefit basis for individual behavioral choices, and make little or no allowance for moral considerations (e.g. Bisping et al., 2008; Burrus et al., 2007; Magnus et al., 2002; Schwieren & Weichselbaumer, 2010). Econometric models of cheating are predicated on the view that humans are an inherently utility-maximizing and cost-minimizing species, or '*homo economicus*', whose behavior is a matter of econometric probability, modeled by 'probit' models. Probit models examine the extent to which the probability of cheating reflects perceived benefits of cheating, suggested by variables such as current GPA and perceived competitive pressure, *versus* the perceived risks, suggested by variables such as the perceived likelihood of detection and severity of punishment. Such models may also typically include demographic factors such as age and gender, as well as contextual factors, such as the prevalence of cheating among peers.

The probit models tested by Bisping et al. (2008) indicated that background factors such as gender, year in school, and age related differently to different forms of academic misconduct. These differences appeared, moreover, to depend largely upon whether the students recognized that various acts were, in fact, misconduct. Students who fail to realize that their instructor views a particular behavior as cheating cannot, the authors argue, correctly assess the risks associated with those behaviors. Instructors should, therefore, be clearer with students about what acts constitute cheating, and should elevate the level of risk that students associated with it. This recommendation echoes policy initiatives advised many times in the literature, as exemplified by honor policy research (e.g. McCabe et al., 2002, 2003). The effect of such policies appears, however, to be moderate, at best (Evans & Craig, 1990a;

Passow, Mayhew, Finelli, et al., 2006; Whitley, 1998). To wit, driving up the risks associated with cheating appears to explain some, but far from all, of its variance.

While rational-cognitive models tested in econometric studies generate helpful guidance for how various factors alter the probability of cheating, the *Homo economicus* view of human nature has been challenged in recent psychological literature by growing evidence that people often forego their own interests for the sake of higher moral principles, such as peace and justice (Haidt, 2001). Haidt (2007) refers to this as the '*homo moralis*' view of human behavior (p. 998), while other scholars have asserted that humans have fundamental, non-rational moral drives (Cushman et al., 2010; Machery & Mallon, 2010; Wilson, 1993). In an experimental study of ultimatum bargaining, Boles, Croson, and Murnighan (2000) found that participants punished bargaining partners who lied by rejecting their offers, even when it resulted in less total gain for themselves. The authors observed that, overall, "the bargainers were little like those depicted by rational economic models. They offered too much, they rejected offers that they should have accepted, and emotions rather than simple profits seemed to have important effects on their behavior" (p. 255) (see also Pillutla & Murningham, 1996). An experimental study by Gneezy (2005) found, similarly, that participants turned down opportunities to deceive others, even when they would have benefited from the deception and could not have been detected (see also Gino & Pierce, 2009).

Results such as these have helped bring about a recent shift in moral psychology research from a strictly rational-cognitive view to a dual-process conception by which non-rational determinants of moral judgment such as emotion and intuition interact with, and are controlled by, cognitive processes such as moral reasoning (Green et al., 2008; Haidt, 2001, 2003, 2007; Narvaez, 2010; Shalvi et al., 2012). The dual-process view of moral judgment posits an evolutionary basis for moral judgment (Greene et al., 2006; Knoch et al., 2006; Machery & Mallon, 2010) that has also been characterized in terms of social contracts (Cosmides & Tooby, 2013; Rettinger, 2007). An experimental study by Cosmides (1989) found, for instance, that people appear to possess an innate ability to perceive violations of social contracts, such as the rules that govern alcohol consumption. Participants easily understood rules that were related to social structure, and readily identified violations. The rules of abstract logic systems, while fundamentally similar, were distinctly more challenging for subjects to understand and monitor for violations.

While contractual thinking is highly rational-cognitive in spheres such as law, it is associated in social spheres with mental processes such as 'recognizing', 'weighing' and 'sensing', that occur too quickly to be explained by rational cognition (Kahneman, 2011). "Once students see the social contract against cheating as violated," explains Rettinger (2007, p. 158), "…they do not see the decision to copy another student's homework as one of academic integrity because the social prohibition against it is not in force." When a social contract is violated, in other words, its bindingness may be nullified for all parties. With regard to cheating, the official rules of the school or classroom may be inconsistent with what students believe the *de facto* rules to be, based on their observations of interpersonal and social factors. In a study of cheating among 164 American undergraduates, O'Rourke et al. (2010) found, for instance, that "seeing others cheat increases cheating behavior by causing students to judge the behavior less morally reprehensible, not by making rationalization easier" (p. 47).

In response to the apparent disjunction between cheating-related cognition and cheating behavior, references to the dual-process perspective have been made in several recent cheating studies (Harding et al., 2012; Murdock et al., 2008). For instance, O'Rourke et al. (2010) entertain in their discussion the possibility that "automatic emotional responses determine cheating behavior" (p. 63). Factors such as mental overload and fatigue have, moreover, been found to impair cognitive control, which leads to poorer moral judgment (Barnes et al., 2011; Greene et al., 2008). Gino et al. (2011) found that cognitive tasks such as writing short essays without using words that include the letters N and A, depleted subjects' self-control, which resulted in both lower moral awareness and lower resistance to temptation.

Of the many integrated rational-cognitive models for cheating tested at the secondary level, none accounts adequately for cheating behavior from a purely rational-cognitive perspective. The more recent dual-process paradigm of moral psychology (Haidt, 2007; Kahneman, 2011) suggests that both cognitive factors such as reasoning and self-regulation, and non-cognitive factors such as emotion and intuition may underlie the noted incongruence between abstract moral beliefs and contextualized moral behaviors, or BBI (Ajzen & Sexton, 1999; Stephens & Nicholson, 2008). While dual-process perspectives have been emergent in moral psychology research for more than two decades (Chaiken & Trope, 1999; Damasio, 1994; Rettinger, 2007; Wilson, 1993), no expressly dual-process framework has yet been devised for academic cheating.

2.6 Neutralization techniques

Neutralization theory, originally developed by Sykes and Matza (1957), holds that individuals are able to break certain rules that they otherwise value by reasoning that certain circumstances alleviate the moral imperative to follow those rules. Such reasoning techniques include denying that one's misbehavior harms others, or ascribing responsibility to external factors (see Table 2.1). Interpreting the justifications offered by students for cheating behaviors as neutralization techniques reflects, therefore, an underlying assumption that cheating is a rational-cognitive act. Rule-breakers must be aware of the immorality of their behavior in order to neutralize the "disapproval flowing from internalized norms and conforming others" (Sykes & Matza, 1957, p. 666). Because neutralization techniques theoretically enable individuals to violate their own moral standards, they have been widely accepted as an adequate way to explain the BBI (e.g. Blasi, 1983; Murdock, & Stephens, 2007; Olafson et al., 2013; Stephens & Nicholson, 2008). If moral awareness must be conscious in order for an individual to neutralize the effects of an immoral act, however, which is a position taken in most published research on the neutralization of cheating behavior, then neutralization must involve intentional self-deception. To the extent that moral awareness may be unconscious, neutralizing justifications may, in fact, be valid in the minds of those who assert them, which raises the question of whether they would then have anything to neutralize.

Categorizing all of the justifications that students give for cheating as intentional self-deception seems likely to preclude efforts to seek deeper insight into why students cheat and how these reasons can be addressed. Rule-breaking is not, after all, inherently immoral, just as acts that are legal are not always ethical (Crittenden et al., 2009a). Research related to domain theory (Turiel, 1983) indicates that adolescents do, in fact, distinguish between moral imperatives and conventional rules (Thornberg, 2008), and may, at times, view rules that prohibit cheating in a conventional, or a-moral, light (Eisenberg, 2004). Inasmuch as students fail to recognize the moral validity of rules regarding honor, classifying their justifications for cheating as neutralization techniques does not resolve the BBI, but only replaces it with another, subtler incongruity – between rules and morals.

Interpretations of neutralization techniques as (1) valid in the minds of rule-breakers and as (2) intentional self-deceptions can both be drawn from the original framework of Sykes and Matza (1957). Neutralization techniques are compared, for instance, to the legal institution of "defenses to crimes", such as insanity and self-defense (p. 666), which, under Anglo-American law, may absolve individuals of culpability for infractions of law. A criminal justification concedes, for instance, that an individual has broken a law, yet it also challenges whether the infraction was immoral and thus whether the individual has acted criminally (Morawetz, 1986). It would appear from this comparison that, like criminal defenses, neutralizing rationales are also potentially legitimate, at least in the minds of those who assert them. The view that neutralizing justifications may be valid in the minds of cheaters emerges several times in the literature on academic integrity. Justifications for cheating may, for instance, be interpreted as valid by a student who "believes circumstances permit or require violating the norm" (Galloway, 2012, p. 382), or when they result from "immature moral reasoning" (Diekhoff et al., 1996, p. 500). Justifications that a student 'believes' are legitimate, and that result from honest judgment, albeit immature, are unlikely to involve intentional self-deception.

The characterization of neutralization techniques as 'intentional self-deception' is otherwise dominant in the literature of academic integrity. Examples include "strategies used to justify dishonesty" (Murdock & Anderman, 2006, p. 137), "excuses to reduce the amount of personal blame associated with cheating" (Olafson et al., 2013, p. 149), and "attitudes [that] allow people to justify behavior they know to be wrong" (O'Rourke et al., 2010, p. 49). These characterizations clearly imply conscious intent behind the 'use' of neutralization techniques to 'reduce', 'allow' and 'justify', in relation to behavior that is 'known to be wrong'. This view is rooted in the assumption that neutralization is accompanied by moral awareness. Inasmuch as morality is a matter of intent (Blasi, 1980), students who do not recognize that cheating is immoral cannot, in fact, be said to act immorally. Thus before students can neutralize the immorality of cheating, they must recognize that cheating is immoral. Individuals who neutralize must be aware that their rationalizations serve immoral purposes, and must fool themselves willingly. This underlying assumption is implied most obviously by Sykes and Matza's (1957) choice of the term 'techniques', as distinct from 'beliefs' or 'misapprehensions'. Students use neutralization techniques, by this conception, to outwit their own moral sensibilities and the moral sensibilities of others, so as to engage in immoral acts with minimal damage to their self-image. The paradox of 'intentional self-deception' is expressed in the following excerpt:

> "In this sense, the delinquent both has his cake and eats it too, for he remains committed to the dominant normative system and yet so qualifies its imperatives that violations are 'acceptable' if not 'right.' Thus the delinquent represents not a radical opposition to law-abiding society but something more like an apologetic failure, often more sinned against than sinning in his own eyes. We call these justifications of deviant behavior techniques of neutralization." (Sykes & Matza, 1957, p. 667)

This excerpt sets up an incongruence between the beliefs and behaviors of a 'delinquent' that resembles the BBI problem identified in cheating literature (Stephens & Nicholson, 2008). The incongruence is described as a conflict between the 'dominant normative system', to which the delinquent is committed, and his or her transgressions against that system. Because the delinquent is aware of violating 'the imperatives' of the system to which he or she is committed, he or she employs neutralization techniques to justify that deviance. While the delinquent may 'often' view him or herself as 'sinned against', the fact of being aware of the immorality of his or her deviance implies that techniques of neutralization are being employed intentionally, i.e. as means of intentional self-deception.

Evidence for a relationship between neutralization techniques and cheating has been identified in many studies. Justifications for cheating that fit neatly within the neutralization categories provided in Table 2.1 have emerged in several qualitative studies at both the secondary level (Galloway, 2012; Taylor, Pogrebin, & Dodge, 2002; Zito & McQuillan, 2011) and tertiary level (Beasley, 2014; Brent & Atkisson, 2011; LaBeff et al., 1990; McCabe, 1992; Olafson et al., 2013). An experimental vignette study conducted at the tertiary level by Rettinger and Kramer (2009) produced evidence that neutralizing attitudes *cause* cheating, and numerous quantitative studies at both levels support this finding (Davy, Kincaid, Smith, and Travick, 2007; Murdock et al., 2007; 2008; Pulvers & Diekhoff, 1999; Stephens & Gehlbach, 2007; for a review, see Whitley, 1998).

Table 2.1

Neutralization techniques

Technique	**Characterization**
Denial of injury	Cheating does not hurt anyone.
Denial of the victim	Those harmed when I cheat deserve the harm. They are not victims.
Appeal to higher loyalties	I cheat in order to serve higher moral principles.
Denial of responsibility	I am impelled to cheat due to forces beyond my control.
Condemnation of the condemners	The authorities cheat, too; their judgment of me when I cheat is hypocritical and irrelevant.

The neutralization framework has also been criticized as a "stylistic convention rather than a genuine theory" that makes poorly-examined assumptions about how students judge the morality of individual instances of cheating (Bouville, 2007, p 7). Students are, for instance, frequently seen to neutralize cheating when they claim that factors beyond their control impel them to cheat, such as unfair workload, low teacher quality, and low interest in the class – even when the relationship between these factors and cheating is well-corroborated by statistical evidence. In the only study of neutralization to have been conducted at the middle school level, Zito and McQuillan (2011) found, for instance, that students justified cheating based on feeling that a teacher either assigned too much work or did not explain concepts well. At the tertiary level, Pulvers and Diekhoff (1999) found, moreover, that "neutralization of cheating accompanied perceptions of the classroom as less personalized, less involving, less cohesive, less satisfying, less task oriented, and less individualized" (p. 495).

Substantial correlational research reviewed in earlier sections shows, moreover, that students who genuinely do have such perceptions of learning environments are, in fact, statistically more likely to cheat. These statistical patterns align quite closely with many of the justifications that are categorized as neutralization techniques. Deeming such justifications to be intentional self-deceptions seems, therefore, not so much of a theoretical advance, as a rather arbitrary accusation.

An outstanding exception to the view that every justification for cheating is a self-serving self-deception is Brent and Atkisson's (2011) assertion that the ways students neutralize cheating "at least sometimes display a consistent logic" (p. 655). While some students in the study appeared able to rationalize cheating in virtually any situation, others appeared to work within "a rational framework for justifying cheating that has some coherence", which the authors describe as a "student-teacher contract" (p. 656). They assert that such contractual frameworks may embody students' expectations for what teachers should do and how well they should do it, such that cheating may become justifiable when these expectations are not met.

A psychological basis for the contractarian style of reasoning detected among participants in Brent and Atkisson's (2011) study can be found in domain theory (Nucci, 2001; Richardson, Mulvey, & Killen., 2012; Turiel, 1983, 2002, 2006), which holds that children distinguish between two broad domains of activity: the moral, and the conventional. Rules and activities belong to the moral domain when they involve harm or benefits to others, whereas they belong to the conventional domain when they arise from custom or social norms and are related to conformity, such as taking one's hat off upon entering a building (Murdock

& Stephens, 2007). Inasmuch as students' contractual frameworks comprise the benefits that students expect from teachers, i.e. what they believe to be teachers' obligations, such as caring, fairness, and pedagogical skill, then breaking those contractual expectations may negate a teacher's moral authority in the eyes of students. By neglecting his or her obligations, a teacher may be perceived to convey insufficient benefit, or perhaps even harm, to students. Research in domain theory suggests that when children perceive something to be less beneficial or more harmful, they are less likely to respect it (Thomson & Holland, 2002). An ethnographic study of Swedish school children by Thornberg (2008) found, for example, that students did not passively accept school rules as inherently moral, or even necessarily view them as worthy of being followed. They tended, instead, both to actively judge the morality and legitimacy of rules and teachers, as well as to "judge moral transgressions as wrong regardless of the presence or absence of rules" (p. 49). Upon perceiving that a teacher neglects his or her professional obligations, students may cease to view rules governing their interaction with that teacher in a moral light, and therefore cease to feel morally obliged by them. While breaking conventional rules may risk formal consequences, it would not risk, at least in the mind of the actor, the taint of moral disgrace. When students break rules that they perceive as conventional, they may feel little or no remorse. The justifications they give for such infractions would, as such, have little or nothing left to neutralize.

While Sykes and Matza (1957) theorized that neutralization techniques may lie "behind a large share of delinquent behavior" (p. 669), the framework has often been applied with a broad brush to all of the justifications that students give for cheating, as a theoretical means of squaring abstract moral beliefs with specific behaviors that contradict them. According to Blasi (1983), for instance, "not to act according to one's judgment should be perceived as a substantial inconsistency, as a fracture within the very core of the self, unless neutralizing devices are put into operation" (p. 201). While the neutralization framework likely does explain variance in how cheaters avoid damaging their self-image, there is no compelling reason to believe that most justifications given for cheating are neutralization techniques. By that logic, Bouville (2007) argues, even "Robin Hood could be said to neutralize his wrongdoing by shifting blame to the wronged" (p. 3). Inasmuch as students differentiate between moral and conventional rules, as domain theory suggests (Turiel, 1983), there may, in fact, be two types of cheating: cheating as a moral infraction and cheating as a conventional infraction. While neutralization appears relevant to the former type of cheating, it does not, on face value, pertain to the latter type. The contract-like framework for cheating justifications described by Brent and Atkisson (2011) may, in fact, operate as a mechanism by which students judge whether rules occupy the moral or the conventional domain. When students think a teacher has broken the teaching-learning contract, i.e. by failing to meet his or her obligations, the students may no longer feel obliged to relate morally to that teacher, thus shifting their view of rules that forbid cheating in his or her class to the conventional domain, where breaking them does not feel like a moral offense. The fundamental difference between the contractarian and neutralizing views of cheating appears to be that, by the contractarian view, the a-morality of cheating may be a genuinely-held belief, whereas by the neutralizing view it is an intentional self-deception.

2.7 Psychological teaching-learning contracts

The term 'teaching-learning contract' is adopted from the work of Murdock and colleagues (2001, 2004), who first introduced the contract metaphor for students' tendencies to justify cheating in terms of teacher-learner reciprocity. This term is modified as 'psychological' in order to emphasize that teaching-learning contracts are, in the present work, subjective constructs held by students that may have no objective validity. The expectations for learning experiences that students include in their notions of reciprocal fairness may not be objectively valid. Similar to legal contracts that comprise mutual obligations, or promises, agreed to by at least two parties (Mather, 1999), psychological teaching-learning contracts (PTLCs) comprise students' expectations of, and reciprocal sense of moral obligation within, learning contexts. PTLCs are held to be implied socially, constructed psychologically, judged both intuitively and cognitively, and wired into the neural structure of the human brain.

While contract metaphors have appeared in the literature of academic integrity for over a decade, little has been written about them. In addition to the three works cited above in relation to PTLCs (Brent & Atkisson, 2011; Murdock et al., 2001, 2004), Colnerud and Rosander (2009), who studied cheating among Swedish university students, used the contract metaphor to emphasize students' obligations, as "defined by the curriculum and the courses and the forms of examinations defined by the university" (p. 514). Students who cheat fail, therefore, "to fulfill [their] side of the contract" (p. 514).

Similar contract metaphors have also been described in literature on children and education more broadly. 'Didactical contracts', introduced by Brousseau (1984), pertain, for instance, to contract-like expectations that evolve between students and teachers within particular class contexts, that "serve to delimit 'legitimate' activity by the teacher", such as precedents set for disciplinary action, and for the types of knowledge teachers require students to learn (e.g. lower- *vs.* higher-order) (Black & Wiliam, 1998, p. 56). PTLCs extend the concept of 'didactical contracts' to all of the beliefs and expectations of what should be entailed by the roles of teacher and student that are implied by broader socio-cultural contexts. While such beliefs and expectations may vary widely across societies and cultures, it is likely that teachers in most, if not all, cultures and societies, are expected to *help their students*. The contract-like nature of the duty to help children, and its many implications in the realm of parenting, was highlighted by Baumrind (1987), who referred to the 'implicit contact' between parent and child, according to which parental authority is, like teacher authority, bestowed and exercised primarily for the benefit of children. When the parent-child contract is violated by parents over time, "children are less likely to attribute responsibility to themselves as moral agents" and may, therefore, not "feel obligated to abide by the explicit or implicit contracts they have with their parents or with society" (p. 111).

All such contract metaphors emphasize the high premium that students, and arguably all people, place on reciprocal fairness. The notion of fairness is raised numerous times in the literature of cheating, and always with the same basic message: teachers who do not "play by the rules" tend to relieve students of the felt obligation to do the same (Murdock et al., 2001, p. 110; also Calabrese & Cochran, 1990; Evans et al., 1993; Jensen et al., 2002; McCabe & Katz, 2009; Murdock et al., 2008; Thorkildsen, Golant, & Richesin, 2007). Shirk and Hoffman (1961) argued, in reference to cheating, that students who perceive unfairness on the part of the teacher may be "tempted to offer an eye for an eye and a tooth for a tooth" (p.132). This reference to Hammurabi's Law Codex (Roth, 1995), which may be the first expression of negative reciprocity under a formal system of law in human history (Fehr & Gächter, 1998), helps illustrate the universality of reciprocal fairness to conceptions of justice (Brüne et al., 2013).

Since King Hammurabi's time (1792-1750 BCE), contractarian notions of reciprocal fairness have infused leading theories in numerous fields of social science such as natural jurisprudence (Grotius, 1625; Locke, 1689; Rawls, 1971; Rousseau, 1762), moral philosophy (Kant, 1797), and economics (Fehr & Gächter, 1998). The human sense of positive reciprocal fairness, or 'reciprocal altruism' (Brosnan & de Waal, 2002; Cosmides, 1989; Cosmides & Tooby, 2013; Trivers, 1971), has more recently been implicated as an outcome of natural selection that permits two-party cooperation and recognition of social contract violations (Cosmides, 1989; Griggs & Cox, 1983; Cosmides & Tooby, 2013). The contention that social contract-based judgment is an evolved function of the brain is supported by findings in neuroscience research that judgments of reciprocal fairness involve brain regions responsible both for cognition, such as the dorsolateral prefrontal cortex (Knoch et al., 2006), and for emotion, such as the limbic region (Greene et al., 2004; Haidt, 2007; Haidt & Greene, 2002). Sensibilities analogous to social contract awareness have also been observed in primates such as chimpanzees (de Waal, 1991, 2014), and capuchin monkeys (Brosnan & de Waal, 2003). De Waal (1991) observed, for instance, that chimpanzees engaged in 'moralistic aggression' in response to "dissatisfaction about the cost/benefit balance of [a] relationship (e.g. lack of reciprocation)" (p.342). Capuchin monkeys were found, similarly, to "respond negatively to previously acceptable rewards if a partner [got] a better deal", which Brosnan and de Waal (2003, p. 299) interpreted as evidence of "social emotions… known as 'passions' by economists" that "guide human reactions to the effort, gains, losses and attitudes of others."

Researchers in the field of organizational behavior have found that employees who feel *un*fairly treated in the workplace often reciprocate with harmful behaviors such as unnecessary absenteeism (Pillutla & Murnighan, 1996), vandalism, and theft (Greenberg, 1993; Skarlicki, & Folger, 1997). A similar ethos of negative reciprocity also emerges in qualitative research on why students cheat, such as a pair of 'typical' rationalizations offered by a secondary student interviewee of Stephens and Nicholson (2008): "This class sucks. I'm cheating the system" and "he doesn't spend any time making up new tests, I don't have to spend any time studying" (p. 367). What comes through from these quotes is not aggression, however, but passivity. The student has absolved himself of the moral obligation, and thus the 'moral motivation' (Schroeder et al., 2010), to be honest, in reciprocation of his teachers' perceived poor performance. This is the sort of response expected to perceived PTLC violations by teachers: not aggression, but moral absolution.

The role of PTLCs in cheating may have been overlooked in prior scholarship due to the overwhelming dominance of the rational-cognitive paradigm in educational research.

Studies of Kohlberg's framework for the development of moral cognition indicate, for instance, that adolescents are seldom cognitively developed enough to engage in contractual reasoning, a 'postconventional operation' (Colby et al., 1983; Colby & Kohlberg, 1987; Rest, 1986). Research in the solely cognitive paradigm has, however, not been sufficient to explain behavior-related anomalies such as the BBI, and the fact that cheating becomes more prevalent as children mature through high school, which runs counter to Kohlberg's thesis that immoral behavior should decrease as individual moral cognition develops in tandem with age. The field of cheating psychology may, for these reasons, benefit from adopting an expanded conception of moral judgment.

The dual-process paradigm of moral psychology allows for the possibility that an emotional-intuitive sense of morality is a key feature of non-rational psychological processes that are neither age- nor stage-dependent. As individuals develop cognitively, these non-rational moral processes are brought increasingly under cognitive control (Haidt, 2007), albeit only to the extent that individuals choose to engage such control (Bandura, 1999; Thorkildsen et al., 2007). The ability to fully comprehend and articulate one's non-rational sense of morality may, therefore, also depend heavily upon one's level of cognitive development, potentially making it more difficult for less cognitively mature adolescents to express, or to even be fully aware of, the reasons why they cheat. The inability to articulate rational bases for 'felt' moral judgments has been referred to as 'moral dumbfounding' (Bjorklund, Haidt, & Murphy, 2000; Sneddon, 2007). Consider, as an example, the following statement in defense of cheating from an adolescent student interviewed by Galloway (2012):

> "…basically it's just like, it's not necessarily that we're compromising our morals and values; it's like you're compromising for like a just reason. It's like hard to say, but like you're compromising it for sort of a good." (p. 393)

PTLCs are hypothesized to embody both the expectations that students have, and the obligations they feel, in learning situations. The perception by a student that an academic context is unfair, inappropriate, or harmful to students may accompany a shift in his or her judgment of rules and authority figures, from the moral to the conventional domains. Rules perceived in the conventional domain may be violated without contrition. The BBI is recast, by this view, from an issue of why students behave in ways they know to be immoral, to the question of why, in certain situations, they fail to regard cheating as immoral. Brent and Atkisson (2011) argue that recognizing the potential for coherence in students' justifications for cheating provides valuable insight into how the problem may be addressed by educators. "A rational framework for justifying cheating that has some coherence", they write, "…might also mean that cheating behavior can be constrained by that rationality" (p. 656). Inasmuch as students' justifications for cheating are valid within their own minds, and are genuinely "constrained by a legitimate set of expectations" (p. 656), their fundamental integrity is still intact. Such students can still, therefore, be trusted to learn and evolve when their concerns are addressed with equal earnestness.

2.8 Chapter summary

This chapter offered a critical review of published research on cheating. A detailed review of an on-going conversation in the literature about how cheating should be defined was reviewed first, in order to locate the most appropriate definitional basis for a study concerned with addressing the belief-behavior incongruence (BBI), an inherently moral issue. Definitions that emphasize the abstract, unifying properties of acts referred to as 'cheating' were identified as the most morally relevant.

Numerous constructs were then reviewed that have appeared during the last 110 years of published empirical research on cheating, including personological variables (e.g. gender, grade-level, self-beliefs, and learner characteristics) and prominent situational variables (pedagogical quality, interest, assessment, workload, and peer influence). Rational-cognitive models of deviant behavior, which have dominated the psychology of academic cheating since the 1970s, were reviewed in light of growing evidence that moral judgment often involves automatic, emotional-intuitive processes that are not consistent with a strictly rational-cognitive view. A 'dual-process' paradigm of moral psychology, associated with such scholars as Kahneman (2011) and Haidt (2007), was introduced that accommodates both rational-cognitive and emotional-intuitive factors that may explain variance in academic cheating.

The neutralization framework, associated with the work of Sykes and Matza (1957), was criticized, in particular, for entailing the automatic assumption that every justification given by students for cheating is an opportunistic self-deception. This interpretation of the reasons students give for cheating, which ignores corroborating correlational and experimental evidence of its contextual antecedents, appears to have risen in popularity as a means of explaining how students cope with the cognitive dissonance theorized to result when their behaviors, such as cheating, contradict the abstract moral beliefs they profess to hold.

A contractarian perspective for how students seem to justify cheating, raised in several studies, and related to other uses of contract metaphors found in research on how children and adolescents relate to adult authority, was identified as an alternative to neutralization theory for explaining why students' beliefs and behaviors may, in some proportion of instances, appear to be incongruent, as a result of what are actually incongruences between their abstract moral beliefs and the moral validity they ascribe to rules within specific contexts. Contractarian metaphors for how students judge the justifiability of rule-breaking, in general, were grouped under the rubric 'psychological teaching-learning contracts' (PTLCs). PTLCs provide a plausible theoretical perspective on how students may come to judge rules, such as those that forbid cheating, as purely conventional, and thus preclude any incongruence between their cheating behavior and moral beliefs.

CHAPTER 3

EMPIRICAL RATIONALE

> Because they first broke the oath we swore together, there has been no injustice at all in our entering their land…. And there will be no injustice in what we are about to do now.
>
> -Archidamus' prayer before the siege of Plataea, 429 BCE
>
> (Thucydides, c420 BCE)

The psychological teaching-learning contract (PTLC) hypothesis of academic disintegrity is explicated in this chapter, and developed into a general PTLC framework, which is situated within the ecological view of learning as a dynamic system, as portrayed by the *Presage, Process, Product Model* (3-P Model; see Figure 3.1) (Biggs et al., 2001), a key conceptual model in student learning theory (Biggs & Tang, 2011). The dual-process paradigm of moral psychology is presented as the theoretical basis for positing contractarian moral judgment among adolescents, which would, according to the strictly rational-cognitive paradigm (e.g. Colby & Kohlberg, 1987), be beyond their developmental limits. Finally, a contractarian structural equation model of disintegrity is developed according to the general PTLC framework that will be tested in the present research program. This structural equation model (hereafter 'PTLC model') incorporates many of the personological and environmental antecedents of cheating that were emphasized in the preceding literature review. The set of specific hypotheses and constructs included in the PTLC model is also presented and justified.

3.1 Definitions

Variation in how cheating is defined by students and teachers has led to disagreement among scholars about whether the concept should be measured as an abstraction or as an inventory of behaviors (Garavalia et al., 2007). The present study resolves this issue by classifying cheating as *intentional* academic deviance. Cheating is defined thus as *any act that a student believes would result in negative consequences if detected, because it contravenes the spirit and/or letter of rules related to honorable academic conduct*. 'Accidental cheating' is, by this definition, an 'honest mistake' that does not reflect intentional wrongdoing. When students describe behavior as 'cheating', they refer to acts that explicitly break rules by which they understand cheating to be defined. Self-reported cheating may or may not, in this sense, be immoral in the eyes of a student, but is nonetheless understood to be a rule violation. The study assumes that students are never under the impression that what is called 'cheating' is acceptable according to school or classroom rules.

The concept of *disintegrity*, which includes cheating as well as acts that "lack integrity or subvert the goals of education", but that are not labeled as cheating (Miller et al., 2011, p. 170), will be used to expand the concept of academic dishonesty in the present work. Disintegrity includes behaviors such as surface learning strategies that do not violate rules, but that nonetheless involve falsifying knowledge, understanding, and skill in order to obtain grades. The term 'legal cheating', used in a similar sense by Kohn (2007b), is evidence that this perspective is emergent in cheating scholarship. Surface learning strategies, described in student learning theory as being oriented to the symbols of learning, instead of to its substance (Marton and Säljö, 1976), help, like cheating, to minimize effort, and over-represent actual intellectual accomplishment. Surface learning strategies are, in this sense, means of academic deception that, while not referred to as 'cheating', meet Miller et al.'s (2011) criteria for disintegrity, by subverting the goal of meaningful learning.

3.2 The PTLC hypothesis

The overarching PTLC hypothesis is that *the degree of moral obligation that students feel to work hard and be honest in a given class context fluctuates directly with how well they think the basic obligations of teachers and classes are met in that context*. This wording intentionally includes non-teacher factors, such as external testing, resource availability, and administrative policies that are beyond teachers' control. Referring, in this sense, to 'classes' and 'academic contexts' as though they bear moral responsibilities is meant to implicate the administrators at all levels of an educational bureaucracy whose effects are felt in classrooms.

The language of the PTLC hypothesis also emphasizes the subjective nature of students' judgments. A student's view of his or her contractual relationship with a given class context is psychological, and may or may not be consistent with reality. That such judgments may be inconsistent with reality does not necessarily make them invalid to students who hold them, which is what distinguishes contractarian judgments of the moral validity of rules from neutralization techniques. Neutralizing the cognitive dissonance that one experiences when violating his or her own moral beliefs implies intentional self-deception, as argued in section 2.6 of the preceding chapter, whereas violating rules that one does not recognize as being moral precludes any possibility of cognitive dissonance.

Domain theory (Turiel, 1983, 2002, 2006) holds that adolescents tend to view rules as being either moral or conventional in character, and that their view of a given rule may shift within specific contexts (Thomson & Holland, 2002; Thornberg, 2008). The fact that students' views on specific rules may shift between the moral and conventional domains suggests that their judgment of the morality of those rules is an on-going process within the dynamic system that learning entails. In dynamic learning systems, as illustrated by Biggs' 3-P Model (Biggs, 1987, 1993; Biggs et al., 2001) (see Figure 3.1), "all components interact to strive towards equilibrium" (Biggs, 1993, p. 76). This is schematized in the 3-P Model with double-headed arrows between all components, such that every component interacts directly, and through all possible mediated pathways, with every other component. Applying this dynamic systems view to the overarching PTLC hypothesis casts the hypothesized reciprocity between (1) students' perceptions of learning context quality and (2) their felt moral obligation as the outcome of on-going equilibration between students' achievement behaviors (product variables) and perceptions of context (presage variables), through processes of moral judgment (process variables). Product variables in the present work include cheating and surface learning strategies, grouped together as 'disintegrity'; as well as deep learning strategies, which involve striving for personal meaning and understanding that is genuine, i.e. learning with integrity.

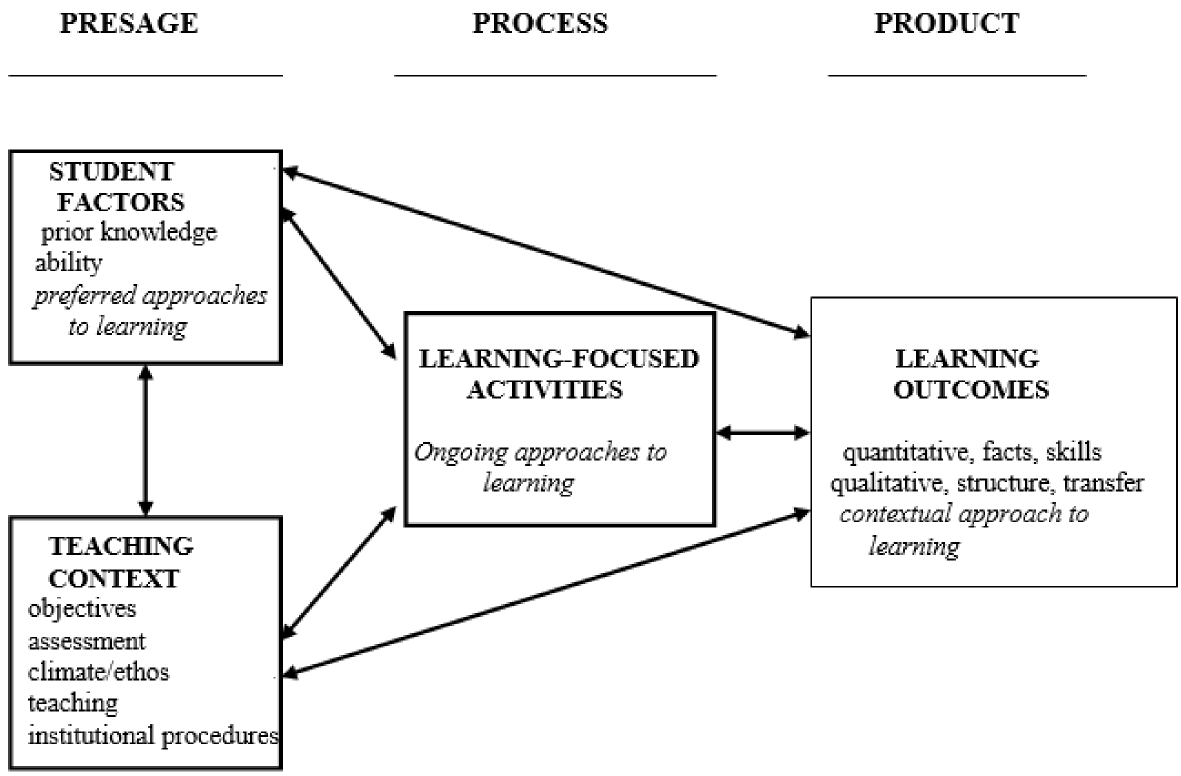

Figure 3.1. The *Presage, Process, Product* model of a dynamic learning system (Biggs et al., 2001), within which the PTLC framework is situated.

Shifting one's view of a given rule to the conventional domain entails an erosion of the moral validity of that rule. Rules in the conventional domain may be broken without moral qualm. Students who view rules that forbid cheating as conventional may, therefore, cheat without concern for their moral self-concept. Acknowledging that students may view integrity-related rules as conventional abandons the assumption implied by neutralization techniques that students apply their abstract moral beliefs to all educational contexts, such that acts of cheating result in cognitive dissonance that must then be neutralized. Students who break rules they hold to be conventional produce no incongruity between their moral beliefs and behaviors, or BBI, *per se*. The incongruity would, instead, be between their abstract moral beliefs, and their view of particular rules.

3.2.1 The PTLC hypothesis within the dual-processing paradigm of moral psychology

PTLCs propose to explain why a student's view of rules might shift from the moral domain to the conventional domain as a result of judgments based on contractarian reciprocal fairness. Students are hypothesized, by this perspective, to judge their moral obligation to follow rules in a given learning context according to how well they think the context fulfills its moral obligations to them, i.e. its 'quality'.

Obligations created by promises are the quintessence of contractual relations in jurisprudence (Mather, 1999) and politics (Medina, 1990). Contractual obligations may be accepted by individuals explicitly, as in rental contracts, or may be implied by the cultural, social, or interpersonal frameworks within which individuals operate (MacNeil, 1974). Contracts of an implied nature have been referred to in jurisprudence as 'relational contracts' (MacNeil, 1974). The obligations that students and teachers have toward one another are overwhelmingly relational (Buckley et al., 2004). The role of the teacher, for instance, includes both explicit professional duties as well as social and cultural responsibilities related to protecting and promoting student welfare. Beyond the responsibilities of individual teachers, formal academic contexts also carry obligations, such as to be worthwhile, credible, interesting, and safe. Students are, by contrast, obliged to follow explicit school and classroom rules, and to observe socio-cultural norms such as working hard and respecting teachers. A student's personal sense of this set of mutual obligations is his or her PTLC.

By assuming that social contract-based moral judgments are common among adolescents, the PTLC hypothesis moves beyond the strictly rational-cognitive paradigm of moral psychology associated with Bandura (1977, 1999) and Kohlberg (1958, 1968), which holds that contractarian judgment is beyond the developmental limits of most adolescents. A crucial aspect of the rationale for the PTLC hypothesis is, therefore, the assertion that social contract-based judgment is an evolved function of the human mind (Cosimdes & Tooby, 2013).

The ability to recognize social contract violations, as embodying the concept of reciprocal fairness, has been described as an "evolved 'Darwinian algorithm'" (Cosmides, 1989, p. 195) that operates with a high degree of automaticity and "independent of general cognitive resources", standing "in shrill contrast with the classical view that states that all behavior is based on one general learning mechanism (i.e., general cognitive capacity, intelligence, rationality)" (Van Lier, Revlin, & De Neys, 2013, p. 2). Automaticity is, moreover, characteristic of emotional-intuitive judgment processes that have been traced to emotional centers of the brain such as the limbic region (Knoch et al., 2006). Dual-process theories of moral psychology generally assert that both emotion and rationality are involved in moral judgment (Chaiken & Trope, 1999; Cushman et al., 2010; Narvaez, 2010), where rationality is necessary for marshaling the 'rapid-fire' of emotional-intuitive impulses (Greene et al., 2008; Kahneman, 2011; Mallon & Nichols, 2010). The assertion that social contract thinking actively involves both types of process is supported by its direct association in *f*MRI research with the left and right medial frontal gyri of the human frontal lobe (Fiddick, Spampinato, & Grafman, 2005), where the "integration of emotion into decision-making and planning" takes place (Greene & Haidt, 2002, p. 520). Adolescents may be especially vulnerable to faulty moral judgments, in general, inasmuch as their capacity for cognition is still developing. They may have less ready access, for instance, to the rational tools needed to analyze and possibly discredit the validity of perceived social contract violations, and/or to restrain their emotional-intuitive judgments of how to respond to social contract violations that they perceive as real. Adolescents may also, as argued in Chapter Two (see section 2.7), be prone to 'moral dumbfounding' (Bjorklund et al., 2000; Sneddon, 2007), or the inability to clearly articulate feelings, senses, and other non-rational experiences that they associate with moral judgments.

The PTLC model avoids the pitfall of moral dumbfounding, by focusing exclusively on the end-products of judgment processes. Instead of asking respondents for their reasons and justifications for cheating, which they may or may not be equipped to provide, the PTLC model is itself a diagram of the actual hypothesized mechanism of contractarian moral judgment. It articulates, on students' behalf, a contractarian heuristic for the justifiability of cheating, in terms of class context quality. Testing the PTLC model thus uses the lens of statistical regression to examine the validity of linkages among a suite of psychological factors that should exist if social contract thinking does, in fact, underlie students' judgments of whether cheating is justifiable.

Framed as an outcome of social contract-based judgment, student cheating is a primal response to a students' perception that he or she is being cheated by academic contexts. Cosmides and Tooby (2013) explain that "'Cheaters'… violate social contracts by taking the benefit offered without satisfying the requirement on which it was made contingent." (p. 216). A student's experience of perceiving a teacher as shirking or incompetent may, for instance, be the emotional-intuitive equivalent of seeing the teacher cheat. The perception that authority figures, such as teachers and parents, violate their contractual obligations by doing insufficient good or undue harm, may lead young people to believe that their own moral obligations are no longer binding (Baumrind, 1987; Thornberg, 2008).

A student who perceives a learning context to be unfair or of low quality may feel that cheating is justifiable, i.e. not *im*moral, even though he or she still recognizes that cheating is against the rules. The student's perceptions of what is immoral or inappropriate in the academic context become, by extension, his or her justifications for cheating. A student who believes that a particular teacher performs poorly because he or she lacks effort, for example, might feel that cheating is justifiable because the teacher is cheating at his or her professional obligations. While this justification of cheating would be labeled 'condemning the condemners' under the neutralization framework, the PTLC perspective allows that it may stem from what is genuinely perceived to be a social contract violation. It is reasonable to expect that individuals are likely to react negatively to such perceptions and feelings, especially if they are cognitively ill-equipped to debunk them, which may help explain why the incidence of student cheating tends to increase as adolescence runs its course in high school, and to decline thereafter (Miller et al., 2007).

The PTLC model is the first structural model of cheating psychology to be positioned expressly within the dual-process paradigm of moral psychology, in that it avoids moral dumbfounding among adolescents in an investigation of what is held to be an evolved, non-rational process, i.e. social contract-based moral judgment. Mainstream research on moral cognition indicates that adolescence is a time of significant cognitive maturation when, by a dual-process view, non-rational processes involved in moral judgment should be under less cognitive control.

3.3 The general PTLC framework

The PTLC perspective posits that the moral obligation students feel to follow rules is largely contingent upon whether their expectations for what constitutes a fair exchange between themselves and a given learning context are met. The basic structure of PTLCs is common to all contracts: counterparties take on obligations to one another (see Figure 3.2), and failure by either party to fulfill their obligations may reduce the obligations of the counterparty. For Medina (1990), "a contract is roughly understood as an agreement between two or more independent parties who voluntarily choose (consent) to abide by certain rules provided that the other party does not violate it" (p. 3).

Obligations borne by party 1 ⟷ Obligations borne by party 2

Figure 3.2. Basic structure of a contract.

A key difference between PTLCs and legal contracts is that the former are purely psychological constructs. PTLCs are held to include, on one side, the relational variables by which students evaluate how well the obligations of academic contexts are fulfilled (obligations borne by Party 1), and on the other side, the degree of moral obligation that they judge appropriate to take upon themselves (obligations borne by Party 2) (see Figure 3.3).

Student's perception of academic context quality ——(+)⟶ Student's felt moral obligation

Figure 3.3. The general PTLC framework

The factors by which students evaluate academic contexts are held to be 'relational', in that they involve social exchange (MacNeil, 1974), and reflect how positively or negatively students relate to such contexts. The obligations borne by the student, are held, by contrast, to be subjective *felt obligations*.

Hypothesis 1. The degree of obligation that students feel to work hard and be honest fluctuates positively with the perceived quality of an academic context (see Figure 3.3).

Students' prior academic experiences lead them both to hold certain expectations of formal academic contexts, such as what constitutes fair treatment and competent pedagogy, and to understand what behaviors and attitudes are expected of them, such as respect for authority figures, exertion of effort, and honesty. According to domain theory (Turiel, 1983, 2002, 2006) students may view the obligations they carry as being either moral or conventional in character. Feeling morally obliged to follow rules may reflect a student's intrinsic motivation to act in a manner that is consistent with a positive self-concept (Mazar et al., 2008; Shalvi et al., 2011), whereas feeling little or no moral obligation to follow rules implies that a student has identified them with the conventional domain. Conventional obligations are externally enforced, and violating them poses no threat to a student's moral self-concept.

The PTLC framework poses students' expectations of the quality of academic contexts as being reciprocal to their own senses of moral obligation. When students believe that a learning context fails to meet reasonable expectations for quality, they may feel that what should be expected of them is correspondingly reduced, thereby shifting their view of rules and responsibilities from the moral to the conventional domain. By a conventional view, rules that forbid cheating are legitimate only insomuch as they are enforced. Violating rules that forbid cheating but that are identified with the conventional domain should pose little or no threat to a student's moral self-concept, but may instead be constrained by the perceived risk of detection and punishment.

Hypothesis 2. Felt moral obligation is hypothesized to partially mediate the influence of how a student relates to a particular class context on whether he or she engages in disintegrity behaviors in that context. Students' perceptions of class quality factors, such as teacher performance, evaluated in terms of 'better' or 'worse', reflect how positively or negatively students relate to academic contexts. As schematized in Figure

3.4, students who relate less positively to a given learning context will feel less moral obligation to be honest, and will, therefore, engage in more disintegrity.

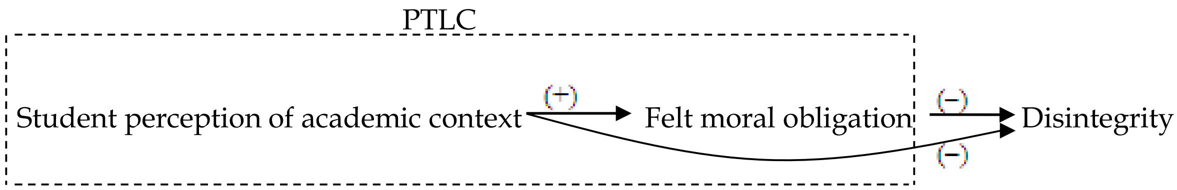

Figure 3.4. The general PTLC framework as an antecedent to behavior

The 3-P Model (see Figure 3.1) asserts that dynamic learning systems involve product components (e.g. behaviors), process components, and presage components, where the latter entail contextual and personological variables (see Figure 3.5). An individual student is held, by this view, to construct and monitor the terms of his or her relationships with teachers against a unique backdrop of personal factors such as demographics, personality structures, and self-beliefs.

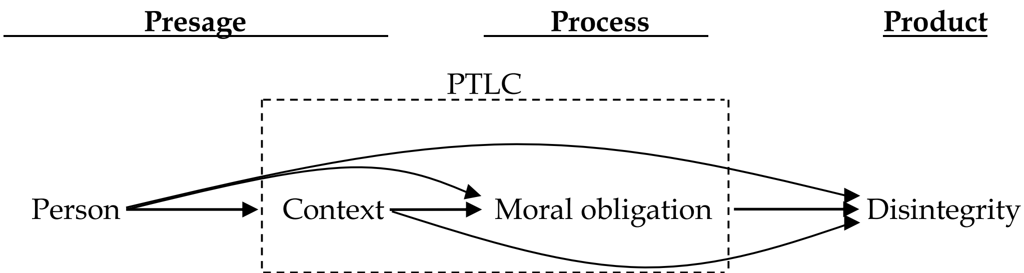

Figure 3.5. The general PTLC framework situated within the 3-P Model framework, emphasizing a left-to-right flow (Biggs, 1993; Biggs et al., 2001).

While the 3-P Model depicts all components in dynamic equilibrium with each other, the left-to-right flow traditionally emphasized in student learning theory (Biggs, 1993) is also emphasized in the PTLC framework (see Figure 3.5). Personological factors are hypothesized, generally, to influence what expectations individual students evaluate most strictly, and according to what specific criteria; what degree of moral obligation they feel initially and how resistant they are to relinquishing it; and what academic beliefs and behaviors have become habituated due to past experience.

3.4 The hypothesized PTLC model

The hypothesized PTLC model is theorized to occur within the broader scope of learning as a dynamic system, as depicted by Biggs' (1987, 1993; Biggs et al., 2001) 3-P Model (see Figures 3.1 and 3.5). The principle mechanism of the hypothesized model, i.e. social contract-based judgment, is consistent with Biggs' (1987) concept of *process* as involving "learning-related activity" (Biggs et al., 2001, p. 138), which mediates the influence of learners' perceptions of academic situations, on achievement behaviors (e.g. 'contextual approaches to learning' in Figure 3.1).

The 'processes' underlying social contract-based judgment are held, in the PTLC framework, to be both cognitive and non-cognitive, i.e. thoughts and feelings by which students adjudge the degree of moral obligation they bear to work hard and be honest within a given class context. As in the 3-P Model, these judgment processes are held to at least partially mediate the influence of personological and contextual perceptions on disintegrity (Figure 3.6). The PTLC is, as such, the hypothesized basis for context-specific moral flexibility (McCabe & Katz, 2009), or 'situation ethics' (Brent & Atkisson, 2011; Crittenden et al., 2009a; Fletcher, 1966), in that students may view cheating in certain situations as a justifiable violation of rules that have been reduced in their eyes to social conventions.

The learning system portrayed by the 3-P Model complements the hypothesized multivariate PTLC model of disintegrity presented in Figure 3.6, in that (A) both involve dynamic interactions between presage, process and product variables; (B) behaviors categorized as disintegrity include surface learning strategies, which are a principal focus of the 3-P Model in student learning theory; and (C) both models involve equilibration, a basic characteristic of dynamic systems (Biggs, 1993), occurring between presage perceptions of context (quality), and learning-related behaviors (disintegrity), as mediated by learning-related processes (moral judgment).

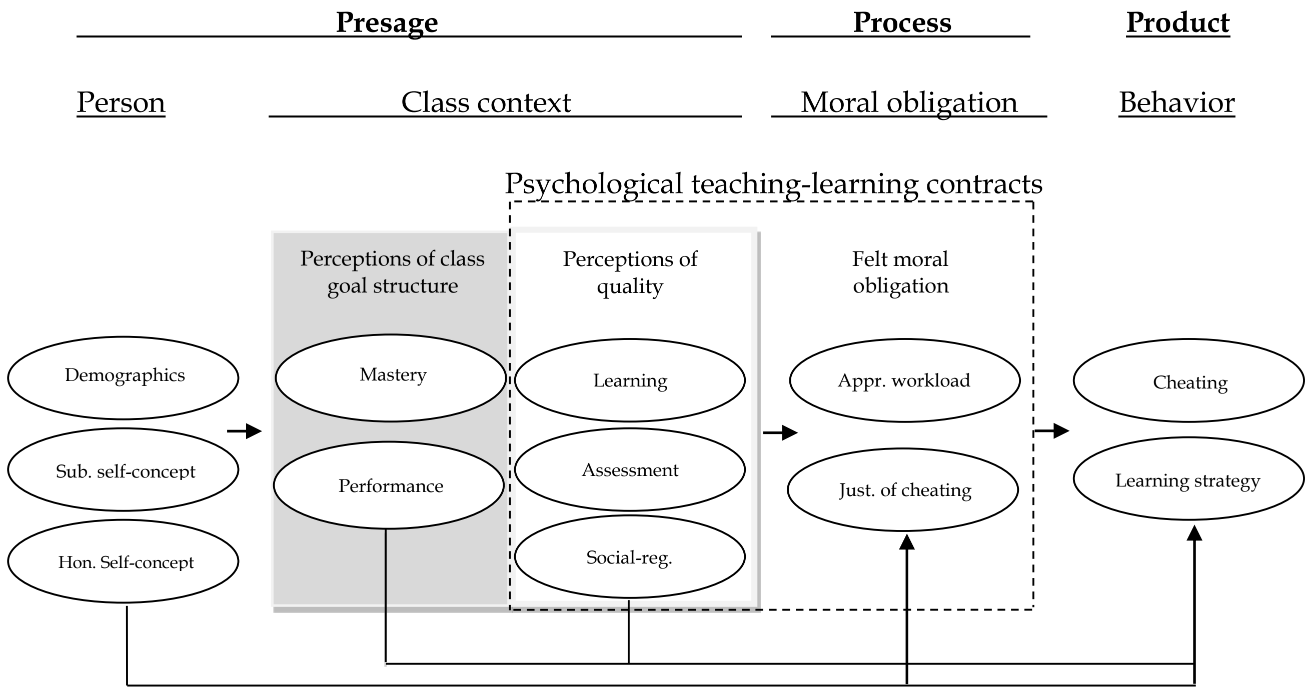

Figure 3.6. The hypothesized PTLC model situated within the 3-P Model framework (Biggs et al., 2001) for dynamic learning systems (see Figure 3.1).

The *presage* category of the hypothesized model includes students' perceptions of motivational goal structure, in addition to the quality of three broad aspects of academic contexts: (1) learning, (2) assessment, and (3) social-regulatory (i.e. rules enforcement and peer norms related to cheating). These perceptions of quality are depicted separately from motivational goal structures in Figure 3.6 to indicate the different roles they are hypothesized to play with respect to moral obligation. While motivational goal structure is theorized to encourage students to adopt either performance or mastery learning goals, students are not necessarily held to relate more positively to either type of goal. Students may, for example, feel positive about performance-oriented classes, and negative about mastery-oriented ones.

The PTLC perspective is envisaged, therefore, to complement, but not accommodate, motivational goal structure.

The *Presage* category also includes personological factors, located within students, that exert both direct and indirect effects on their perceptions of class context, and on the particular motivational goals and learning strategies they tend to adopt (Biggs, 1987; Biggs et al., 2001). Personological presage variables in the present model include demographics such as age, socio-economic status, gender, and English proficiency, as well as two self-belief factors: *Subject self-concept* and *Honesty-trustworthiness self-concept*.

In the following subsections, the variables and measures used to populate each of the hypothesized PTLC model's four components (behavior, moral obligation, context, and person) are introduced and justified. For the sake of clarity, these variables, and related hypotheses, are explained from right to left across the model, or 'backwards', from cheating and learning strategies, to their moral, situational, and personological antecedents. All questionnaire items are presented in Appendices B and C.

3.4.1 Behavioral variables

Self-reported cheating. The present study extends a body of self-report-based research on academic integrity conducted by Murdock and various colleagues (2001, 2004, 2007, 2008), and Anderman and various colleagues (1998, 2004, 2010) that has focused on achievement goals and goal structures, in addition to learning context quality factors at the secondary school level. Cheating is measured in the present work with a three-item self-report scale developed by Midgley et al. (2000) that was used in secondary-level studies conducted by both Murdock et al. (2001) and Brown-Wright et al. (2013). This three-item measure is augmented with a single item from a conceptually similar scale used by Anderman and colleagues (1998, 2004, 2010): "I have cheated on my Science work this year" (see Appendices B and C)

While a number of scholars have pointed out the likelihood that self-report measures underestimate actual cheating due to socially acceptable responding (Miller, Shoptaugh, & Parkerson, 2008; Walker, 2010), others report that students are generally willing to admit cheating on questionnaires that are anonymous (McCabe, 2005; Rettinger & Kramer, 2009). Self-report scales appear, in fact, to be the only means available for measuring cheating by students on real assessments, in real class contexts, over the course of an academic year.

Learning strategy. Approaches to learning involve both motivational components and strategic components, namely *surface* motivation *vs. surface* strategy, and *deep* motivation *vs. deep* strategy (Biggs et al., 2001). In view of the relatively large amount of research, and inconsistency of findings, on how academic integrity relates to students' achievement goal motivations (see section 2.3.4), the present study focuses exclusively on the strategic components of deep and surface approaches to learning.

Surface learning strategies, by which students seek to obtain grades with minimal intellectual effort, are measured with a six-item instrument developed by Simon et al. (2004) as an aspect of the broader notion of disintegrity (Miller et al., 2011). Deep strategies reflect, by contrast, that students seek to truly understand, make personal connections to, and master academic material. Deep learning, measured with a seven-item instrument developed by Anderman et al. (1998), has *integrity* in the sense that it involves the construction of genuine personal meaning, and the integration of that meaning with a learner's background knowledge (Ramsden, 1992). Surface and deep learning strategies are positioned in the PTLC model, therefore, as behavioral correlates of *Self-reported cheating* under 'behaviors' (see Figures 3.6 and 3.8).

Hypothesis 3: *Surface learning strategies* and *Self-reported cheating* will, as forms of disintegrity, share a large positive correlation with each other, and large negative correlations with *Deep learning strategies*.

3.4.2 Moral obligation variables

The hypothesized model is predicated on the claim made by domain theory (Turiel, 1983, 2002, 2006) that rules may be viewed as either moral or conventional, and that children and adolescents may shift between these views in reference to rules in a given context. PTLCs offer a framework for the judgments that underlie such shifts, insomuch as students believe such judgments are valid. Students may interpret a teacher's failure to fulfill his or her PTLC obligations (e.g. pedagogical effort and skill) as a type of cheating, in the sense that social contract violations are often interpreted as cheating (Cosmides & Tooby, 2013). Students who feel cheated by a teacher may reciprocally feel lower moral obligation to respect that teacher, the content of his or her class, and the legitimacy of his or her rules. Students may feel, in other words, that when teachers do a poor job, academic cheating becomes more justifiable. 'More justifiable', in this sense, is equivalent to 'less immoral', which implies that the obligation to be honest has shifted from the moral domain to the conventional domain.

Justifiability of cheating. Neutralization techniques are, like 'defenses to crimes' (Morawetz, 1986; Sykes & Matza, 1957), by which violations of law are defended in legal settings, dichotomized into excuses and justifications (Brent & Atkisson, 2011; Morawetz, 1986; Scott & Lyman, 1968). Justifications, such as reasons for cheating that correspond to the neutralization category 'condemnation of the condemners' (see Table 2.1), are of principal interest to this study. An act is justifiable if it is right, reasonable, or defensible (*New Oxford American Dictionary*), even if it may otherwise be designated as delinquent or criminal (Morawetz, 1986). Inasmuch as academic honesty is morally imperative, therefore, justifications for cheating cannot be valid. That academic honesty is understood by most students to be morally imperative is a key assumption of the neutralization view of cheating that leads to the expectation that cheating behaviors should be incongruent with cheaters' abstract moral beliefs. Conversely, if justifications for cheating are valid in the mind of a student, then honesty, and the rules that forbid dishonesty, are not morally imperative. Thus, the four-item measure of *Justifiability of cheating* adapted in the present study from Murdock et al. (2004) is, in effect, a measure of the degree to which honesty is felt to be morally imperative. To the extent that cheating is seen to be justifiable, rules against it must be viewed as conventional, rather than moral. The overarching PTLC hypothesis holds, moreover, that shifting from a moral to a conventional view of rules reflects diminished commitment on the part of students to their PTLC obligations, which may come about when they feel cheated in academic contexts that fail to meet their expectations for pedagogical, assessment, social, and regulatory quality. *Justifiability of cheating* is hypothesized, therefore, to mediate the effect of perceptions of class quality on academic integrity (see Figures 3.6 and 3.8).

Hypothesis 4: Students will judge cheating to be more justifiable in class contexts that fail to meet their expectations for quality. Perceptions of class quality will negatively predict *Justifiability of cheating*.

Hypothesis 5: Students who judge cheating to be more justifiable will report more disintegrity, and less use of deep learning strategies. *Justifiability of cheating* will positively predict *Self-reported cheating* and *Surface learning strategies* and negatively predict *Deep learning strategies*.

Appropriate workload. Appropriate workload will be measured in the present study using a five-item scale from the *Course Experience Questionnaire* (CEQ) (Wilson et al., 1997), which has been used to study surface and deep learning strategies in a large body of previous research associated with student learning theory (Biggs & Tang, 2011), to be discussed in more detail in section 3.4.3. While the appropriateness of workload has traditionally been seen as a contextual predictor of both cheating (Smith et al., 1972; Jurdi et al., 2011a) and surface learning strategies (Ramsden & Entwistle, 1981; Wilson et al., 1997; Diseth, 2007), it will be positioned in the hypothesized model as a dimension of moral obligation that reflects student commitment (see Figures 3.6 and 3.8).

As argued in the literature review (see section 2.4.4), an association between excessive workload and disintegrity may reflect either students' concerns about their *preparedness* to succeed in a given class (Evans & Craig, 1990a; Galloway, 2012; Sisti, 2007; Zito & McQuillan, 2011), diminished self-control as a function of cognitive overload (Greene et al., 2008), or low *commitment* to a particular learning context (Curry, 1984; Kember, 2004). This highlights the fact that perceptions of whether workload is appropriate may entail a more complex set of considerations than the sheer volume of required work (Kember, 2004). Higher-order factor analyses of tertiary students' course evaluation data, collected using the CEQ, have indicated, for instance, that the appropriateness of workload is psychometrically distinct from environmental variables that are related to class quality, such as *Good teaching* and *Appropriate assessment* (Richardson, 1994; Trigwell & Prosser, 1991; Wilson et al., 1997). Whether the workload in a given class is perceived as 'appropriate' may, in fact, reflect the amount of time and effort that a student believes the class is worth. If a student's ability and level of commitment are uniform across all classes, the appropriateness of workload should vary directly with the amount of time the work takes to complete. If every class required one hour of homework per week, they would all, by this view, be equally appropriate. If, however, a student finds certain subjects more conceptually challenging, as most students do, then appropriateness of workload should also vary according to ability. This is referred to here as the 'ability' aspect of appropriate workload. Students with lower aptitude for a given subject will have to devote more time and effort in order to achieve the same amount of success as students with higher aptitude, which may lead to a perception that the workload of more challenging classes is heavier, and therefore less appropriate.

Appropriateness of workload may also have a 'commitment' aspect. Students who relate more positively to a given class may feel more committed to it, and may perceive the workload as being more appropriate than students who feel less committed. This, in addition to aforementioned findings in prior research that *Appropriate workload* is psychometrically distinct from measures of class quality (Wilson et al., 1997), advocates for hypothesizing its relationship to cheating in terms of *commitment*, as a proxy for the moral obligation to work hard. The 'commitment' aspect of whether a student believes the workload in a given class is appropriate is, moreover, isolated from the 'ability' aspect of *Appropriate workload* in the hypothesized model, by controlling for *Subject self-concept*, a measure of self-perceived ability (see section 3.3.4).

Hypothesized as an aspect of moral obligation, *Appropriate workload* is held to reflect students' perceptions of how much work is appropriate to do for a given class. In terms of the PTLC framework, *Appropriate workload* is held to represent the amount of effort that is morally compelled by a student's PTLC for a given class, which is hypothesized to fluctuate directly with how positively the student perceives contextual elements of the class, i.e. that constitute a useful, well-taught, well-managed, and fair academic experience. A class perceived to be useless, badly-taught, poorly-managed, and generally unfair should, by contrast, compel less effort, implying a shift from a moral to a conventional view of whether students should strive for meaningful learning in a given class. A conventional view of the duty to strive for meaningful learning might entail that effort should be minimized, i.e. surface learning strategies should be adopted to the extent that meaningful understanding is not explicitly required on assessment tasks. *Appropriate workload* will serve in the PTLC model, therefore, as a mediator for the effects of contextual factors on cheating and learning strategy.

Hypothesis 6: Students will perceive the workload to be less appropriate in class contexts that fail to meet their expectations for quality. Perceptions of class quality will positively predict *Appropriate workload*.

Hypothesis 7: Students who perceive the workload in a given class to be less appropriate will report more disintegrity, and less usage of deep learning strategies. *Appropriate workload* will negatively predict *Self-reported cheating* and *Surface learning strategies* and positively predict *Deep learning strategies*.

3.4.3 Learning context variables

The central role that justifications and excuses play in cheating behavior is well recognized in the literature (e.g. Day et al., 2011; Murdock et al., 2001, 2004). Students are often observed to justify cheating by blaming class context factors such as teaching quality, assessment quality, and interest (Galloway, 2012; Olafson et al., 2013; Stephens & Gehlbach, 2007; Taylor et al., 2002; Zito & McQuillan, 2011). Such justifications have been almost uniformly dismissed as techniques of neutralization (e.g. Diekhoff et al., 1996; Galloway, 2012; Haines et al., 1986; LaBeff et al., 1990; Murdock & Stephens, 2007; Rettinger & Kramer, 2009), despite being corroborated by correlational and experimental evidence. Higher incidence of cheating is statistically associated with assessments characterized as 'high stakes' (Jensen et al., 2002; Nichols & Berliner, 2007), or that are perceived as inauthentic or poorly designed (Heckler et al., 2013); with teacher quality perceived as poor (Anderman et al., 2010; Evans et al., 1993; Murdock et al., 2001; 2004; Shipley, 2009; Stearns, 2001); with classes and learning tasks perceived as useless and/or un-interesting (Baird, 1980; Ma et al., 2007; McCabe et al., 2002; Sisti, 2007; Schraw et al., 2007); with the perception that cheating is the norm among peers or classmates (Bowers, 1964; Burrus et al., 2007; Carrell et al., 2006; Eisenberg, 2004; Galloway, 2012; Gino et al., 2009; Hartshorne & May, 1928; McCabe & Treviño, 1997; McCabe et al., 2008; Nora & Zhang, 2010; O'Rourke et al., 2010; Teodorescu & Andrei, 2009; Walker et al., 1966); and with learning and assessment situations perceived as unfair (Brent & Atkisson, 2011; Evans & Craig, 1990a; Evans et al., 1993; Jensen et al., 2002; McCabe et al., 2001, 2008; Rettinger, 2007; Vowell & Chen, 2004).

Learner perceptions of context are organized into two parts in the PTLC model developed here (see Figure 3.6): (1) perceived fairness and quality, and (2) classroom goal structure. Each of these parts of the PTLC model includes dimensions of student experience that have been emphasized in research on both cheating and learning strategy. To answer calls for a more nuanced examination of the principal dimensions of learner experience such as teaching and assessment (Murdock et al., 2004; 2008), students' evaluations of fairness and quality are modeled with eight measures. The present model also extends integrity research related to achievement goal theory (Ames & Archer, 1988; Anderman et al., 1998; 2010), by including measures for mastery and performance classroom goal structure.

The *Course Experience Questionnaire* (CEQ) (Entwistle & Ramsden, 1983; Wilson et al., 1997), which has been developed and used principally in the field of student learning theory (Biggs & Tang, 2011), is a multidimensional measure of students' course evaluations that originated in a set of grounded exploratory studies at Lancaster University, UK (Entwistle & Ramsden, 1983). The CEQ is a theoretically coherent framework for student experience that has been related empirically to surface and deep learning strategies for more than thirty years, and is currently one of the most widely used measures of tertiary course experience in the world (Marsh, Ginns, Morin et al., 2011; Richardson, 2005). Factors measured by the CEQ such as *Good teaching*, *Clear goals and standards*, and *Appropriate assessment* are also aspects of academic experience that students often blame when they justify cheating, as outlined at the beginning of this subsection. This set of CEQ factors is augmented in the present work by one measure of the perceived usefulness of the curriculum of a given class, i.e. *Usefulness of* *curriculum* (Rowe & Hill, 1998), as a proxy for student interest, as well as of two measures of assessment context: *Transparency* and *Authenticity* (Dorman & Knightley, 2006).

CEQ in a secondary setting. The present study adapts the wording of CEQ scales to the experience of secondary school students. The original CEQ items (Wilson et al., 1997), are compared to their modified equivalents used in the present study in Appendix C. Table 3.1, below, broadly summarizes the evolution of the modern CEQ and demonstrates the consistency of its constructs over time. A vast, varied, and current body of research supports the validity and reliability of the CEQ at the tertiary level (e.g. Ginns et al., 2007; Lawless & Richardson, 2002; Ning & Downing, 2010; Ramsden, 1991; Richardson, 1994, 2005; Wilson et al., 1997). Modified CEQ scales have also been adapted to the secondary level in the form of the *School Experience Questionnaire* (SEQ) (Ramsden, Bowden & Martin, 1988; Ramsden, Martin, & Bowden, 1989). As can be seen in Table 3.1, however, reliability estimates for three of the four SEQ scales reported by Ramsden et al. (1988) fall below .70. Reliability estimates of .70 - .80 are widely considered to be "adequate" for purposes of structural equation modeling (Kline, 2011, p. 70), by which standard the reliability of these SEQ scales is *in*adequate for structural equation modeling. CEQ measures (see also Table 3.1), which are analogous to the SEQ scales in question, but show substantially better reliability, will be used for the present study.

Teacher quality. The eight-item CEQ measure *Good teaching* examines the degree to which students perceive their teacher as supportive and able to deliver lessons effectively. Teacher quality has been associated in prior work with self-reported cheating (Anderman et al., 2010; Murdock et al., 2004, 2007; Shipley, 2009), the perceived likelihood of cheating (Evans & Craig, 1990a, 1990b; Evans et al., 1993), the acceptability and justifiability of cheating (Anderman et al., 1998; Day et al., 2011; Murdock et al., 2004, 2007). Teacher quality has also been associated with whether students use deep or surface learning strategies (Ramsden, 1991; Wilson et al., 1997; Diseth, 2007; Ning & Downing, 2010). Teachers who provide insufficient support for learning, who are perceived as unhelpful at explaining material, or as unable to guide students through potentially confusing ideas, may leave students feeling 'on their own' at overcoming academic challenges. Students may view poor teaching as a failure on the part of a teacher to fulfill his or her socially and professionally implied obligations. In viewing the student-teacher relationship as a social contract, the perception that a teacher neglects students, or undermines students' efforts to succeed, may be interpreted as cheating on the part of the teacher (Cosmides & Tooby, 2013) that obviates the moral obligation students feel to work hard and be honest.

Clear goals and standards. The five-item CEQ measure *Clear goals and standards* pertains to the clarity of the specific purposes of work in a given class, the study requirements of its particular curriculum, and the criteria by which student performance will be assessed (Entwistle & Ramsden, 1983; Ramsden, 1991). The clarity of goals and standards is associated with both learning strategy (Diseth et al., 2010; Wilson et al., 1997) and cheating in secondary education settings (Brent & Atkisson, 2011; Evans & Craig, 1990a). Academic goals and standards express the intended purpose and meaning of student effort, and guide students' judgments of their own progress. Learning contexts perceived as lacking clear goals and standards may convey insufficient benefit or undue harm to students by failing to inform them of how to direct their efforts to succeed at required work. Students may feel, as such, that when goals and standards are unclear, work requirements are invalid and can be flouted without violating moral imperatives.

Usefulness of curriculum. The four-item measure *Usefulness of curriculum* (Rowe & Hill, 1998) is used to replace the SEQ measure *Preparation for study in higher education*, which achieved a reliability estimate of just .58 in Ramsden et al. (1988) (see Table 3.1). The *Usefulness of curriculum* measure was developed to query secondary student perceptions of the value of the curriculum of a given class. It evinced good reliability (.86) in the study reported by Rowe and Hill (1998), and is employed in the present study as a proxy for students' overall interest in Science class, the research setting for the present work. Intrinsic interest has been seen to characterize deep learning since the earliest days of student learning theory (Fransson, 1977; Marton, 1976), and has also been connected to higher levels of engagement and lower levels of cheating (Shraw et al., 2007). Viewing the subjects or topics covered in a class as 'useless' suggests a pronounced lack of interest. Being required to learn material perceived as useless also violates the basic assumption that formal education is meant to be helpful. Students may feel it is unjust to have to learn what they perceive to be useless, not merely because it is non-beneficial, but because it wastes their time, which, in social contract terms, cheats the fundamental expectation that one's education should promote student welfare. Students who feel cheated by useless learning experiences may take the conventional view that disintegrity is justifiable in order to achieve the grades by which they will nonetheless be judged to have succeeded or failed.

Measures of learning context factors such as these (teacher quality, goal clarity, and curriculum usefulness) are employed in the present study to query respondents' perceptions of the quality of their Science classes. A significant amount of research reviewed above suggests that how well or poorly students perceive learning contexts influences their cheating. The contractarian perspective asserted here holds that moral obligation is an important mechanism by which this influence is exerted. Moral obligation is hypothesized to mediate the relationship between student perceptions of class quality and cheating behavior, as shown above in Figures 3.5 and 3.8 (see also hypotheses 4 – 7).

Hypothesis 8: Students who perceive teacher quality as low, goals and standards as unclear, and the curriculum as useless in a given class will be more likely to engage in disintegrity, and less likely to engage in deep learning strategies. *Good teaching, Clear* *goals and standards,* and *Usefulness of curriculum* will negatively predict *Self-reported cheating* and *Surface learning strategies* and positively predict *Deep learning strategies*. These predictive effects will be at least partially mediated by *Justifiability of cheating*.

Table 3.1 presents sample items and alpha reliabilities of measures used on various versions of the CEQ since its development (Entwistle & Ramsden, 1983). The CEQ serves, in the present study, as a theoretically coherent measure of the principal dimensions of students' evaluations of class context.

Table 3.1

Development of the Course Experience Questionnaire over the past three decades.

Course Perceptions Questionnaire Entwistle & Ramsden, 1983, p. 124	**Student Experience Questionnaire (SEQ)** Ramsden,et al., 1988, p.4	**CEQ** Ramsden, 1991, p.134	**CEQ** Wilson et al., 1997	**SCEQ** Ginns et al., 2007, p.605	**This study**
Relationships with students: *closeness of student/lecturer relationships; help and understanding show to students.*	Teaching support: *the extent to which pupils think the teaching they experience is supportive of their learning* (α = .81)	Good teaching: *clarity of explanation, level at which material pitched, enthusiasm and help with study problems* (p. 132) (α = .87)	Good teaching: *same as Ramsden (1991)* (α = .86-.88)	Good teaching: *same as Ramsden (1991)* (α = .83)	Good teaching: *Wilson et al. (1997)*
Workload: *pressure placed on students in terms of demands of the syllabus and assessment tasks*		Appropriate workload: *The sheer volume of work to be got through in this course means you can't comprehend it all thoroughly (neg.)* (α = .77)	Appropriate workload: *same as Ramsden (1991)* (α = .74-.75)	Appropriate workload: *same as Ramsden (1991)* (α = .76)	Appropriate workload: *Wilson et al. (1997)*
Clear goals and standards: *extent to which standards expected of students are clear and unambiguous*	Structure, climate, and cohesiveness: *the extent to which goals are clearly defined and pupils and staff share similar aims* (α = .64)	Clear goals: *You usually have a clear idea of where you're going and what's expected of you in this course* (α = .80)	Clear Goals and standards: *same as Ramsden (1991)* (α = .82)	Clear goals and standards: *the staff made it clear right from the start what they expected from students.*(α = .80)	Clear goals and standards: *Wilson et al. (1997)*
	Formal achievement: *the extent to which pupils feel they are bring encouraged to perform highly in external examinations* (α = .68)	Appropriate assessment: *Staff here seem more interested in testing what we have memorised than understood (neg.)* (α = .71)	Assessment: *same as Ramsden (1991)* (α = .73-.74)	Appropriate assessment: *same as Ramsden (1991)* (α = .72)	Appropriate assessment: *Wilson et al. (1997)*
Freedom in learning: *amount of discretion possessed by students in choosing and organizing academic work*	Independence in learning: *the perceived stress on developing the capacity to learn independently* (α = .64)	Emphasis on independence: *Students here are given a lot of choice in the work they have to do* (α = .72)	Independence: *same as Ramsden (1991)* (α = .67-.68)		
Vocational relevance: *perceived relevance of courses to students' careers*	Preparation for study in higher education: *the extent to which pupils feel they are being prepared for learning in higher education* (α = .58)		Generic Skills: *the extent to which graduates perceive their courses as developing a number of generic skills and abilities (p.36)* (α = .79-.80)	Generic skills: Same as Wilson et al. (1997) (α = .77)	Usefulness of curriculum: *Rowe & Hill (1998)*

Classroom goal structure. Achievement goal theory has been studied in relation to cheating at three hierarchical levels in educational settings: the school, the classroom, and the individual student. The most consistent and significant association with cheating across much of the research upon which the present study builds has been at the classroom level, i.e. classroom goal structure (Anderman et al., 1998; Anderman & Midgley, 2004; Murdock et al., 2001, 2004; Stephens & Gehlbach, 2007). Achievement goal theory traditionally portrays classroom goal structure as comprising the key factors by which a teacher directs students towards either mastery goals or performance goals. More than a decade of research has indicated that cheating tends to be positively related to perceptions of performance goal structure (e.g. Anderman et al., 1998), and/or negatively related to perceptions of mastery goal structure (e.g. Stephens & Gehlbach, 2007). Performance goal structures are held to influence students to orient themselves to performance goals (Meece et al., 2006), which, as reviewed in section 2.3.4, have been associated with higher levels of cheating (Anderman et al., 1998; Koul et al., 2009; Murdock et al., 2001; Olafson et al., 2013; Rettinger & Cramer, 2009; Stephens & Gehlbach, 2007) and surface learning strategies (Fellonar et al., 2007). Mastery goal orientations have, by contrast, been associated negatively with cheating (Anderman & Midgley, 2004; Murdock et al., 2001), and positively with deep learning strategies (Fellonar et al., 2007; Phan, 2008, 2009a, 2009b).

Neither type of classroom goal structure is hypothesized to represent a breach of students' PTLCs within a given class context. While many scholars have concluded, through decades of research, that performance goals are less desirable than mastery goals, it is not clear why students would, with any uniformity, perceive one goal structure as more moral than the other. Goal structures do, however, convey "messages about the purposes of instruction" (Anderman & Midgley, 2004, p. 501) that could affect whether students believe they are expected to *learn* as opposed to *earn grades*. Achievement goal structures might, in other words, influence what duties students feel morally obliged to undertake in a given class. The emphasis on grade-achievement that characterizes performance goal structures may, in particular, convey to students that disintegrity is more justifiable if it produces better grades.

Perceptions of mastery and performance goal structures, measured respectively in the present study by two, five-item scales developed by Midgley et al. (2000), were hypothesized to exert effects on self-reported cheating and surface learning strategies that were at least partially mediated by moral obligation (see Figures 3.6 and 3.8). Performance goal structures may increase the likelihood of cheating purely by intensifying competition for good grades, whereas mastery goal structures may be antithetical to cheating by encouraging a self-referential view of achievement.

Hypothesis 9 - 10: Students in performance goal-oriented class contexts will be encouraged to view cheating as more justifiable and workload as less appropriate. (9) *Performance goal structure* will positively predict *Justifiability of cheating*, and (10) negatively predict *Appropriate workload*.

Hypothesis 11 - 12: Students in mastery goal-oriented class contexts will be encouraged to view cheating as less justifiable and workload as more appropriate. (11) *Mastery goal structure* will negatively predict *Justifiability of cheating*, and (12) positively predict *Appropriate workload*.

Hypotheses 13 - 15: Students who perceive a performance goal structure in Science class will engage in more disintegrity and less deep learning, (13) *Performance goal structure* will positively predict *Self-reported cheating*, and (14) *Surface learning strategies*; and (15) will negatively predict *Deep learning strategies*. These effects will be at least partially mediated by moral obligation.

Hypotheses 16 - 18: Students who perceive a mastery goal structure in Science class will engage in less disintegrity and more deep learning. (16) *Mastery goal structure* will negatively predict *Self-reported cheating*, and (17) *Surface learning strategies*; and (18) positively predict *Deep learning strategies*. These effects will be at least partially mediated by moral obligation.

3.4.4 Assessment context variables

Assessment is included in the hypothesized model as an extension of the learning context that, itself, comprises multiple dimensions including appropriateness, authenticity, and transparency. The term *assessment* refers herein to teacher-evaluated assignments, where evaluation may take such forms as verbal or written comments, marks along continua, check marks, or letter grades. Assessments by this definition include homework assignments, papers, projects, portfolios, quizzes, tests and, where applicable, behavioral criteria such as participation and citizenship.

The power of assessment to affect learning strategies has long been recognized by student learning theorists at the tertiary level (Segers & Dochy, 2006; Struyven, Dochy & Janssens, 2002). At the secondary level, a similar body of argument and evidence for the broad and potent influence of assessment on student experience and learning outcomes has coalesced in the literature of assessment *for* learning (AFL) (Black & Wiliam, 2006a, 2006b; Harlen, 2006). AFL has a strong tradition of considering assessment in multidimensional terms. Crooks (1988) concludes his review of assessment research with an appeal for assessment practices that (1) utilize timely feedback, (2) expressly facilitate student progress, and (3) encourage deep learning by emphasizing understanding, transferable learning, and thinking skills. More recently, Dorman and Knightley (2006) added to the AFL literature a five-dimensional model of student perceptions of assessment: 1. congruence with planned learning, 2. student consultation, 3. diversity, 4. authenticity, and 5. transparency. As shown in Appendices B and C, the present study uses the latter two of Dorman and Knightley's (2006) scales, in addition to the CEQ scale for *Appropriate assessment* (Wilson et al., 1997), in order to measure a nuanced picture of student perceptions of assessment.

Appropriateness. The four-item CEQ scale *Appropriate assessment* measures students' perceptions of how much freedom of intellectual self-expression and self-determination is permitted by the assessment methods in a given class. Assessments that emphasize rote memorization and 'regurgitation' are inappropriate by this conception; appropriate assessments entail higher-order thinking and acknowledge individual understanding. By denying students opportunities to think independently and demonstrate the fullness of their intellectual accomplishment, fact-focused and/or highly directive assessment tasks may characterize achievement in a manner that seems unjustly narrow. Students who feel that their achievement is under-represented, i.e. who feel 'cheated' of due credit, by assessment tasks perceived as inappropriate may feel that it is justifiable to reciprocate by representing themselves in dishonest ways.

Transparency. Drawing from Dorman and Knightley (2006), the seven-item measure *Transparency of assessment* is hypothesized to extend to the assessment context the dimension of *Clear goals and standards* from the learning context (see Figure 3.7), in that the goals of a given class are largely embodied by its assessments. Transparency reflects the "extent to which the purposes and forms of assessment tasks are well-defined and clear to the learner" (Dorman & Knightley, 2006, p. 52). A lack of transparency may give rise to the sense that an assessment is unfair, in that important information about it is not made available, such as what will be covered, how it will be conducted, and when it will take place. Students who feel that teachers are supposed to provide such information may feel that the 'rules of the game' change when teachers do not provide it. A student who feels that lacking such information unfairly prejudices assessments of his or her achievement may judge that disintegrity is a justifiable countermeasure for attaining a fair grade.

Authenticity: The term 'authentic assessment' describes assessments of knowledge, understandings, and skills that are perceived by students to be meaningful (Waldrip, Fisher, & Dorman, 2009). Authentic assessments signify accomplishment that is of genuine importance to students, as distinct from the importance of grades. The seven-item measure of *Authenticity of assessment* (Dorman & Knightley, 2006) extends to the assessment context the dimension of *Usefulness of curriculum* from the learning context (see Figure 3.7) (Rowe & Hill, 1998). Students who are genuinely convinced that an assessment task misses the point of what they have learned, is irrelevant to their lives, or has been assigned as 'busywork' may take a conventional view that it serves solely to produce a grade. In the same sense that surface learners pursue the sign over what is signified (Marton & Säljö, 1976), students who perceive a learning task to be inauthentic may judge disintegrity to be justifiable for the purpose of obtaining a good grade, i.e. the sign, because they do not respect what it signifies. Moral obligation is hypothesized, as such, to mediate the effects of assessment authenticity on disintegrity in Figures 3.5 and 3.8 (see also hypotheses 4 – 7).

Hypothesis 19: Students who perceive the assessment context in a given class to be inappropriate, inauthentic, and/or non-transparent will be more likely to engage in disintegrity, and less likely to engage in deep learning strategies. *Appropriate assessment, Authenticity of assessment,* and *Transparency of assessment* will negatively predict *Self-reported cheating* and *Surface learning strategies* and positively predict *Deep learning strategies*. These effects will be at least partially mediated by moral obligation.

Figure 3.7 presents the dimensions of classroom teaching and assessment contexts included in the present study. Alignment of the dimensions of the two contexts (dashed lines) is meant to indicate theorized relationships between constructs.

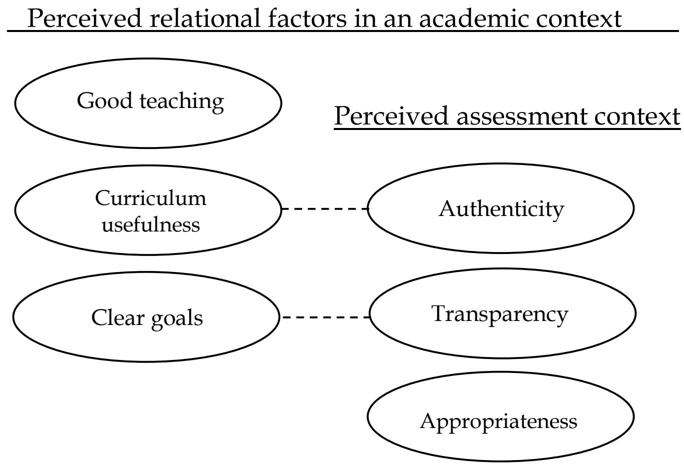

Figure 3.7. Perceived learning and assessment variables in the hypothesized model.

3.4.5 Social-regulatory context

Social and regulatory factors, such as the behavioral tone of a class, and whether rules are clear and appropriately enforced, are important non-academic aspects of the backdrop to learning and assessment in any academic context (Emmer & Stough, 2001). A large body of research, reviewed in section 2.4, indicates that students are more likely to cheat when classroom discipline is lax (e.g. Briggs et al., 2013; Houston, 1977; Tittle & Rowe, 1973; Whitley, 1998), and when their peers demonstrate supportive attitudes towards cheating (e.g. Magnus et al., 2002; Nora & Zhang, 2010; Schraw et al., 2007).

Peer norms. Peer norms play a prominent role in influencing cheating behavior (Carrell et al., 2008; Gino et al., 2009; McCabe & Treviño, 1997; Nora & Zhang, 2010; Stephens & Gehlbach, 2007; Walker et al., 1966). The seven-item scale *Peer norms*, developed by Mayhew et al. (2009), measures student perceptions of whether their peers believe cheating is justifiable. Peer norms perceived as favorable to cheating are hypothesized, according to social comparison theory (Festinger, 1954), to communicate un-favorable contextual perceptions and to define cheating as an appropriate in-group behavior. *Peer norms* is positioned, therefore, as a mediator in the hypothesized model, between perceptions of class quality, and the justifiability of cheating (see Figure 3.8).

Social comparison theory (SCT) (Broeckelman-Post, 2008; Festinger, 1954), reviewed in section 2.4.5, asserts that opinions become increasingly uniform within groups as people appraise their own opinions against those of their peers. Individuals whose opinions differ from those of their peer group find themselves pressured either to bring their own opinions into line with the group, to change their peers' opinions, or to leave the group. Evaluations by students of a given learning context arise, by this view, at least partly as a matter of group consensus. Several scholars have suggested similarly that peer norms mediate the influence of student experience upon determinations of whether or not to cheat (Teodorescu & Andrei, 2009; Whitley, 1998). The context-specific moral flexibility associated with peer norms (McCabe & Katz, 2009) may, therefore, reflect the degree to which students think their peers have favorable attitudes towards cheating.

Students see through their own eyes, judge according to their own standards, and are undoubtedly responsible for their own actions. When they infer that their peers think cheating is acceptable because a class is unfair or of low quality, however, they may themselves adopt more negative opinions of the class, and view cheating as more justifiable as a result. *Peer norms* is positioned, therefore, as a mediator of the effects of perceived class context on *Justifiability of cheating* and *Self-reported cheating* (see Figure 3.8).

Hypothesis 20: Peer norms related to cheating will mediate the influence of perceived class context on whether individual students judge their own acts of cheating to be justifiable.

An individual's perceptions of class context factors will negatively predict *Peer norms*, which will, in turn, positively predict *Justifiability of cheating*.

Rules. A scale for 'experience of school rules', developed for secondary education research (Gregory, Cornell, Fan, et al., 2010), is adapted to classrooms in the present study. The modified six-item scale measures students' perceptions of rules as clear, fair, and enforced effectively. Educators, like parents, exercise authority over students for the purpose of protecting and promoting student welfare (Baumrind, 1987). Rules that seem to make no sense, or that seem to be applied in harsh or arbitrary ways, may be viewed as failures of a teacher or a school to exercise authority for the benefit of students. Students may feel that rules applied in an inconsistent, spiteful, or arbitrary manner in a given class cease to be morally legitimate, and may furthermore reject the notion that they have a moral obligation to heed such rules (Thomson & Holland, 2002). When adolescent students feel that rules are inappropriate, unclear, or applied unjustly, they may come to view them as conventional in character (Murdock & Stephens, 2007; Thornberg, 2008), and not, therefore, morally binding.

A second important aspect of students' experience of classroom rules is whether the rules are enforced effectively enough to create a sense of real risk associated with cheating. While breaking moral imperatives carries internal consequences (Aronson, 1968; Blasi, 1980; Mazar et al., 2008), the risks of breaking conventional rules are external (Turiel, 1983, 2006). When students feel alienated in a given class, external consequences may be the last line of defense for the integrity of assessment processes. Students who view rules as conventional are more likely to break them when they are poorly enforced. Higher rates of cheating due to poor enforcement of rules in a given class may lead, in turn, to a 'contagion' effect by signaling to classmates that cheating is easy to get away with and potentially necessary in order to compete (Gino et al., 2009; Walker et al., 1966). The experience of classroom rules is hypothesized, therefore, to affect cheating both indirectly, as a function of whether students view rules as morally legitimate in a given class context, i.e. as mediated by moral obligation (see Figures 3.5 and 3.8, and hypotheses 4 – 7), and directly, as a function of the perceived risk associated with cheating.

Hypothesis 21: Students will cheat more when they perceive that rules are not enforced effectively, especially if they view rules against cheating as conventional. *Experience of classroom rules* will negatively predict *Peer norms*, *Justifiability of cheating*, and *Self-reported cheating*.

3.4.6 Person

Demographic variables and stable intra-psychic factors are mainstays of academic integrity literature (Anderman & Murdock, 2007), underscoring broad consensus behind the idea that individual differences interact with context to produce behavioral outcomes (Mischel, 2004). Intra-psychic factors related to cheating in the hypothesized model include two domain specific measures: *Subject self-concept* and *Honesty-trustworthiness self-concept*, in addition to six demographic variables: age, grade-level, gender, language(s) spoken in the home, and English language proficiency.

Age and grade-level. While cheating behavior is seen to vary somewhat predictably with age, the nature and extent of the relationship is bound up with a wide range of correlates that age shares with grade-level (Anderman & Midgley, 2004; Miller et al., 2007), such as physical and cognitive development, and changes in educational goals (Franklyn-Stokes & Newstead 1995; Newstead et al., 1996). In the present study, age and grade-level may serve as control variables for distinguishing between patterns of change rooted in physiological growth and maturation, and/or in educational context. Grade-level is viewed, in particular, as a potential source of factorial non-equivalence (Cheung & Rensvold, 2002) in the measurement model for Figure 3.8, based on the findings of Anderman & Midgley (2004) related to changes in cheating behavior occurring over the Grade Eight – Grade Nine transition.

Gender. While research on the relationship between gender and academic integrity has generated mixed results, the most common finding is that females tend to look less favorably on cheating than males (Miller et al., 2007). A meta-analysis of 48 integrity studies conducted by Whitley et al. (1999) found an effect size of d = .21 for gender and cheating-related attitudes, but a mere d = .08 for gender and actual cheating behavior. The present study treats gender as a potentially important independent variable for understanding personological patterns in academic contexts, and, like grade-level, as a potential source of factorial non-equivalence in the measurement model for Figure 3.8.

English-language proficiency. Respondents will be asked to indicate their proficiency with English, as either *fluent, high, intermediate, low*, or *beginner*. Respondents' ability to understand survey items may affect the meaning of their responses. Beginning English speakers might, for instance, provide idiosyncratic answer patterns because they misunderstand nuanced language. Respondents who rate themselves at a *beginner* or *low* level of English may need to be screened-out in order to clarify results.

Subject (Science) self-concept. The five-item scale *Subject self-concept* (Marsh, Ellis, Parada, et al., 2005), pertains specifically to the subject of Science, in which cheating has often been found to be more common than in other subject areas (Meade, 1992; Miller et al., 2007; Murdock et al., 2001; Newstead et al., 1996; Schab, 1991).

Subject-related self-concept, formed through self-evaluations of past performance, social comparison, and self-judgment (Bong & Clark, 1999; Bong & Skaalvik, 2003), reflects the ways in which students *relate* to a given subject area (Kornilova, Kornilov, & Chumakova, 2009). A student's belief in the likelihood of his or her performing successfully on academic tasks appears to be a key influence on cheating behavior, as illustrated by the fact that 90% of the students surveyed by Nora and Zhang (2010) indicated that they would not cheat if they felt confident of success. The one study to have investigated the relationship between self-concept and academic integrity at the secondary school level (Rost & Wild, 1994) reported a significant negative correlation (r = -.27, p < .05) in a population of 197 German high school students. The authors interpreted this effect to represent an affinity among students with high academic self-concept for "attributing their successes to internal sources" (p. 129), and a corresponding tendency among students with low academic self-concept to cheat as a means of coping with performance-related anxiety. *Subject self-concept* is hypothesized in the present work to positively predict favorable perceptions of class context, adaptive moral judgment, and more honest behavior.

Hypotheses 22 – 25. Subject self-concept will positively predict (22) perceptions of academic relational variables (i.e. excluding *Experience of school rules* and *Peer norms*), (23) perceptions of *Mastery goal structure*, (24) *Appropriate workload*, and (25) the use of *Deep learning strategies*.

Hypotheses 26 - 29. Subject self-concept will negatively predict (26) perceptions of *Performance goal structure*, and (27) *Justifiability of cheating*, (28) use of *Surface learning strategies*, and (29) *Self-reported cheating*.

Honesty-trustworthiness self-concept. The ten-item measure *Honesty-trustworthiness self-concept,* developed for the *Self-Description Questionnaire II* (Marsh, 1992), investigates student self-perception in relation to integrity. While self-concept is viewed as being relatively stable across contexts, it tends to be shaped more readily by experience than by personality structures (Bong & Skaalvik, 2003). One's honesty-trustworthiness self-concept is expected, as a self-belief shaped by past-looking self-assessment, to embody the effects of past cheating behavior as well as moral identity, both of which have been prominent predictors of cheating behavior in prior research (Gino et al., 2011; Harding et al., 2012; Mayhew et al., 2009; Wowra, 2007b). The present study would be the first to use an integrity-related self-concept scale to directly investigate academic cheating at the secondary school level. Students with higher honesty-trustworthiness self-concept are hypothesized to be less likely to believe that cheating is acceptable among their peers, less likely to judge cheating to be justifiable, and less likely to cheat.

Hypotheses 30 - 32. *Honesty-trustworthiness self-concept* will negatively predict (30) the perceived amenability of peer norms to cheating, i.e. *Peer norms* (31) *Justifiability of cheating*, and (32) *Self-reported cheating*.

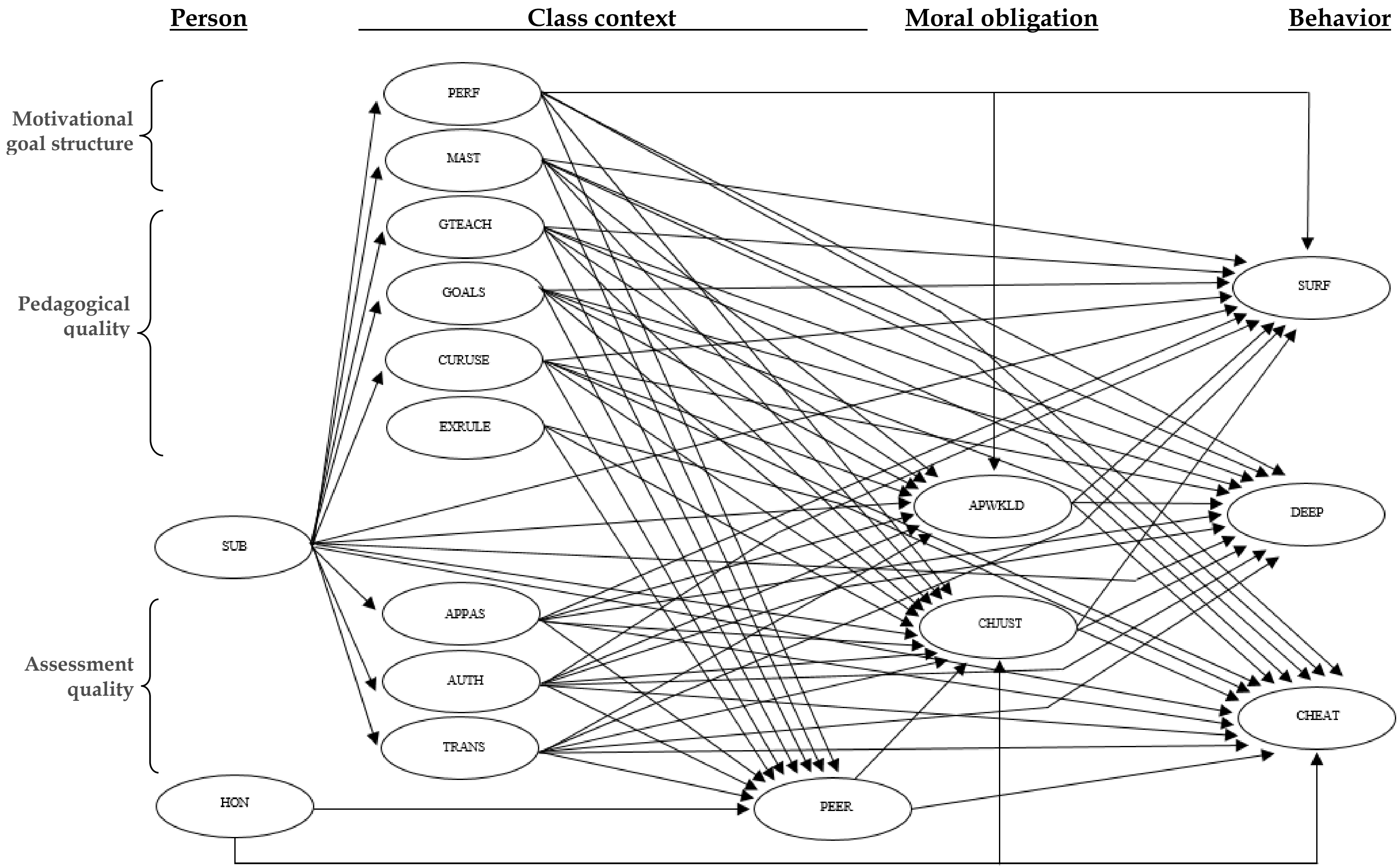

Figure 3.8. Model 1: The hypothesized PTLC structural model. SUB = Subject Self-Concept, HON = Honesty-Trustworthiness Self-Concept, PERF = Performance Goal Structure, MAST = Mastery Goal Structure, GOALS = Clear Goals and Standards, GTEACH = Good Teaching, APWKLD = Appropriate Workload, CURUSE = Usefulness of Curriculum, APPAS = Appropriate assessment; AUTH = Authenticity of Assessment, TRANS = Transparency of Assessment, EXRULE = Experience of Classroom Rules, PEER = Peer Norms Related to Cheating, DEEP = Deep Learning Strategies, SURF = Surface Learning Strategies, CHJUST = Justifiability of Cheating, CHEAT = Self-Reported Cheating.

The initial hypothesized PTLC model (Model 1). Figure 3.8 presents the initial multivariate PTLC structural equation model to be tested in the present research program. The model is organized into four broad components (person, class context, moral obligation, and behavior), which are populated, in total, with seventeen individual variables, among which seventy-five beta paths have been inserted to represent the thirty-two hypotheses given above. The complexity of this initial model is expected to diminish with the discovery of second-order factors that underlie contextual factors, as would be indicated by multicollinearity between their measures. CEQ scales (see Table 3.1) have, for instance, been found to produce a second-order structure for 'teacher quality' (Diseth, Palleson, Brunborg, & Larson, 2010; Richardson, 1994; Trigwell and Prosser, 1991; Wilson et al., 1997) that, if found in the present work, would allow the first-order measures to be related through that broader concept in the model. A second-order factor pertaining to 'assessment quality' is also anticipated, due to the inclusion of three assessment-related scales.

3.5 Chapter summary

The conceptual model for psychological teaching-learning contracts (PTLCs) was developed in this chapter. The overarching PTLC hypothesis was explicated in relation to domain theory (Richardson et al., 2012; Turiel, 1983, 2002, 2006; Thomson & Holland, 2002; Thornberg, 2008), and within the view that learning systems are characterized by dynamic equilibrium, as depicted in the 3-P Model (see Figure 3.1). The PTLC hypothesis was then posed as the basis of a four-category PTLC framework that abandons the assumption that students apply their abstract moral beliefs rigidly across contexts, as would be necessary for the BBI to accompany acts of cheating.

The PTLC hypothesis proffers an alternative account of cheating viewed as a conventional infraction that would not entail incongruity between moral beliefs and behaviors, but between moral beliefs and rules. Cheating acts viewed as conventional infractions by the students who commit them fall outside the scope of the BBI, and do not, therefore, need to be neutralized. The PTLC hypothesis is based, instead, on the assumption that humans innately understand and respond to reciprocal fairness (Cosmides & Tooby, 2013; Machery & Mallon, 2010), such that students may feel that cheating is justifiable when they think they have been cheated by teachers or teaching contexts. The perceived moral legitimacy of academic rules is, by this view, contingent upon the perceived moral legitimacy of academic contexts.

It was explained that the PTLC hypothesis marks a departure from the strictly rational-cognitive paradigm of moral psychology that has dominated much of the prior research on academic cheating. Adolescent cognition is, according to the rational-cognitive paradigm, rarely advanced enough for social contract-based judgment (Rest, 1986). The PTLC finds solace, instead, in the dual-process paradigm of moral psychology (Cushman et al., 2010; Haidt, 2007), which allows that social contract-based judgment may involve irrational and non-cognitive mental processes. Social contract-based judgment may, in fact, be an evolved function of the human sense of reciprocal fairness, which would explain why humans tend to interpret social contract violations as 'cheating' (Cosmides, 1989; Cosmides & Tooby, 2013). Inasmuch as PTLCs are social contacts, therefore, it is plausible to juxtapose, in a contractarian framework, the obligation students 'feel' to be honest against what they may feel to be cheating by teachers, schools, or educational contexts.

The general PTLC framework was next developed within the broader 3-P Model (Figures 3.1 and 3.5) based on the fundamental structure of a contract. It was argued that the failure of one party to fulfill contractual obligations could reciprocally relieve the counterparty from having to fulfill his or her corresponding obligations, such as refraining from certain types of behavior. The tendency to perceive and judge contract fulfillment, and to engage in disintegrity behaviors, was further held to depend on intrapersonal factors such as personality and self-beliefs. The mutual obligations of teacher and learner, background personological factors, and integrity-related behaviors were assembled into a four-category framework (Person, Class context, Moral obligation, and Behavior; see Figure 3.5)

Next, the hypothesized structural equation model to be tested in the present program of research was developed. This was done by populating each of the four components of the general PTLC framework with specific variables that have been emphasized in the literature on cheating. Measures of the specific variables included in the model were introduced and interrelated with hypotheses to be tested in the empirical phase of this research. The complete structural model was presented in Figure 3.8.

The purpose of the PTLC framework is to articulate a mechanism that has been suggested in recent scholarly works for how students adjudge the moral imperative to follow rules. A clearer picture of the role that contractarian judgment plays in cheating behavior is envisaged to suggest measures by which educators can further the cause of academic honor in their schools and classrooms.

CHAPTER 4

PILOT STUDY

4.1 Purpose

Two key purposes were addressed in the Pilot Study: (1) examination and, where necessary, modification of scale measures, and (2) identification of instances of multicollinearity and, where appropriate, either removal of constructs or creation of higher-order factors. Pilot Study data was collected in parallel with data for Time 1 of the Main Study. Schools selected for the Pilot Study were located in four separate countries on three continents, with the intent to represent the diversity expected in Main Study sample. None of the data analyzed in the Pilot Study was included in analyses conducted in the Main Study.

The questionnaire used to collect Time 1 data included seventeen measures drawn variously from the empirical literatures of both American secondary educational research and a mix of British and American tertiary educational research. Measures developed in secondary educational research contexts include *Subject* and *Honesty-trustworthiness self-concept*, *Performance* and *Mastery goal structure*, *Authenticity* and *Transparency of assessment*, *Usefulness of curriculum*, *Justifiability of cheating*, *Experience of classroom rules*, and *Self-reported cheating*; measures developed principally in British tertiary educational research contexts include *Good teaching*, *Clear goals and standards*, *Appropriate assessment*; and *Peer norms* was developed in an American tertiary research context. The wording of items that composed these measures was altered in numerous cases, in order to help convey their intended meaning to students at

American international secondary schools. While care was taken to preserve the original meaning of all items (see Appendix C), their psychometric validity and reliability could not be reasonably assured without first testing them on a sample similar to that used in the Main Study. The Pilot Study was conducted, therefore, in order to identify and correct factor misfit and multicollinearity in the measurement model, prior to the Main Study.

4.2 Participants

The Pilot Study sample (N = 96) included a mix of students aged 13 (16%), 14 (53%), 15 (29%), and 16 (2%), of which 64 were male (67%) and 32 were female (33%). These students were drawn from four private international schools in Europe (32%), East Asia (33%), and Africa (32%). 61% of respondents indicated that the predominant language in their home was not English.

Of the fifteen international schools that ultimately agreed to participate in the present research, the four chosen for the Pilot Study had, as a group, similar geographical distribution to the remaining eleven schools, which constituted the Main Study sample. This helped ensure that, for the sake of comparability, the Pilot and Main Study samples would entail similar ethnic and linguistic composition. Two of the schools in the Pilot Study were American international schools registered with the US State Department Office of Overseas Schools; one was a Japanese/English international school in Tokyo that follows a blend of the Japanese National Curriculum and the International Baccalaureate program (Doherty & Shield, 2012); and one was an all-boys boarding school that follows a modified British educational model.

All participants and their parents signed consent forms that expressly stated their willingness to participate in a Pilot Study that would involve a single instance of data collection. These consent forms, and attendant participant information forms, were, along with the overall design of the Pilot Study, approved by the University of Sydney Human

Research Ethics Committee (Protocol Number 14193; see Appendix AG). Participants filled out online questionnaires hosted by surveymonkey.com that were designed to require an answer to every question. Not answering a question triggered a prompt requesting an answer before proceeding to the next question. Participants could stop taking the questionnaire at any point, but could not proceed to the end without answering each question in sequence. Data screening thus entailed eliminating all but complete questionnaires, from which no data was missing.

4.3 Analysis

The psychometric properties of individual measures intended for the Main Study were assessed, as reported in this section, first with exploratory factor analysis (EFA), and then with confirmatory factor analyses (CFA). Problematic measures were modified.

4.3.1 Exploratory Factor Analysis

Exploratory factor analysis (EFA) is used to explore the latent structure of a data sample when the statistical relationships between observed variables are uncertain (Byrne, 2012). Most of the measures employed in this study have not been tested in prior research conducted in international secondary schools. The wording of many items designed for these measures has, moreover, been slightly modified to meet the specific aims of the present study (see Appendix C). EFA estimation was, therefore, conducted for each measure, using the program *FACTOR*, *version 7.00* (Lorenzo-Seva & Ferrando, 2006, 2007), to ensure that each measure represented a single, uni-dimensional factor. All EFAs were conducted with factor analysis using Pearson correlation matrices and parallel analysis (Horn, 1965) with marginally bootstrapped samples (Lattin, Carroll, & Green, 2003) to determine what number of factors should be extracted.

4.3.2 Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) is a methodology for assessing a hypothesized latent variable structure by comparing its covariance matrix to the covariance structure observed in the data. Constructs showing uni-dimensionality in EFA were analyzed in the confirmatory mode with the M*plus*, *version 7* software package (Muthén & Muthén, 2012), using maximum likelihood estimation with robust standard errors (Chou, Bentler & Satorra, 1991). This procedure was used in the Pilot Study to assess one-factor congeneric models (see Figure 4.1) according common fit indices (see Table 4.1).

Congeneric modeling, which is explained in more detail in section 5.5.2, was used, additionally, as a basis for creating weighted composite factor scores for all latent variables in the hypothesized model. Weighted composite scores reduce model complexity by converting respondents' answers on the multiple indicators of a given measure into a single weighted average. Weighted composite scores were used to generate a bivariate correlation matrix in section 4.6.

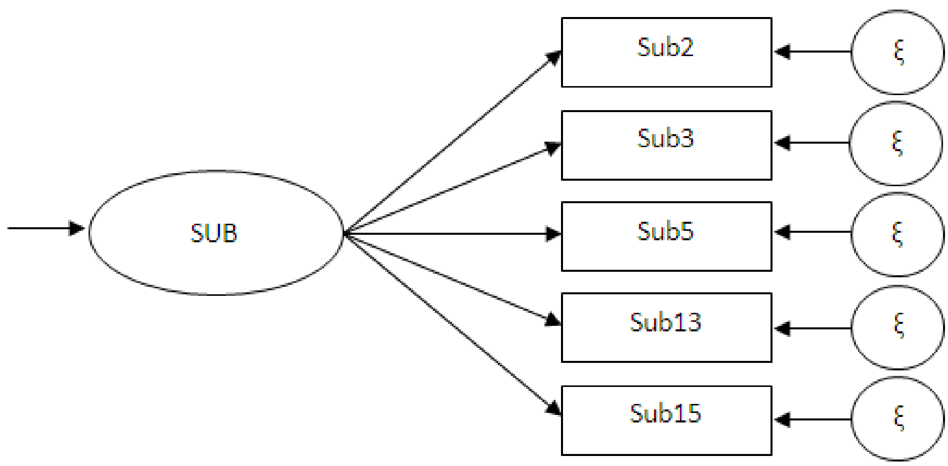

Figure 4.1. The congeneric model for the factor *Subject self-concept.*

Figure 4.1 presents a congeneric CFA model of the latent factor *Subject self-concept*. The factor includes five observed variables, for which the raw factor score coefficients are .156,

.249, .102, .360 and .219, respectively. Factor score coefficients represent the regression weights of individual items. Composite scores for *Subject self-concept* are calculated by first dividing each factor score by the sum of all five, which standardizes them to a scale of 1 (creating scores of .144, .229, .094, .331, and .202 respectively). Each standardized factor score coefficient is then multiplied by the corresponding score in an individual's response set, according to the formula: *Subject self-concept* = (Sub2 * .144) + (Sub3 * .229) + (Sub5 * .094) + (Sub13 * .331) + (Sub15 * .202) (Holmes-Smith, 2012; Holmes-Smith & Rowe, 1994). Factor loadings for composite scores are computed as the square root of the Rho reliability statistic for the corresponding measure, and error variances are computed as 1 – Rho, for that measure. Fixing these values with the M*plus* syntax shown in Appendix A expresses the latent variance of each factor in a simplified format, such that models of greater complexity can be fit with smaller samples (Holmes-Smith & Rowe, 1994; Raykov, 2009).

4.4 Congeneric model fit criteria

The purpose of the fit criteria used for congeneric analyses in the Pilot Study was to identify significant model misfit, in order to determine which scales and items were in greatest need of modification prior to the Main Study. To improve the certainty of identifying truly poor fit, cutoff thresholds were adopted (see Table 4.1) that lessened the likelihood that correct models would be rejected due to the idiosyncrasies of a small sample. This ran counter to logic of adopting stricter cutoff thresholds for ensuring good fit. The fit criteria in Table 4.1 were thus chosen to establish a more liberal standard for acceptable fit among simple congeneric models.

Table 4.1

Cutoff criteria for congeneric models in the Pilot Study

χ^2	>	.00
p of χ^2	$\geq$	.05
CFA Factor loading (λ)	$\geq$	.30
RMSEA	$\leq$	.10
Lower bound *CI* of *RMSEA*	$\leq$	.06
CFI	$\geq$	.95
Rho	$\geq$	.60

χ^2 (Chi-squared). Non-significant χ^2 statistics indicate the exact fit of a hypothesized covariance structure of to the covariance structure observed in an actual data set (Kline, 2011). χ^2 *p*-values are, however, more likely to be non-significant for smaller samples (Schumacker & Lomax, 2010). Significant *p*-values should, therefore, be more reliable indicators of poor fit to the small sample of data used for the present Pilot Study. Thus, to achieve the stated aim of identifying significantly poor fit, the threshold of $p > .05$ was adopted.

Root mean-square error of approximation. RMSEA values tend to be inflated in small samples, especially for models with few degrees of freedom (Chen, Curran, Bollen, Kirby & Paxton, 2008; Hu & Bentler, 1999, Kline, 2011). A threshold of RMSEA $\leq .10$, interpreted by MacCallum et al. (1996) as “mediocre” (Byrne, 2012), was adopted as the cutoff threshold for the present study. A threshold of $\leq .06$ was additionally adopted for its lower-bound 90% confidence interval, which Sivo, Fan, Witta, and Willse (2006) found to reject no correct models in a sample of $N = 150$.

Comparative fit index. CFI expresses the degree fit of a hypothesized covariance structure to the data on a scale of 0 to 1 (Bentler, 1990), and tends to vary directly with sample size (Sivo et al. 2006; Hu & Bentler, 1999). A threshold of CFI $\geq$ .95 was adopted for this study, based on Sivo et al.'s (2006) finding that CFI $\geq$ .95 rejected no correct models at N = 150.

Reliability. Factor reliability was calculated in the present study as the *Rho coefficient* ("Rho") (Raykov, 2009), a weighted reliability index, i.e. that accounts for the relative contribution of items in a measure. Judging by the common convention that reliability estimates below .60 are 'very low', .60-.69 are 'low', .70-.79 are 'adequate', .80-.89 are 'very good', and .90-1.00 are 'excellent' (Kline, 2011, p. 70), the threshold for reliability in the pilot study was set at .60.

Factor loadings. While many scholars recommend a factor loading threshold of .400 for retaining items in a factor model (Stevens, 2002; Field, 2009), a more relaxed threshold of .300 was adopted for the present study (9% of variance explained in the item by the latent factor).

4.5 Psychometric properties of the central constructs for the Pilot Study

In this section the psychometric properties of individual factors included in the hypothesized model were examined using both EFA and CFA of congeneric models. All measures that violated congeneric model fit criteria (highlighted values in Table 4.2) were analyzed independently in order to determine how the indicated misfit should be addressed prior to the Main Study.

The small size of the Pilot Study sample (N = 96) and low number of degrees of freedom typical of single-factor congeneric models was seen to necessitate relaxing the thresholds for certain fit statistics in the preceding section. The small size of the Pilot Study sample additionally increased the likelihood that the statistical features of the models presented below would be idiosyncratic. This potential risk was reduced by retaining all congeneric models that could be modified in theoretically defensible ways to satisfy the fit thresholds described above, for re-analysis at Time 1 of the Main Study.

Table 4.2

Summary of initial EFA and CFA fit statistics, Pilot Study (N = 96)

	EFA	CFA								
						RMSEA				
Scale (# items)	Advised dimensions	χ^2	*p*	*df*	Loading range	*Value*	Low 90%*CI*	High 90%*CI*	*CFI*	Rho
Subject self-concept (5)	1	9.5	.091	5	.704-.885	.097	.000	.190	.98	.89
Honesty-trust. self-concept (10)	2	EFA advised 2 dimensions								
Performance structure (5)	1	11.3	.046	5	.336-.638	.114	.014	.205	.87	.66
Mastery structure (5)	1	2.3	.807	5	.613-.854	.000	.000	.089	1.00	.85
Appropriate workload (5)	1	1.9	.859	5	.032-.947	.000	.000	.078	1.00	.56
Good teaching (8)	1	17.6	.618	20	.493-.846	.000	.000	.076	1.00	.91
Usefulness of curriculum (4)	1	1.1	.592	2	.667-.931	.000	.000	.167	1.00	.87
Clear goals and standards (5)	1	3.2	.665	5	.546-.776	.000	.000	.112	1.00	.78
Appropriate assessment (4)	1	6.8	.033	2	.200-.746	.159	.039	.296	.84	.58
Transparency of assessment (7)	1	24.4	.031	14	.374-.811	.092	.028	.148	.93	.82
Authenticity of assessment (7)	1	18.7	.178	14	.484-.782	.059	.000	.122	.96	.82
Peer norms (7)	1	34.2	.002	14	.568-.841	.123	.071	.175	.91	.88
Experience classroom rules (6)	1	15.2	.085	9	.494-.726	.085	.000	.156	.94	.76
Surface learning strategies (6)	2	EFA advised 2 dimensions								
Deep learning strategies (7)	1	26.2	.025	14	.406-.800	.095	.034	.151	.93	.83
Justifiability of cheating (4)	1	14.8	.001	2	.560-.964	.258	.146	.388	.90	.76
Self-reported cheating (4)	1	5.1	.078	2	.750-.914	.127	.000	.269	.94	.89

4.5.1 Honesty-Trustworthiness Self-Concept

The ten-item scale *Honesty-trustworthiness self-concept*, developed by Marsh, Parker, and Barnes (1985), rendered two dimensions using EFA. The original scale contained ten items, of which five were reversed. Four of the five reversed items formed a second dimension in the EFA. The author of this scale has, himself, recently been critical of reversed items (Marsh, Nagengast, & Morin, 2013). A single reversed item, Hon11, did, however, load with the non-reversed items in Dimension 1. As pointed out by Raykov (2012), a clear separation of loadings within a single scale suggests distinct factors, but with two items loading below .400, it was unclear what latent construct the selection of items in Dimension 2 would be sufficient to represent.

Table 4.3

Honesty-Trustworthiness Self-Concept EFA, Pilot Study (N = 96)

	EFA				
	2 dimensions advised				
Item	Dimension 1	R^2	Dimension 2	R^2	Item wording
Hon1	.604	.36	.025	.00	People can really count on me to do the right thing.
Hon4 (R)	-.131	.02	.756	.57	Cheating on a test is OK if I do not get caught.
Hon6	.683	.47	.036	.00	I always tell the truth.
Hon7 (R)	.283	.08	.336	.11	I sometimes take things that belong to other people.
Hon8	.668	.45	.043	.00	When I make a promise I keep it.
Hon9	.837	.70	-.232	.05	I am honest
Hon10	.638	.41	.159	.03	I often tell lies.
Hon11 (R)	.654	.43	.219	.05	Honesty is very important to me.
Hon12 (R)	.082	.01	.820	.67	I sometimes cheat.
Hon14 (R)	.267	.07	.367	.13	I sometimes tell lies to stay out of trouble.

Modification. Dimension 1, which included six of the original ten items including one reversed item, appeared to be the best approximation of the intended measure of honesty-trustworthiness self-concept. Removing items Hon4, Hon7, Hon12 and Hon14 reduced *Honesty-trustworthiness self-concept* to a 6-item scale. In CFA, this modified factor demonstrated good reliability (Rho = .84), with a non-significant chi-squared statistic ($\chi^2(9)$ =

11.1, p = .270), and otherwise good fit statistics (CFI = .99; $RMSEA$ = .049, CI_{lower} = .000), and strong factor loadings (.640 - .779).

Table 4.4

Honesty-trustworthiness Self-Concept CFA, Pilot Study (N = 96)

	CFA			
	Factor loadings (λ)			
Item	*Est.*	*S.E.*	*p*	Item wording
Hon1	.519	.102	.000	People can really count on me to do the right thing.
Hon6	.616	.080	.000	I always tell the truth.
Hon8	.594	.095	.000	When I make a promise I keep it.
Hon9	.679	.100	.000	I am honest
Hon10	.608	.122	.000	Honesty is very important to me.
Hon11	.677	.108	.000	I often tell lies. (R)

Theoretical considerations. The original *Honesty-trustworthiness self-concept* measure produced two dimensions that were characterized largely by whether item-wording expressed admissions of dishonesty (reversed items) or professions of honesty. All of the items in Dimension 2 in Table 4.3 are reversed. With the single exception of Hon11 (I tell lies), all of the items in Dimension 1 are positively worded professions of honesty. Items in Dimension 1 tend, moreover, to allude to more abstract notions of 'being honest', whereas items in Dimension 2 refer to specific acts such as 'taking things that belong to other people', 'telling lies to stay out of trouble', and 'cheating'. This conceptual disparity is consistent with the theoretical argument presented in the literature review that acts involving dishonesty may be contextualized in students' minds as either conventional or moral. Specific infractions may, even when they involve blatant deception, be viewed as part of the conventional domain, and may, therefore, be dissociated from moral conceptions of honesty and dishonesty (Turiel, 2002). This application of domain theory could help explain why the idea of 'telling lies' (item

Hon11: I often tell lies) is statistically distinct from 'lying to stay out of trouble' (item Hon14: I sometimes tell lies to stay out of trouble). The six items of Dimension 1 appear, like item Hon11, to pertain to conceptions of honesty as a moral abstraction, whereas the four items on Dimension 2 query students' judgments of specific rule-breaking behaviors.

Conclusion. Based on the foregoing statistical and theoretical reasoning, items composing Dimension 2 in Table 4.3, including Hon4, Hon7, Hon12 and Hon14, were removed from the *Honesty-trustworthiness self-concept* measure. Dimension 1, which appears to pertain to notions of honesty as a moral abstraction, was retained for the Main Study.

4.5.2 Performance Goal Structure

The five-item scale *Performance goal structure* rendered a single dimension in EFA, and demonstrated passable reliability (Rho = .66) and an acceptable range of factor loadings (.336 - .638). In CFA, however, the chi-squared statistic was significant ($\chi^2(5) = 11$, $p = .046$), and most other fit indicators were also unacceptable ($CFI = .87$; $RMSEA = .114$, $CI_{lower} = .014$).

Modification. Removing the weakest item (Perf36) minimally improved reliability (Rho = .67), but significantly improved the fit of the congeneric model ($\chi^2(2) = .263$, $p = .877$; $CFI = 1.00$; $RMSEA = .000$, $CI_{lower} = .000$).

Theoretical considerations: Item Perf36 describes a pattern of teacher behavior that could be read to suggest favoritism of 'smart' students (My science teacher calls on smart students more than on other students). Other items on this scale differ from item Perf36 in that they specify teacher behaviors that clearly express student comparisons, such as 'telling' students how they compare, 'pointing out' students who make good grades, and making 'obvious' those students who do not do well. Item Perf36 may have been a weak contributor to this factor because its language does not expressly frame 'calling on smarter students' as a performance comparison, as is done in the other four items.

Table 4.5

Performance Goal Structure CFA, Pilot Study (N = 96)

	CFA Factor loadings (λ)			
Item	*Est.*	*S.E.*	*p*	Item wording
Perf36	.336	.133	.012	My Science teacher calls on smart students more than on other students.
Perf61	.552	.129	.000	My Science teacher tells us how we compare to other students.
Perf69	.638	.112	.000	My Science teacher points out those students who get good grades as an example to us all.
Perf74	.477	.151	.002	My Science teacher makes it obvious when certain students are not doing well on their science work
Perf75	.626	.114	.000	My Science teacher lets us know which students get the highest scores on a test

Conclusion. The perception that a teacher 'calls on smart students' may not do as much to imply a performance goal structure as would the perceptions of explicit comparison-making described in the other items on this measure. Item Perf36 contributed little in the way of variance to the overall factor, damaged congeneric model fit, and appeared, moreover, to provide little unique information about a teacher's tendency to compare and/or favor high performing students. Item Perf36 was, therefore, excluded from the measure of *Performance goal structure* used in the Main Study.

4.5.3 Appropriate Workload

The five-item scale *Appropriate workload* rendered a single dimension in EFA, but demonstrated very low reliability (Rho = .56). In CFA, factor loadings fell, moreover, significantly outside the acceptable range (.032 - .947), with the estimate for item Apwkld49 failing to achieve significance (λ = .032, *p* = .756). Despite these issues, the congeneric model managed to achieve good fit ($\chi^2(5) = 1.9$, $p = .859$; $RMSEA = .000$, $CI_{lower} = .000$; $CFI = 1.00$). Low reliability and the poor performance of items Apwkld49 and Apwkld21, which explained very little variance in the factor were, however, in need of correction prior to the Main Study.

Modification. Removing the weakest item, Apwkld49, improved reliability (Rho=.64), while having little effect on other fit statistics ($\chi^2 = .407$, $p = .816$; $CFI = 1.00$; $RMSEA = .000$, $CI_{lower} = .000$). The range of factor loadings (.291 - .956), remained unacceptable, however, due to the loading of item Apwkld21 ($\lambda = .291$, $p = .016$), which fell below the minimum value for factor loadings of .300. Removing item Apwkld21 resulted in a small negative residual variance for item Apwkld52 (-.033), known as a 'Heywood case' (Byrne, 2012), that caused the residual covariance matrix to be non-positive definite. The Heywood case was addressed by fixing the residual variance of item Apwkld52 to .00001 (Muthen & Muthen, 2014). This resulted in model of fit similar the models tested above ($\chi^2(2) = .396$, $p = .820$; $CFI = 1.00$; $RMSEA = .000$, $CI_{lower} = .000$; Rho = .67), but with a very imbalanced 3-item factor structure, composed of a factor loading of $\lambda = 1.00$ (Apwkld52), accompanied by factor loadings of less than half that magnitude for Apwkld30 ($\lambda = .405$) and Apwkld35 ($\lambda = .460$).

Theoretical considerations: Item Apwkld49 queries respondents' evaluations of the scope of coverage in a given class (It seems to me that my Science teacher tries to cover too much material). Coverage of material could imply breadth of curriculum, as distinct from workload. The fact that this item failed completely to load on the factor may indicate, further, that the evaluations of *scope* of coverage are distinct from evaluations of the *amount* of work, which is the construct's intended conception (Entwistle & Ramsden, 1983). While the breadth or scope of material covered in a given class seems likely to correspond to how much work is assigned, the judgments of the appropriateness of that amount of work are inherently subjective. The possibility that the breadth of coverage in a given class is unrelated to students' perceptions of how much work is appropriate is, in fact, consistent with the treatment of this factor as a measure of students' felt moral obligation to work hard in the hypothesized PTLC model.

Table 4.6

Appropriate Workload CFA, Pilot Study (N = 96)

	CFA			
	Factor loadings (λ)			
Item	*Est.*	*S.E.*	*p*	Item wording
Apwkld21	.295	.125	.018	There's a lot of pressure on you as a student in my Science class.
Apwkld30	.417	.114	.000	The large amount of work you have to do in my Science class means you can't understand it all completely.
Apwkld35(R)	.498	.119	.000	In my Science class, we are usually given enough time to understand the things we have to learn.
Apwkld49	.032	.102	.756	It seems to me that my Science teacher tries to cover too much material.
Apwkld52	.947	.168	.000	The amount of work in my Science class is too large.

Item Apwkld21 queries respondents' experience of 'pressure on students' in a given class. While in tertiary contexts, for which this measure was developed, the main source of 'pressure ' in a class may be workload, in secondary contexts the pressures exerted on students often emanate from a wider array of sources, such as peer and disciplinary regimes. As suggested by its consistently low loadings, item Apwkld21 may have only partly pertained to respondents' notions of workload in the Pilot Study.

Conclusion. The measure *Appropriate workload* demonstrated a number of psychometric problems. After removing two of the five original items, factor structure remained poorly balanced between the highest loading item (Apwkld52) and the other two, resulting in a Heywood case. Reliability estimates were also uniformly low (.56 - .67). These observations suggested, overall, that this measure of appropriate workload may not be appropriate for the

American international secondary school population. The measure was tentatively retained for further analysis at Time 1 of the Main Study.

4.5.4 Appropriate Assessment

The four-item scale *Appropriate assessment* achieved a single dimension in EFA, but demonstrated very low reliability (Rho = .58), and generally unacceptable congeneric fit. Chi-squared was significant ($\chi^2(2) = 6.8$, $p = .033$); $CFI = .84$ and $RMSEA = .160$ both fell wide of their respective cutoff thresholds; and the factor loading of Appas76 fell below the threshold of .300 ($\lambda = .200$, $p = .259$). In sum, the four-item congeneric model for *Appropriate assessment* demonstrated very poor fit.

Modification. The weakest item, Appas76 (It would be possible to succeed in my Science class just by studying for tests and quizzes the night before), seems to assume that tests and quizzes are the only bases for success in a given class. This is, however, seldom true in secondary Science classes, which frequently involve writing assignments, presentations, and laboratory work. The item could be read, moreover, to query whether a class is easy enough to pass with minimal effort, rather than whether a teacher's assessments are appropriate. Removing this item improved Rho reliability to .64. Fit statistics were obtained in CFA by constraining the two most similar residual variances, items Appas16 (My Science teacher asks us too many questions just about facts) and Appas17 (To do well in my Science class, all you really need is a good memory) to be equal, which rendered good fit ($\chi^2(1) = .257$, $p = .612$; $RMSEA = .000$, $CI_{lower} = .000$; $CFI = 1.00$). The 3-item factor structure was, however, markedly imbalanced, with a factor loading of $\lambda = .899$ (Appas47), accompanied by loadings of $\lambda = .496$ (Appas16) and $\lambda = .360$ (Appas17), and low reliability (Rho = .64).

Table 4.7

Appropriate Assessment CFA, Pilot Study (N = 96)

	CFA			
	Factor loadings (λ)			
Item	*Est.*	*S.E.*	*p*	Item wording
Appas16	.586	.149	.000	My Science teacher asks us too many questions just about facts.
Appas17	.410	.133	.002	To do well in my Science class, all you really need is a good memory
Appas47	.746	.170	.000	My Science teacher seems to care more about what you've memorized that what you've understood.
Appas76	.200	.177	.259	It would be possible to succeed in my Science class just by studying for tests and quizzes the night before.

Conclusion. The measure *Appropriate assessment* demonstrated very poor initial fit to pilot data. Dropping the weakest-loading item produced excellent fit indices, but with low reliability and noted weakness in the factor loading range. While these observations suggest that the measure *Appropriate assessment* may not be suitable for the student population presently under consideration, it was tentatively retained for further analysis at Time 1 of the Main Study.

4.5.5 Transparency of Assessment

The seven-item scale *Transparency of assessment* achieved a single dimension in EFA, and demonstrated good reliability (Rho = .82). While in CFA, the factor's *RMSEA* = .092 (CI_{lower} = .028), and factor loadings (.372 - .811) were acceptable, its χ^2 statistic was significant ($\chi^2(14)$ = 24.4, p = .031), and its *CFI* = .93 fell below the threshold of .95.

Modification. Removing the weakest item, Trans41 (λ = .374, p = .000) had no effect on reliability (Rho = .82), but improved χ^2 to non-significance ($\chi^2(9)$ = 16.3, p = .0612); and while the point-estimate for *RMSEA* was unaffected by this modification (*RMSEA* = .092), the lower bound of its confidence interval fell to .000. *CFI* also improved from .93 to .95, and the range of factor loadings remained above the established limit (.447 - .801).

Table 4.8

Transparency of Assessment CFA, Pilot Study (N = 96)

	CFA			
	Factor loadings (λ)			
Item	*Est.*	*S.E.*	*p*	Item wording
Trans22	.441	.111	.000	I know what is needed to successfully accomplish graded assignments in my Science class.
Trans28	.594	.057	.000	I understand the purpose of graded assignments in my Science class.
Trans32	.811	.053	.000	I understand what is needed in all graded assignments in my Science class.
Trans41	.374	.112	.000	I am told in advance WHY I am being asked to do graded assignments in my Science class.
Trans45	.670	.082	.000	I am told WHAT science topics and information I will be graded on in my Science class.
Trans63	.679	.083	.000	I am told in advance WHEN I will be graded in my Science class.
Trans66	.772	.055	.000	I know in advance HOW I will be graded in my Science class.

Theoretical considerations. Item Trans41 is one of two items on this scale that pertains to whether students understand the purposes of graded assignments in science class (I am told in advance WHY I am being asked to do graded assignments in my Science class). Item Trans28 asks students to respond to a very similar statement (I understand the purpose of graded assignments in my Science class). The key difference between these two items appears to be whether students are 'told in advance', or whether they more broadly 'understand' the purpose of graded assignments. Item Trans28 (λ = .594, p = .000) explained more than twice as much variance in the factor (R^2 = .353) as Trans41 (R^2 = .140), which may have reflected this broader language.

Conclusion. Determining whether students are 'told in advance' appears to arbitrarily privilege a single aspect of how students may come to understand why graded assignments are assigned. Item Trans41 contributed, as such, little in the way of variance or unique information about whether students understand the purpose of graded assignments in science class and was, therefore, eliminated from the *Transparency of assessment* measure used in the Main Study.

4.5.6 Peer Norms Related to Cheating

The seven-item measure *Peer norms* achieved a single dimension in EFA, and demonstrated good reliability (Rho = .88). In CFA, however, the fit of the congeneric model was unacceptable. The χ^2 statistic was highly significant ($\chi^2(14) = 33.1$, $p = .003$); and *RMSEA* = .123, and its lower confidence interval of .071, fell considerably wide of thresholds for appropriate fit in the Pilot Study. *CFI* = .91 also fell below the threshold of .95.

Modification. Since the two weakest items in this model, Peer51 and Peer65, had identical factor loadings ($\lambda = .568$, $p = .000$), the removal of each was explored, separately. Removal of Peer51, which had the greatest positive effect on model fit, did not affect reliability (Rho = .88), but did improve *CFI* to .95; yet while both *RMSEA* and its lower confidence interval improved, the former still exceeded the threshold (*RMSEA* = .107, $CI_{Lower} = .036$), and χ^2 remained significant ($\chi^2(9) = 18.9$, $p = .026$).

When peer65 was removed, while still retaining peer51, there was, again, no effect on reliability (Rho = .88). Other fit statistics, however, either deteriorated or demonstrated marginal improvement. χ^2 was still significant ($\chi^2(9) = 22.9$, $p = .007$); *RMSEA* increased (*RMSEA* = .127, $CI_{Lower} = .063$); and *CFI* improved marginally, from .92 to .93.

Table 4.9

Peer Norms CFA, Pilot Study (N = 96)

	CFA			
	Factor loadings (λ)			
Item	*Est.*	*S.E.*	*p*	Item wording
Peer24(R)	.814	.041	.000	Most of my classmates think that I should NOT cheat in Science class.
Peer31(R)	.699	.071	.000	My classmates will look down on me if I cheat in Science class this year.
Peer40(R)	.706	.075	.000	None of my classmates think it is okay to cheat in my Science class.
Peer51	.568	.077	.000	Most of my classmates expect me to cheat in my Science class this year.
Peer55	.760	.055	.000	Most of my classmates in Science class this year would be willing to cheat on a Science test to avoid failing.
Peer58	.841	.050	.000	If I cheated on a Science test this year, most of my classmates would think that's okay.
Peer65(R)	.568	.108	.000	Most of my classmates would NOT think it's okay if I cheated in Science class this year.

The possibility of removing both Peer51 and Peer65 was explored next. Removal of both items had no effect on reliability (.88) but improved *CFI* to .96. The change in *RMSEA* was, however, still only a marginal improvement over the original 7-item factor model (*RMSEA* = .114, CI_{Lower} = .011); and χ^2, while improved, was still significant ($\chi^2(5) = 11.2$, p = .048).

The factor was next analyzed without each of the remaining non-reversed items, Peer55 and Peer58, in turn. As in the case of the four items removed from the *Honesty-trustworthiness self-concept* measure, method effects are often associated with reversed items (Marsh et al., 2013). While CFA indicated that removing either Peer55 or Peer58, in addition to Peer51, resulted in acceptable fit, removing the combination of Peer51 and Peer55 generated the best fit, overall. While the reliability of this 5-item factor (.86) was slightly lower than the original 7-item factor (.88), χ^2 improved dramatically ($\chi^2(5) = 4.41$, $p = .492$); and other fit statistics fell comfortably within established limits ($RMSEA = .000$, $CI_{Lower} = .011$; $CFI = 1.00$).

Theoretical considerations. Items Peer51 (Most of my classmates expect me to cheat in my Science class this year) and Peer55 (Most of my classmates in Science class this year would be willing to cheat on a Science test to avoid failing) both refer to anticipated behaviors, either on the part of the respondent or on the part of his or her peers. Every other item in the measure refers, by contrast, to how the respondent believes her or his peers would judge cheating. While all of these items share a common referent, i.e. cheating, they appear to represent two sides of the well-known discrepancy between behaviors and beliefs, or the 'belief-behavior incongruity' (BBI) (Bergman, 2002; Stephens & Nicholson, 2008). The significant χ^2 distribution for the original 7-item model, indicated by both the highly significant χ^2 statistic and the inflated *RMSEA* statistic, appeared to reflect this incongruous combination of items that pertain to anticipated cheating behavior *versus* those that pertain to judgments of whether cheating is right or wrong.

Conclusion. Items Peer51 and Peer55 appear to represent the behavioral component of the BBI, whereas the other items in the measure appear to pertain to beliefs. While both items had large and significant loadings on the overall factor, removing them greatly improved its χ^2 distribution. Both items were, therefore, removed from the *Peer Norms* measure used in the Main Study.

4.5.7 Experience of Classroom Rules

The six-item scale *Experience of classroom rules* rendered a single dimension in EFA, and was acceptably reliable (Rho = .76). While in CFA, the congeneric model demonstrated satisfactory fit according to most measures ($\chi^2(9) = 15.2$, $p = .085$; $RMSEA = .085$, $CI_{lower} = .000$), and factor loadings that fell within an acceptable range (.494 - .726), an unacceptable degree of misfit was indicated by the *CFI* statistic (*CFI* = .94), which fell below the threshold of .95.

Modification. Removing the item with the lowest loading, Exrule59 (see Table 4.10), increased *CFI* to .95 with no effect on reliability (.76) and minimal effect on other fit statistics ($\chi^2(5) = 9.5$, $p = .092$; $RMSEA = .096$, $CI_{lower} = .000$). The range of factor loadings for the modified factor also remained above the threshold level (.557 - .686).

Table 4.10

Experience of Classroom Rules CFA, Pilot Study (N = 96)

	CFA			
	Factor loadings (λ)			
Item	*Est.*	*S.E.*	*p*	Item wording
Exrule20	.534	.092	.000	The rules in my Science class are fair.
Exrule23	.592	.101	.000	My Science teacher makes sure that everyone follows the rules in class.
Exrule27	.726	.088	.000	If a student breaks the rules in my Science class, the teacher will do something about it.
Exrule43	.632	.105	.000	If a rule is broken in my Science class, students know what the teacher will do about it.
Exrule54	.611	.091	.000	Everyone knows the rules for how students should behave in my Science class.
Exrule59	.494	.105	.000	The punishment for breaking rules in my Science class is the same no matter who you are.

Theoretical considerations. While Exrule59 shares in common with most other items on this measure the themes of fairness and consistency, it is the only item to explicitly mention punishment (The punishment for breaking rules in my Science class is the same no matter who you are). Item Exrule59 is, therefore, only able to characterize teachers who routinely punish students for breaking rules. The lack of clarity around how item Exrule59 pertains to teachers who embrace positive and supportive forms of behaviour management (e.g. Sugai & Horner, 2002) may have been a source of misfit in the congeneric model for this measure.

Conclusion. Querying students' experience of punishment does not necessarily conform to their experience of rules in a given classroom. For students whose teachers strive to manage behavior in non-punitive ways, item Exrule59 may have been a source of model misfit. Item Exrule59 was, therefore, eliminated from the measure of *Experience of classroom rules* used in the Main Study.

4.5.8 Surface Learning Strategies

The measure *Surface learning strategies*, developed by Simon, Dewitte, and Lens (2004), rendered two dimensions in EFA. The original measure included six items describing three surface strategies: memorizing what is not understood, strategically skipping information, and rehearsing information so as to be able to reproduce it. The two items that formed Dimension 2, Surf82 and Surf90, described the third of these strategies, rehearsing (I study for Science class by rehearsing (repeating over and over) important information; and, I study for Science class by rehearsing and repeating the material over and over until I can write it exactly, word-for-word).

Modification. While Dimension 1 appeared to be the better of the two measures of *Surface learning strategies*, it was not certain that removing both items Surf82 and Surf90 was necessary or helpful. For this reason, the advised number of dimensions was ascertained by EFA when each of the items in question was removed, respectively. When Surf82 was removed, 2 Dimensions were still advised. When Surf90 was removed, and Surf82 retained, however, a single dimension was advised. A CFA of this latter modification is discussed below.

Table 4.11

Surface Learning Strategies EFA, Pilot Study (N = 96)

	EFA				
	2 dimensions advised				
Item	Dimension 1	R^2	Dimension 2	R^2	Item wording
Surf82	.173	.03	.541	.29	I study for Science class by rehearsing and repeating the material over and over again until I can write it exactly, word-for-word.
Surf87	.559	.31	.168	.03	I study for Science class by skipping parts I do not understand.
Surf88	.362	.13	.283	.08	I study for Science class by memorizing things I do not understand.
Surf90	-.153	.02	.629	.40	I study for Science class by rehearsing (repeating over and over) important information.
Surf91	.777	.60	-.012	.00	I study for Science class by skipping over parts I think the teacher will not ask questions about.
Surf97	.620	.38	-.151	.02	I study for Science by skipping parts I do not find important.

The removal of item Surf90 rendered a factor of passable reliability (Rho=.61), and acceptable fit with respect to other statistics ($\chi^2(5) = 6.74$, $p = .241$; $RMSEA = .060$, $CI_{lower} < .000$; $CFI = .95$). Factor loadings were, however, outside the acceptable range (.169-.801), with the estimate for item Surf82 failing to achieve significance ($\lambda = .169$, $p = .222$).

Dimension 1 of the original EFA presented in Table 4.11, which excludes both Surf82 and Surf90, rendered a factor of improved reliability (Rho = .65), acceptable fit ($\chi^2(2) = 3.43$, $p = .180$; $RMSEA = .083$, $CI_{lower} = .000$; $CFI = .96$), and factor loadings within the acceptable range (.325 - .798). While both of the items that compose Dimension 2 were strongly correlated with each other, Dimension 2 was not appropriate for structural equation modelling for two reasons: (1) it was not clear that these two items would suffice in terms of content validity for measuring *rehearsing* as a learning strategy; and (2) congeneric models with fewer than three items have negative degrees of freedom and cannot, therefore, be modelled straightforwardly.

Table 4.12

Surface Learning Strategies CFA, Pilot Study (N = 96)

	CFA			
	Factor loadings (λ)			
Item	*Est.*	*S.E.*	*p*	Item wording
Surf87	.520	.119	.000	I study for Science class by skipping parts I do not understand.
Surf88	.325	.130	.013	I study for Science class by memorizing things I do not understand.
Surf91	.798	.141	.000	I study for Science class by skipping over parts I think the teacher will not ask questions about.
Surf97	.630	.084	.000	I study for Science by skipping parts I do not find important.

Theoretical considerations. The four items composing Dimension 1 measure two common surface learning strategies: memorizing what is not understood, and skipping portions of assigned materials that are deemed unlikely to appear on tests, in order to minimize workload. Another prominent surface strategy instrument, the Study Process Questionnaire (SPQ) (Biggs et al., 2001), also consists, for the sake of comparison, exclusively of items describing memorization and minimizing the scope of study.

Conclusion. Based on the foregoing statistical and theoretical reasoning, the *Surface learning strategies* measure used in the Main Study was reduced to Dimension 1 in Table 4.12, which includes items Surf87, Surf88, Surf91, and Surf97.

4.5.9 Deep Learning Strategies

The seven-item scale *Deep learning strategies* achieved a single dimension in EFA, and demonstrated good reliability (Rho = .83). In CFA, while $RMSEA$ = .095, CI_{lower} = .034, and the range of factor loadings (.406 - .800) were acceptable, two statistics indicated misfit. The χ^2 statistic was significant ($\chi^2(14) = 26.2$, $p = .025$); and $CFI = .93$ fell beneath the established threshold of .95.

Table 4.13

Deep Learning Strategies CFA, Pilot Study (N = 96)

	CFA			
	Factor loadings (λ)			
Item	*Est.*	*S.E.*	*p*	Item wording
Deep80	.406	.102	.000	I spend some time thinking about how to do my Science work before I start it.
Deep81	.673	.073	.000	I ask myself questions when I work on Science to make sure I understand.
Deep83	.790	.053	.000	I try to connect new work in Science to what I've learned before.
Deep85	.607	.102	.000	When working on a Science problem, I try to see how it connects with something in everyday life.
Deep93	.662	.073	.000	I take my time to figure out my work in Science.
Deep94	.800	.045	.000	When I make mistakes in Science, I try to figure out why.
Deep96	.496	.083	.000	If I can't solve a Science problem one way, I try to use a different way.

Modification. Removing the item that contributed the least variance to the overall factor (Deep80) (see Table 4.13) had no effect on reliability (.83), but did improve fit ($\chi^2(9) = 10.9$, $p = .285$); $CFI = .99$; $RMSEA = .046$, $CI_{lower} = .000$). The range of factor loadings for the 6-item factor was, moreover, well above the threshold (.508 - .837).

Theoretical considerations: Item Deep80 pertains to whether students plan how to do work in Science class (I spend some time thinking about how to do my Science work before I start it). Other items on this scale are, by contrast, more specific as to how a respondent 'does' work in terms of self-questioning, active connection-making, self-correction, and applications of time and effort to understanding. The strategies that students plan before doing school work could, in fact, be of a deep or surface character. The activity of 'thinking about how to do work' may be too broad to necessarily connote deep learning strategies.

Conclusion: The fact that students think about their work before they start it does not necessarily mean that they will engage in deep learning strategies. Item Deep80 appeared, therefore, to be a weak contributor to the overall factor in that students may spend time planning to do work in a variety of ways that are not 'deep'. Item Deep80 was, as such, removed from the *Deep learning strategies* measure used in the Main Study.

4.5.10 Justifiability of Cheating

The four-item scale *Justifiability of cheating* achieved a single dimension in EFA, and demonstrated acceptable reliability (.76). In CFA, while factor loadings appeared robust (.560 - .964), the χ^2 statistic for the congeneric model was highly significant ($\chi^2(2) = 14.5$, $p < .000$); *RMSEA* and the lower bound of its confidence interval were also more than twice the threshold (*RMSEA* = .260, CI_{lower} = .146); and *CFI* = .90 was also below the minimum of .95.

Modification. The four-item measure for *Justifiability of cheating* presented in Table 4.14 included items from the original measure (Chjust86 and Chjust99) that have been used in previous secondary-level studies conducted by Murdock and various colleagues (2004, 2008); a third item (Chjust79) that was adapted from a scale used by Anderman et al. (1998), upon which Murdock's aforementioned research built; and a fourth item that was developed strictly for the purpose of this study (Chjust98).

Removing the item that was developed *a priori* for use in this project (Chjust98) retained the three items that have been used successfully in prior, secondary-level research. This three-item measure demonstrated good reliability (.80), and a robust range of factor loadings (.554 - .930). Fit statistics for the three-item measure were obtained by constraining the residual variances of items Chujst79 and Chjust86, which were closest in magnitude, to be equal. This produced a single degree of freedom, and rendered a congeneric model of excellent fit (χ^2 = .690, p = .406; *RMSEA* = .000, CI_{lower} = .000; *CFI* = 1.00).

Table 4.14

Justifiability of Cheating CFA, Pilot Study (N = 96)

	CFA			
	Factor loadings (λ)			
Item	*Est.*	*S.E.*	*p*	Item wording
Chjust79	.751	.041	.000	It's reasonable to cheat in my Science class.
Chjust86	.964	.049	.000	Students would have a good reason to cheat on a test in my Science class.
Chjust98	.743	.055	.000	I can understand why students would cheat in my Science class.
Chjust99	.560	.099	.000	Students would be justified to cheat on an exam in my Science class.

Theoretical considerations: Item Chjust98 queries the 'understandability' of cheating in the context of Science class, whereas the other three items on the *Justifiability of cheating* measure refer specifically to whether, in the respondent's opinion, students have valid 'reasons' to cheat in that context. In contrast, item Chjust98 was developed based on the assumption that cheating is 'understandable' to the extent that there are *reasons* to cheat. This assumption appears to have been flawed. Understandability may, in fact, draw upon a different realm of consideration than reasonability. Cheating might, for instance, be understandable in view of an individual's established habits or susceptibility to temptation. It might be possible to understand why someone cheats, in other words, without agreeing that they have valid reasons for cheating. It is possible that the discrepant chi-squared distribution for this model, evidenced by both the highly significant chi-squared statistic and the inflated *RMSEA* statistic, reflected the conceptual disparity, created by the inclusion of item Cheat98, between understandability and reasonability.

Conclusion: Inasmuch as one can understand why someone cheats, without believing that their cheating is justifiable for valid reasons, the inclusion of an item that queries understandability appears to be poorly suited for this measure. Item Chjust98 was, therefore, removed from the *Justifiability of cheating* measure used in the Main Study.

4.5.11 Self-Reported Cheating

The four-item measure for *Self-reported cheating* achieved a single dimension in EFA, and demonstrated good reliability (Rho = .89). In CFA, factor loadings were robust (.750 - .914), and the χ^2 statistic was non-significant ($\chi^2(2) = 5$, $p = .083$). However, incremental fit indices of *CFI* = .94 and *RMSEA* = .127 fell short of established limits.

Table 4.15

Self-Reported Cheating CFA, Pilot Study (N = 96)

	CFA			
	Factor loadings (λ)			
Item	***Est.***	***S.E.***	***p***	**Item wording**
Cheat84	.809	.074	.000	I sometimes cheat on Science tests this year.
Cheat89	.802	.084	.000	I have cheated in Science class this year.
Cheat92	.914	.063	.000	I sometimes cheat on my Science work this year.
Cheat95	.750	.095	.000	I have cheated on Science class work by copying answers from other students this year.

Modification. The measure *Self-reported cheating* was originally composed of four items, of which three (Cheat84, Cheat92, and Cheat95) constituted a measure developed by Midgley et al. (2000) that has been used in prior studies (Murdock et al., 2001; Brown-Wright et al., 2013). A fourth item (Cheat89), adapted from a measure used in secondary-level studies by Anderman and various colleagues (1998, 2004, 2010), was added as a safeguard against possible measure dysfunction (see Appendix C).

Removing item Cheat89 reverted to the original three-item measure developed by Midgley et al. (2000), which had good reliability (.87), and robust factor loadings (.750 - .914). Fit statistics for the three-item measure were obtained by constraining the residual variances of items Cheat84 and Cheat95, which were closest in magnitude, to be equal. This produced a single degree of freedom, and rendered a congeneric model of excellent fit ($\chi^2(1) = .016$, $p = .900$; *RMSEA* = .000, $CI_{lower} = .000$; *CFI* = 1.00).

Conclusion. Item Cheat89, which was added for the purposes of the present study, served only to detract from the measure *Self-reported cheating* originally developed by Midgley et al. (2000), and was, therefore, excluded from the Main Study.

4.5.12 Summary of fit statistics for modified measures

The purpose of the congeneric model analyses presented in this section was to identify prominent model misfit and modify measures accordingly, so as to avoid, inasmuch as possible, the need for *post-hoc* modifications during the Main Study. Based on the foregoing analyses, seventeen of the 99 items on the original questionnaire instrument were eliminated for the Main Study.

Table 4.16

Modified measures: Summary of EFA and CFA fit statistics, Pilot Study (N = 96)

	EFA	CFA								
						RMSEA				
Scale (# items)	Advised dimensions	χ^2	*p*	*df*	Loading range	Value	Low 90%*CI*	High 90%*CI*	*CFI*	Rho
Subject self-concept (5)	1	9.3	.099	5	.704-.885	.097	.000	.190	.98	.89
Honesty-trust. self-concept* (6)	1	11.1	.270	9	.640-.779	.049	.000	.131	.99	.84
Performance structure* (4)	1	.263	.877	2	.423-.674	.000	.000	.100	1.00	.67
Mastery structure (5)	1	2.2	.815	5	.613-.854	.000	.000	.089	1.00	.85
Appropriate workload* (3)	1	.396	.820	1	.405-1.00	.000	.000	.121	1.00	.67
Good teaching (8)	1	16.9	.661	20	.493-.846	.000	.000	.076	1.00	.91
Usefulness of curriculum (4)	1	1.0	.600	2	.667-.931	.000	.000	.167	1.00	.87
Clear goals and standards (5)	1	3.1	.678	5	.546-.776	.000	.000	.112	1.00	.78
Appropriate assessment* (3)	1	.257	.612	1	.360-.899	.000	.000	.215	1.00	.64
Transparency of assessment* (6)	1	16.3	.061	9	.460-.831	.092	.000	.162	.95	.82
Authenticity of assessment (7)	1	18.1	.204	14	.484-.782	.060	.000	.122	.97	.82
Peer norms* (5)	1	4.4	.492	5	.534-.713	.000	.000	.133	1.00	.86
Experience classroom rules* (5)	1	9.5	.092	5	.557-.686	.096	.000	.156	.95	.76
Surface learning strategies* (4)	1	3.4	.180	2	.325-.798	.086	.000	.237	.96	.66
Deep learning strategies* (6)	1	10.9	.285	14	.508-.806	.046	.000	.129	.99	.83
Justifiability of cheating* (3)	1	.690	.460	1	.664-881	.000	.000	.252	1.00	.80
Self-reported cheating* (3)	1	.480	.785	2	.750-.914	.000	.000	.131	1.00	.87

* Modified scales

4.6 Correlational analysis

The correlation matrix presented in Table 4.17 was estimated next in order to identify multicollinearity between variables. Very high correlations between variables, or multicollinearity, can lead to spurious results in SEM by falsely inflating beta paths (Field, 2009; Schumacker & Lomax, 2010). Instances of high multicollinearity (r > .750) in Table 4.17 were thus addressed by either dropping measures or including them in a higher-order factor structures.

A complete correlation matrix of seventeen latent variables cannot be estimated as a structural model with only 96 cases, because the number of free model parameters would exceed the number of cases. The correlation matrix in Table 4.17 was, therefore, estimated in SPSS using composite scores, which meant that latent factor loadings and error variances were not used. Field (2009) suggests that correlations of $r \geq .800$ should be considered multicollinear. For the purposes of this analysis, however, in which the omission of factor loadings and error variances is likely to reduce the magnitude of correlations between factors, the threshold for what was considered multicollinear correlations was lowered to $r \geq .750$.

The matrix in Table 4.17 presents all bivariate correlations among seventeen composite latent factors. Seven of these factors appeared to form a cluster with correlations of .750 or greater (highlighted in Table 4.17), including *Clear goals and standards* (Goals), *Mastery classroom structure* (Mast), *Authenticity of assessment* (Auth), *Transparency of assessment* (Trans), *Experience of classroom rules* (Exrule), *Appropriate assessment* (Appas), and *Good teaching* (Gteach). The factors in this cluster will either be dropped from the model due to the redundancy implied by extreme multicollinearity, or be fitted into a higher-order factor structure.

Table 4.17

Bivariate Correlations among Variables in Model 1, Pilot Study (N = 96)

	SUB	HON	PERF	MAST	GOALS	GTEACH	APWKLD	CURUSE	APPAS	AUTH	TRANS	EXRULE	PEER	DEEP	SURF	CHJUST
SUB	1															
HON	.256*	1														
PERF	-.037	.059	1													
MAST	.400**	.331**	.119	1												
GOALS	.416**	.335**	.010	.786**	1											
GTEACH	.344**	.312**	.122	.877**	.851**	1										
APWKLD	.499**	.172	-.058	.513**	.485**	.432**	1									
CURUSE	.425**	.265**	.102	.670**	.589**	.606**	.456**	1								
APPAS	.176	.083	-.334*	.811**	.813**	.792**	.486**	.455**	1							
AUTH	.405**	.314**	.030	.726**	.742**	.765**	.370**	.679**	.707**	1						
TRANS	.383**	.323**	.061	.719**	.792**	.785**	.494**	.658**	.755**	.785**	1					
EXRULE	.316**	.425**	.007	.703**	.765**	.777**	.344**	.572**	-.273**	.767**	.752**	1				
PEER	-.118	-.142	.016	-.280**	-.361**	-.353**	-.368**	-.340**	-.306*	-.398**	-.393**	-.400**	1			
DEEP	.464**	.293**	.098	.624**	.551**	.612**	.480**	.602**	.711**	.594**	.611**	.575**	-.270**	1		
SURF	-.470**	-.192	.111	-.259*	-.298**	-.271**	-.331**	-.202*	-.446*	-.318**	-.315**	-.271**	.038	-.413**	1	
CHJUST	-.369**	-.365**	.014	-.519**	-.590**	-.569**	-.497**	-.475**	-.466**	-.559**	-.630**	-.626**	.633**	-.591**	.269**	1
CHEAT	-.213*	-.262**	.005	-.363**	-.449**	-.435**	-.324**	-.376**	-.355*	-.462**	-.524**	-.486**	.517**	-.440**	.302**	.709**

*. Correlation is significant at the 0.05 level (2-tailed); **. Correlation is significant at the 0.01 level (2-tailed).

SUB= Subject Self-Concept, HON= Honesty-Trustworthiness Self-Concept, PERF= Performance Goal Structure, MAST= Mastery Goal Structure, GOALS= Clear Goals and Standards, GTEACH= Good Teaching, APWKLD= Appropriate Workload, CURUSE= Usefulness of Curriculum, APPAS = Appropriate assessment; AUTH= Authenticity of Assessment, TRANS= Transparency of Assessment, EXRULE= Experience of Classroom Rules, PEER= Peer Norms Related to Cheating, DEEP= Deep Learning Strategies, SURF= Surface Learning Strategies, CHJUST= Justifiability of Cheating, CHEAT= Self-Reported Cheating.

4.6.1 Higher-order factor analysis

A higher-order latent factor represents the hypothesis that two or more first-order latent factors, or those measured with observed variables, can be explained in terms of a single, shared source of variance (see Figure 4.2). Higher-order factors are often hypothesized for multicollinear groups of factors, where the large amounts of statistical overlap among them are believed to reflect the higher-order variance source. Higher-order factors are, as such, measured indirectly, by way of the first-order variables they explain (Kline, 2011).

A second-order factor model for all seven multicollinear variables in Table 4.17, estimated with weighted composite scores, demonstrated unacceptably high *RMSEA* ($\chi^2(14) = 32$, $p = .004$; $RMSEA = .116$, $CI_{lower} = .063$; $CFI = .97$; $SRMR = .034$). *Good teaching* was the defining variable in this structure, with an unstandardized loading of $\lambda = 1.000$. *Mastery goal structure* and *Clear goals and standards*, which had the second and third highest loadings ($\lambda = .944$ and $\lambda = .856$), contributed little unique variance, however, due to high multicollinearity with *Good teaching* ($r = .877$ and $r = .851$, respectively). Removing these latter two factors improved parsimony, but did not improve *RMSEA* (.132). Of the five remaining factors, *Appropriate assessment* had the lowest reliability (.64) and weakest factor loadings in the preceding congeneric analyses. Dropping *Appropriate assessment* rendered an excellent second-order model ($\chi^2(2) = .892$, $p = .640$; $RMSEA = .000$, $CI_{lower} = .000$; $CFI = 1.00$; $SRMR = .006$) (see Figure 4.2). The four factors retained in this model captured three distinct dimensions of 'teacher quality': pedagogical skill, assessment quality, and behavior management.

Small sample correction. The second-order factor structure reported above for *Teacher quality* was next cross-validated by estimating it with all observed variables (instead of composite scores). This entailed 82 free model parameters, which biased fit statistics due to the small sample size ($\chi^2(295) = 468$; $RMSEA = .078$, $CI_{lower} = .065$; $CFI = .85$; $SRMR = .065$)

(Herzog et al., 2007). The *Swain-R, version 1.2* small sample adjustment software package (Boomsma & Herzog, 2013) was used to correct this bias, and demonstrate that the model for *Teacher quality* fit the pilot data when estimated with all observed variables ($\chi^2(295) = 420$; *RMSEA* = .067, CI_{lower} = .051; *CFI* = .90; *SRMR* = .065).

Target coefficient. The second-order model for *Teacher quality* was next tested for its ability to account for the covariances between the four first-order factors it comprised, using a fit index introduced by Marsh and Hocevar (1985) known as the 'target coefficient' (see also Cheung, 2000; Cheung & Ng, 2000; Spencer, Barrett & Turner, 2003). The target coefficient (TC) for a second-order model is calculated as "the ratio of the chi-square of the first-order model to the chi-square value of the more restrictive [second-order] model" (Marsh and Hocevar, 1985, p. 571). The TC is scaled, as such, from 0 – 1, where a value of 1 indicates that the second-order structure accounts perfectly for the covariance among factors in the first-order model. Comparing the chi-squared value of $\chi^2(293) = 468.119$ (Swain corrected $\chi^2(293) = 419.53$) for the four-factor, first-order model to the chi-squared value of $\chi^2(295) = 467.970$ (Swain corrected $\chi^2(295) = 419.54$) for the second-order factor model produced a TC of 1.00, which strongly supported the structural validity of *Teacher quality*.

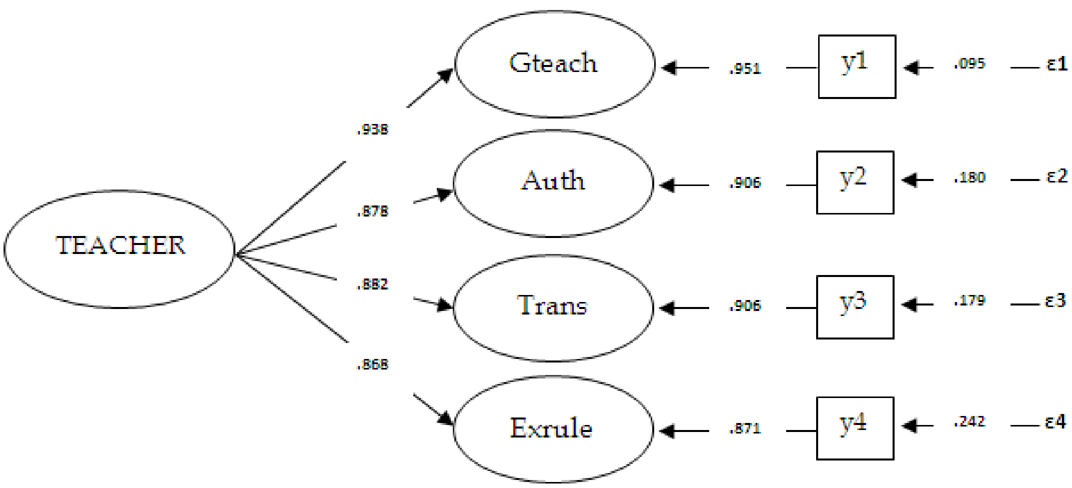

Figure 4.2. The confirmatory factor model for the higher-order factor *Teacher quality*.

The measures *Mastery goal structure* and *Clear goals and standards* contributed very little unique variance to the initial second-order factor structure, due to their extremely high correlations with *Good teaching*. Removing these two items improved parsimony, but rendered a second-order factor structure with unacceptably high *RMSEA*. When the measure *Appropriate assessment* was additionally removed, the second-order factor achieved good fit to the pilot data. The four remaining first-order factors formed a parsimonious second-order structure that represented three distinct dimensions of teacher quality: pedagogical skill, assessment quality, and behavior management.

4.7 Model 2

The originally hypothesized PTLC model (Model 1) was revised significantly based on the foregoing Pilot Study. The revised PTLC model was dubbed Model 2; see Figure 4.3).

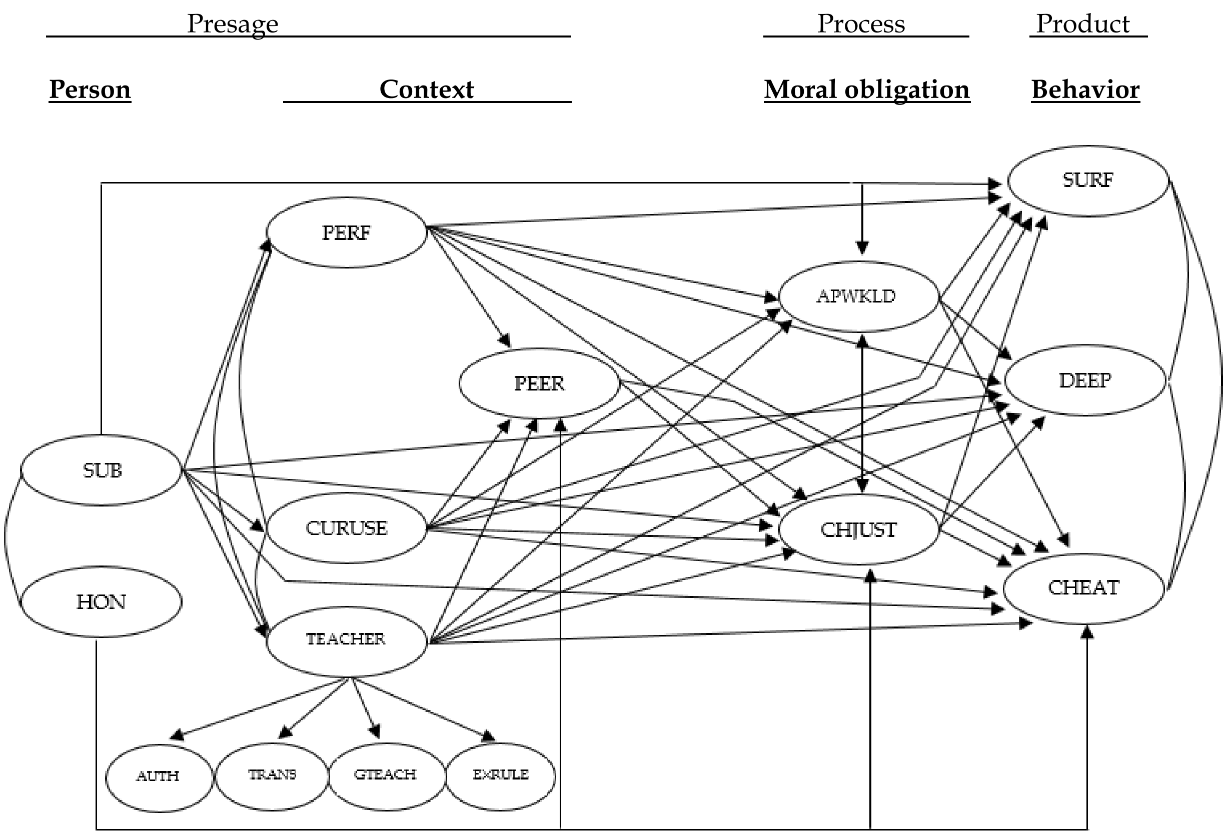

Figure 4.3. Model 2: The revised hypothesized PTLC structural model, situated within the 3-P Model framework (Biggs, 1993; Biggs et al., 2001), including all modifications made based on the results of congeneric modeling and analysis of the bivariate correlation matrix.

Removing *Appropriate assessment*, *Mastery goal structure*, and *Clear goals and standards*, from Model 1 reduced the number of first-order factors in Model 2 to fourteen. Multicollinear relations between *Authenticity* and *Transparency of assessment*, *Good teaching*, and *Experience of classroom rules* were additionally modeled as a single higher-order factor.

4.8 Chapter summary

The purpose of the Pilot Study was to test and refine the set of measures included in the original hypothesized PTLC model. Analysis proceeded in two stages: (1) congeneric model analysis, (2) correlational analysis, and (3) higher-order factor analysis. The congeneric model of each measure included on the original questionnaire instrument for this study was tested for fit, and, where necessary, modified. Modified congeneric models were then used to create weighted composite scores for the purpose of estimating a correlation matrix for the measurement model.

Correlational analysis in section 4.6 identified a cluster of seven multicollinear factors that appeared to represent aspects of students' perceptions of teacher quality. Two of these factors were removed due to extreme multicollinearity. A third (*Appropriate assessment*) was removed as a likely source of misfit in the second-order model. The resulting second-order factor structure, dubbed 'Teacher quality', was retained in Model 2.

CHAPTER 5

METHODOLOGY

Structural equation modeling (SEM) was used in the present study to test a set of causal hypotheses implied by the PTLC perspective on academic disintegrity developed in Chapter Three. Data was collected from a diverse sample of secondary students at American-curriculum international schools located outside of the United States, at two time points, separated by approximately twelve months. These data were used to assess the psychometric properties of the hypothesized measurement model, and to estimate the hypothesized structural model. Following cross-sectional analyses, matched samples from Times 1 and 2 were used to estimate an autoregressive longitudinal model.

5.1 Causal language

The use of causal language in SEM studies has been identified as a significant problem in educational research (Robinson, Levin, Thomas, et al., 2007). Scholars generally agree that non-experimental methodologies, such as SEM, are inappropriate for inferring causality (Biddle & Martin, 1987; Kline, 2009; Martin, 2011). Yet the use of causal language in reference to structural models is widespread for a variety of more and less appropriate reasons (Biddle & Martin, 1987; Robinson et al., 2007). Mueller and Hancock (2010) argue that an appropriate basis for using causal language in SEM research is when it "*is done from within the context of the particular causal theory proposed* and the possibility/probability of alternative explanations is raised unequivocally [italics in original]" (p.382). Causal language may, in other words, be appropriate when grounded in a defensible causal theory that has directly informed a structural model. Discussing a model in causal terms may be more appropriate, for instance, when the hypotheses it tests arise from experimental studies in which (1) cause-effect relations have been isolated from third variable effects, (2) the temporal precedence of the cause over the effect is established, and (3) the direction of causality (A → B; not A ← B) is determined (Kline, 2011; Mulaik, 2009; Pearl, 2000).

Experimental studies heavily inform the PTLC model. These studies indicate that dishonest behavior is determined in part by moral self-concept (Ariely, 2012; Gino et al., 2011; Mazar et al., 2008; Shalvi et al., 2011), and in academic spheres by perceptions of teacher quality (Day et al., 2011; Murdock & Anderman, 2006; Murdock et al., 2004, 2007), extrinsic and performance goal structures (Lobel & Levanon, 1988; Mills, 1958; Shelton & Hill, 1969; Taylor & Lewit, 1966), and perceived peer behaviors (Gino et al., 2009; Walker et al., 1966). The implications of these findings were incorporated in the hypothesized model as measures for *Honesty-trustworthiness self-concept*, the higher order factor for *Teacher quality*, *Performance goal structure*, and *Peer norms*, respectively. These factors were hypothesized, moreover, to predict the degree of moral obligation students feel to work hard and be honest in a manner consistent with non-rational contractarian intuition. This overarching hypothesis is supported by experimental findings that judgments of morality and justice tend to be heavily influenced by non-rational mental processes (Boles et al., 2000; Gino & Pierce, 2009; Gneezy, 2005; Pillutla & Murninghan, 1996; Shalvi et al., 2012), such as social contract thinking (Cosmides, 1989; Knoch et al., 2006).

The present study was designed to test hypotheses that reflect relations theorized to be of a causal nature. While its non-experimental, passive-observational SEM research design was not appropriate for inferring causality, the hypothetical model was heavily grounded in prior experimental research. This considerably strengthened the rationale for advised use of causal language in the present study by meeting Mueller and Hancock's (2010) abovementioned stipulation that such language be situated within the context of causal theory.

5.2 Outline of empirical analyses

The first analyses of data related to this research were conducted in a planned pilot study (see Chapter Four) focused on ensuring the psychometric validity and reliability of measures in the hypothesized measurement model. Measures that demonstrated poor fit or reliability, or that were multicollinear with other measures, were either modified, dropped, or combined into second-order factors. The result of these changes to Model 1 was dubbed Model 2 (see Figure 4.3).

Structural modeling of Main Study data utilized a two-phase approach described by Mueller and Hancock (2010). In Phase One, basic descriptive properties of latent factor measures comprised by Model 2 were analyzed, both individually and in groups, using *version 21* of the *SPSS* statistics package and *version 7.11* of the M*plus* statistical modeling program. The reliability of individual measures was assessed with a method devised by Raykov (2004, 2009) that takes into account the weighted contributions of individual scale items. Confirmatory factor analyses (CFA) were then conducted of one-factor, 'congeneric', models, followed by analyses of the full multivariate measurement model. CFA was also used to test multi-group factorial invariance across gender and grade-level groups, and finally to identify differences in factor means related to demographic variables such as English proficiency and parental educational attainment, in a multiple-indicators multiple-causes (MIMIC) model.

Congeneric factor model misfit detected in Phase One was addressed by either dropping or modifying measures. Prevalent differences in factor means occurring across groups were addressed by incorporating grouping variables into the hypothesized model. These changes to Model 2 resulted in Model 3 (see Figure 6.1).

Phase Two of the modeling process involved analysis of the structural characteristics of Model 3. Factorial non-invariance across gender groups detected in Phase One prompted analyzing the model with gender-specific data prior to estimating it with the combined, or 'co-ed' sample. Prominent differences between male and female structural models were noted at Times 1 and 2.

5.2.1 Missing data

Questionnaires were administered online and by paper. Respondents to online questionnaires provided uniformly complete data sets, as attempting to leave an item unanswered triggered a prompt to respond, and prevented a respondent's progress to the next question. Incomplete electronic questionnaires occurred only when participants simply quit at some point during the questionnaire, leaving the subsequent sections blank. There was a limited amount of missing data on paper-based questionnaires. Sixty paper-based questionnaires including missing data were completed at Time 1, and 20 at Time 2. Data screening at both times entailed eliminating a handful of obviously invalid answer sets (e.g., all 5s), and all incomplete electronic questionnaires. Data missing from paper-based questionnaires was imputed using the Multiple Imputation function in SPSS.

Paper-based data were first tested, at both times, to ensure that data was missing at random. When the randomness of missing data was tested for these samples as wholes, Little's 'missing completely at random' (MCAR) test was highly significant, indicating that data was *not* missing at random. When the data from paper-based questionnaires were analyzed separately, however, Little's MCAR test was non-significant at both Time 1 (χ^2 = 71.417, df = 1846, p = 1.000) and Time 2 (χ^2 = .000, df = 176, p = 1.000). This indicated that, among paper-based questionnaires, data was missing completely at random. Missing data imputation could, therefore, proceed.

5.3 Participants

5.3.1 Time 1 sample

Time 1 data was collected by questionnaire during May and June of 2012 (the end of the 2011-2012 school year for participating schools). 493 students completed questionnaires, of which 201 were male, 292 were female; 277 were in Grade Eight, and 216 were in Grade Nine (see Table 5.1). All eleven participating schools were American-curriculum international schools, hereafter 'American international schools' or 'international schools', listed with the US Office of Overseas Schools, of which seven received direct assistance from the US State Department. All participating schools were also accredited by widely-recognized accrediting agencies in the United States. American international schools were chosen for the present research because their broad ethnic diversity would enhance the generalizability of findings; their administrative autonomy would make them more readily accessible; and because the doctoral researcher for this project was intimately familiar with the American international school community, having served as a teacher and administrator in such schools throughout most of the prior decade. American international schools in the present study were located in nine countries, including two in Eastern Africa, one in Western Europe, two in Eastern Europe, three in Eastern Asia, and one in Western Asia. The names of these countries cannot be given as it would compromise the schools' identities.

Students at Eastern Asian schools were disproportionately represented in the present study, accounting for 71% of the Time 1 sample, as compared to 9% from Eastern Africa, 9% from Western Asia, 9% from Eastern Europe, and 2% from Western Europe. The effect of this regional bias was mitigated, however, by the multicultural makeup of individual international schools. International schools generally cater to a more nationally and ethnically diverse population than host country schools. Less than 38% of respondents reported, for instance, that an Eastern Asian language was the primary language spoken in their homes (Chinese dialects: 22%, Korean: 10%, Japanese: 5%, Tagalog: .4%, and Vietnamese: .4%), as compared to 44% who listed English as the primary language in their homes, and 18% who listed one of twenty other languages from outside of Eastern Asia. Inasmuch as linguistic diversity indicates ethnic and national diversity, the disproportionate representation of schools in Eastern Asia did not carry over to the ethnic and national makeup of the sample. The American international schools in this study adhere, moreover, to curricular programs of American and European origin, rather than those of their host countries. Examples include the *American Education Reaches Out* (AERO) curriculum, the College Board's *Advance Placement* Programs, and the *International Baccalaureate* program. The Time 1 sample was, in sum, broadly representative of ethnically and nationally diverse American international schools.

5.3.2 Time 2 sample

Time 2 data was collected in May and June of 2013 from 297 students who had also completed a questionnaire at Time 1. The retention rate from Time 1 was 60%. The loss of 40% of the sample, year-over-year, is largely attributable to administrative turnover that occurred at several participating schools. Incoming heads of school and divisional leaders appear, in some cases, to have not been fully aware, or enthusiastic, of their school's participation in the project. In such cases, students may not have been given sufficient advanced notice that the questionnaire was coming up, and/or may have encountered scheduling conflicts that prevented them from participating.

All eleven American international schools represented at Time 1 were also represented at Time 2. Sample composition at Time 2 resembled that at Time 1 in terms of the ratio of males (39%) to females (61%), the regional distribution of respondents (73% in Eastern Asia), and parental educational attainment (see Table 5.1). The incidence of languages most spoken in the home shifted, however, away from Eastern Asian languages, which were listed by approximately 34% of respondents (Chinese dialects: 20%, Korean: 7%, Japanese: 6%, Tagalog: 1%; Vietnamese: .3%), in favor of English, which was listed by approximately 53% of respondents. A correspondingly higher proportion of Time 2 respondents also rated their English skills as fluent (48%), as compared to Time 1 (37%), suggesting that Time 1 respondents who were less confident of their English language skills were less likely to participate at Time 2.

Table 5.1

Characteristics of the Time 1, Time 2, and Longitudinal samples

	Time 1 sample (N = 493)	Time 2 sample (N = 297)	Longitudinal sample (N = 225)
	Grade level		
Grade Eight	277 (56%)	N/A	
Grade Nine	216 (44%)	150 (50%)	123 (55%)
Grade Ten	N/A	147 (50%)	102 (45%)
	Gender		
Male	201 (41%)	115 (39%)	72 (32%)
Female	292 (59%)	182 (61%)	153 (68%)
	School location		
E. Asia	349 (71%)	218 (73%)	165 (73%)
E. Africa	44 (9%)	25 (8%)	19 (8%)
W. Asia	44 (9%)	11 (4%)	7 (3%)
E. Europe	44 (9%)	33 (11%)	24 (11%)
W. Europe	12 (2%)	10 (3%)	10 (4%)
	Self-rated English language proficiency		
Fluent	184 (37%)	141 (48%)	72 (32%)
High proficiency	203 (41%)	103 (35%)	128 (57%)
Intermediate proficiency	96 (20%)	45 (15%)	19 (8%)
Low proficiency	7 (1%)	6 (2%)	6 (3%)
Beginner	1 (.002%)	2 (1%)	0 (0%)

Table 5.1, continued

Characteristics of the Time 1, Time 2 and longitudinal samples

	Time 1 sample (*N* = 493)	Time 2 sample (*N* = 297)	Longitudinal sample (*N* = 225)
	Language spoken most at home		
English	217 (44%)	157 (53%)	122 (54%)
Chinese	109 (22%)	58 (20%)	41 (18%)
Korean	51 (10%)	22 (7%)	12 (5%)
Japanese	22 (5%)	19 (6%)	17 (8%)
Arabic	22 (5%)	4 (1%)	2 (1%)
French	8 (2%)	3 (1%)	2 (1%)
Other	64 (13%)	34 (12%)	29 (13%)
	Maternal educational attainment		
Advanced degree	135 (27%)	80 (27%)	67 (30%)
University of college degree	241 (49%)	145 (49%)	109 (48%)
Trade or apprenticeship training	4 (1%)	2 (1%)	3 (1%)
High school diploma or certificate	36 (7%)	36 (12%)	31 (14%)
Less than a High school diploma	9 (2%)	7 (2%)	4 (2%)
I don't know	68 (14%)	27 (9%)	11 (5%)
	Paternal educational attainment		
Advanced degree	252 (51%)	148 (50%)	113 (50%)
University of college degree	168 (34%)	113 (38%)	91 (40%)
Trade or apprenticeship training	3 (1%)	3 (1%)	2 (1%)
High school diploma or certificate	14 (3%)	7 (2%)	6 (3%)
Less than a High school diploma	4 (1%)	2 (1%)	1 (.4%)
I don't know	52 (11%)	24 (8%)	12 (5%)

5.3.3 Longitudinal sample

Students in the longitudinal sample transitioned to higher grade levels, either from Grade Eight to Grade Nine, or Grade Nine to Grade Ten, between Times 1 and 2. Of the 297 respondents to the Time 2 questionnaire, 225 provided personal identification codes that could be matched with their responses at Time 1. This sample included students from nine of the original eleven schools, but remained consistent with the cross-sectional samples in terms of grade-level ratio (55% Grade Eight/Nine students, 45% Grade Nine/Ten students), regional distribution (73% Eastern Asia), and parental educational attainment (see Table 5.1). It is important to note that in all nine participating schools, Grade Eight and Grade Nine are encompassed by different divisions, Middle School and High School, respectively. Students who transitioned from Grade Eight to Grade Nine during the course of the study thus also experienced a broader environmental transition, from Middle School to High School, whereas students who transitioned from Grade Nine to Grade Ten did so within the High School environment.

The proportion of respondents in the longitudinal sample who indicated that English was the language most spoken in their homes (53%) remained consistent with the Time 2 sample (54%), while the proportion who rated their English language skills as 'fluent' was markedly lower among longitudinal respondents (32%). The gender ratio in the longitudinal sample also differed from the cross-sectional samples, with fewer males represented (32%) than females (68%). Table 5.1 presents a comparison of these sample characteristics.

5.4 Procedure and participation

Approval for this research was granted by The University of Sydney Human Research Ethics Committee (HREC) (Protocol Number 14193; see Appendix AG), on the basis of the project's scientific merit and appropriate planning for the safety and dignity of all participants. Schools and individual participants were guaranteed, *inter alia*, both anonymity, and the right to withdraw from the project at any time, without repercussion.

Senior school administrators at approximately thirty international schools were contacted initially by phone, followed by email and postal correspondence. Schools that agreed to participate asked their students in Grades Eight and Nine to sign, and have their parents sign, a consent form on which they either opted-in or opted-out of the two-year study. The consent forms of students who opted-in were returned to the University of Sydney.

Questionnaires were administered at participating schools in 2012 and 2013, during late May and early June, which is the end of the American school year. Questionnaire administration was conducted either online, through SurveyMonkey.com, or by hardcopy. Nine schools opted to use the online format, accounting for 88% of the Time 1 sample, whereas two schools opted for hardcopies, accounting for 12% of the Time 1 sample.

Time 2 questionnaire modification. With the permission of HREC, thirty-three items that had been eliminated either during the Pilot Study (see Chapter Four), or during preliminary measurement analyses conducted at Time 1 of the Main Study, were dropped from the questionnaire used at Time 2, and twenty-three new items were added (marked with asterisks in Appendix D). The twenty-three new items included eleven items composing three psychological need satisfaction measures (autonomy, relatedness, and competence) (Reeve & Tseng, 2011), that have been developed in the field of self-determination theory (Deci & Ryan, 2011; Jang, Reeve, Ryan, & Kim, 2009), and twelve potential candidate-items for measures that demonstrated marginal fit in the Pilot Study, including *Surface learning strategies*, *Appropriate workload* and *Justifiability of cheating*. Analysis of these items is strictly exploratory and is not reported in the present study.

This reduced, in total, the number of items on the questionnaire from 105 at Time 1 (see Appendix B) to ninety-five at Time 2 (see Appendix D). Schools were notified of the change in advance of administering the Time 2 questionnaire. Items added to the Time 2 questionnaire are marked with an asterisk in Appendix D.

5.5 Modeling approach

5.5.1 Preliminary analyses

Many of the measures employed in the present study had not been included previously in structural equation models or in research at the secondary level. Each of these measures was selected, based on a prior track record of good psychometric performance, to represent an important aspect of the hypothesized model. Given their novelty within structural modeling research of secondary school phenomena, the overall plan of the present study incorporated extensive preliminary analyses, including a pilot study, as well as modification or exclusion of measures where necessary, in order to ensure the validity of constructs incorporated in the hypothesized model.

Congeneric modeling. One-factor congeneric models were of crucial importance to the present study because they served as the basis for calculating weighted composite scores (Holmes-Smith & Rowe, 1994; Rowe & Hill, 1998). Composite scores were necessary in the present study to estimate models for which ratios of sample size (N) to free model parameters (q), or '$N:q$ ratios', would otherwise have been inadequate. Unacceptable fit of congeneric models was addressed either with *post-hoc* modifications such as allowing observed indicator error terms to co-vary, or by dropping malfunctioning factors from the study. *Post-hoc* modifications were avoided wherever possible, however, as they cast doubt on the generalizability of structural model results, especially when the structural model is tested with the same data used to modify the latent variables (Bandalos & Finney, 2010; Byrne, 2012; Kline, 2011).

5.5.2 Modeling procedure

Structural equation modeling was used to test the hypothesized PTLC model. Cross-sectional data was analyzed separately at each of two time points, and then used to estimate a matched-samples longitudinal model. Model analysis followed a two-phase approach, in which the validity and reliability of measurement models were tested first, including analyses both of one-factor congeneric models and of the multivariate measurement model, and hypothesized structural models were tested second (Mueller & Hancock, 2010).

Confirmatory factor analysis. Confirmatory factor analysis (CFA) is a method of assessing the construct validity, or 'fit', of a latent measurement model to a particular data set. CFA will be used in the present study to assess both congeneric factor models, and multivariate measurement models.

CFA tests *a priori* measurement hypotheses in which observed variables are generally allowed to load only on the latent construct they were designed to measure. Construct validity, which entails how well a measurement model represents an intended construct or constructs, is assessed by comparing the variance-covariance structure of the hypothesized latent model to the variance-covariance structure freely observed in the data (Byrne, 2012). A higher degree of similarity between these two variance-covariance structures indicates better 'fit' of the hypothesized latent model to the data, and bolsters the argument that the measure or measures in question are valid (Byrne, 2012). The degree of misfit of a measurement model is reported by a variety of fit statistics (discussed in detail in section 5.8, below). CFAs of multivariate models are strict tests of the models' *a priori* measurement hypotheses, because the included factors must fit the data, as a group, while also retaining their predicted factorial structure with respect to one another.

Congeneric CFA. One important type of measurement structure analyzed with CFA in the present study was the one-factor congeneric model. One-factor congeneric model analysis is a special case of CFA, in which the fit of an individual latent model is interrogated. A poorly-fitting congeneric model casts doubt on the validity of both the construct it represents, and any statistical relations between it and the other factors in a hypothesized structural model. Verifying that individual measures adequately represent their intended constructs ultimately informs the validity of the multivariate models they compose.

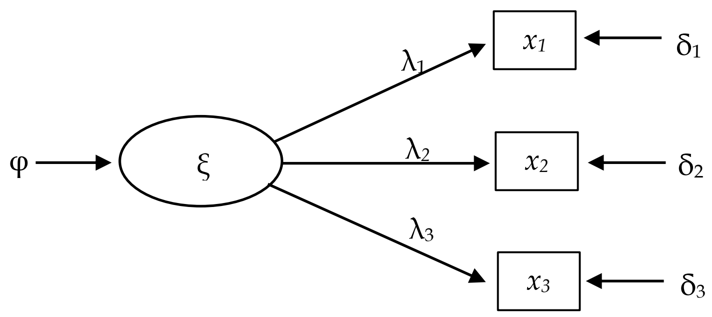

Figure 5.1. A three-indicator congeneric model.

A simple, three-indicator congeneric model is provided in Figure 5.1. The three indicators, represented by x's, are hypothesized to reflect a single unobservable, or 'latent' variable, represented by ξ. The congeneric model in Figure 5.1 can also be expressed as a set of three linear expressions of the form $x_i = \lambda_i\xi + \delta_i$, where the λ's are the regression terms, or loadings, of the observed variables on the latent variable, and the δ's represent the random error associated with each observed variable. These three linear equations may be condensed, furthermore, to a single equation by using matrices to group observed variables (x's), factor loadings (λ's), and error term variances (δ's), respectively (see equation 1).

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_1 \\ \lambda_3 \end{bmatrix} \begin{bmatrix} \xi \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix} \tag{1}$$

In addition to the regression equation for the congeneric model exampled above, a variance-covariance matrix known as the Theta-delta matrix is estimated for the model's error terms. The Theta-delta matrix for the congeneric model in Figure 5.1 (see Figure 5.2) includes estimates of error term variances along the diagonal, and error term covariances beneath the diagonal. Error covariances are usually constrained to zero, as shown in Figure 5.2, but may be permitted, or 'freed', when there is a strong theoretical or methodological reason for doing so (Kline, 2011).

$$\boldsymbol{\Theta}_\delta = \begin{bmatrix} \Theta^\delta{}_{11} & & \\ 0 & \Theta^\delta{}_{22} & \\ 0 & 0 & \Theta^\delta{}_{33} \end{bmatrix}$$

Figure 5.2. A Theta-delta (Θ_δ) matrix, containing variances and covariances for the error terms of observed indicator variables.

Multivariate CFA. In multivariate CFA, a measurement model that includes more than one latent variable is assessed in terms of its fit to a particular data set. This is done by estimating all factors in the set simultaneously, while freeing all possible inter-factor covariance parameters. Researchers may also, under special circumstances, permit observed variables, represented as *x*'s in Figure 5.1, to load onto more than one latent variable, or 'cross-load'. Each observed variable is, however, normally restricted to loading only on the latent variable it was designed to measure.

Higher-order CFA. A special case of multivariate CFA involves the inclusion of higher order factors (Gray, 1997). Higher-order latent factors are hypothesized to give rise to groups of lower-order latent variables, in the same sense that first-order latent factors are hypothesized to give rise to observable variables. Higher-order factors may be hypothesized *a priori*, in order to explain groups of theoretically similar lower-order factors, or may be fitted *a posteriori* as a means of accounting for clusters of highly correlated, or multicollinear, factors (Schumacker & Lomax, 2010). Higher-order factors serve in SEM to specify theoretically sound underlying causes for large statistical associations between latent variables.

5.5.3 Multi-group measurement invariance

Research samples in the social sciences typically comprise multiple subgroups (e.g. gender, grade-level, political affiliation, language). The tendency for questionnaire measures to function differently across subgroups, and thus to convey variations in meaning depending on the subgroup in question, is an increasingly recognized problem in SEM literature (Byrne, 2012). The assumption of measurement invariance becomes especially important when the interpretation of between-group differences is a research objective (Cheung & Rensvold, 2002; Meade, Johnson, & Braddy, 2008; Byrne & van de Vijver, 2010).

Multi-group measurement invariance is also an important assumption of employing single-indicator weighted composite scores in place of multiple-indicator factor models (Holmes-Smith, 2012). Weighted composite scores are computed from factor score coefficients derived from the sample as a whole, and used to represent the variance of multiple-indicator factors as single scores. Inasmuch as patterns of responding on a given measure are inconsistent between the subgroups of a sample, weighted composite scores may misrepresent the factors' operational meanings within the model.

A number of published works indicate the likelihood of two potential sub-groups with respect to academic cheating among secondary students: (1) grade-level (e.g. Brandes, 1986; Anderman & Midgley, 2004; Galloway, 2012; Schab, 1969, 1991), and (2) gender (e.g. Calabrese & Cochran, 1990; Canning, 1956; Davis, 1973; Finn & Frone, 2004; Galloway, 2012; Schab, 1980). A third potential concern was longitudinal invariance, or whether factorial structure was equivalent for respondents at Times 1 and 2.

The strategy for assessing multi-group measurement invariance involves testing a succession of measurement models, in which increasingly restrictive sets of parameters are held invariant across groups. The sequence typically begins with fitting *baseline models* for each of the subgroups, separately. A *configural model* is typically tested next, in which the form, or configuration, of indicators and factors is held constant across both groups, simultaneously, while allowing parameter estimation to occur uniquely for each group. Subsequent models most commonly test the invariance of factor loadings (*metric model*) and observed variable intercepts (*scalar model*) (Byrne, 2012; Vandenberg & Lance, 2000). A fourth model tested in the present study held factor variances equal across groups, in order to assess the appropriateness of representing them with composite scores.

A difference in fit between the configural model and any of the subsequent invariance model indicates that the groups in question are not perfectly invariant. Factorial invariance is not, however, treated as a zero-sum phenomenon, but as a matter of degree (Byrne et al., 1989; Cheung & Rensvold, 2002; Marsh & Hocevar, 1985). A decline in *CFI* of less than .01 (i.e. *CFI* < -.01) is widely cited as a standard for assuming that a measurement model is invariant (Cheung & Rensvold, 2002; Kline, 2011). This standard will be applied in the present study.

Differential item functioning analysis. A second means of analyzing measurement invariance between groups that was utilized in this study is differential item functioning (DIF) analysis (Wang & Wang, 2012). DIF analysis is used to analyze group differences by regressing all factor indicators in a given measurement model on grouping variables, such as male = 0 and female = 1. Factors whose item means differ significantly by group membership are functioning at least somewhat differently across groups. DIF analysis thus enables the identification of where, specifically, factorial non-invariance is concentrated within a measurement model. This is consistent with arguments by Byrne, Muthén, and Shavelson (1989) in favor of methods "for pinpointing the source of inequality within the offending matrix" (p. 457). Identifying where group differences occur at the item-level enables assessment of their relative importance within a model. Group differences in factorial structure that are concentrated in outcome variables may, for instance, be of greater concern than in other variables in a given model.

5.5.4 Covariate analyses

Covariates will be modeled at Times 1 and 2 using a 'multiple-indicators multiple-causes' (MIMIC) methodology. MIMIC models (see Figure 5.3) entail regressing a set of latent variables on a set of covariates and interaction terms, in order to investigate mean-level differences between groups (Grayson, Mackinnon, Jorn, et al., 2000; Wang & Wang, 2012). It is important to distinguish between the two types of mean difference discussed in this chapter: (1) those that occur at the level of whole factors, which imply group differences in extent (e.g. males tend to rate their teachers more favorably than females) and (2) those that occur at the level of individual items, as might imply differences in factor structure, or in factors' operational definitions (e.g. males tend to emphasize different qualities than females in rating teachers). Groups may have different factor means for factors that are structurally and functionally invariant between them, and *vice versa*.

Figure 5.3 presents an example MIMIC model in which two latent factors (*Subject self-concept* and *Honesty-trustworthiness self-concept*) are regressed on a set of covariates and two-way interaction terms. MIMIC models tested at Times 1 and 2 included the full multivariate measurement model regressed on a set of covariates including three demographic variables (English proficiency, maternal educational attainment, and paternal educational attainment) and two dichotomous grouping variables, gender and grade-level, dummy coded respectively as 1 = males, 2 = females; and 1 = Grade Eight, 2 = Grade Nine. MIMIC models also included all possible two-way interactions between these covariates. To avoid the potential for multicollinearity among covariates and interaction variables (Aiken & West, 1990), the three demographic variables mentioned above were mean-centered (M = 0, SD = 1) prior to the calculation of interaction terms.

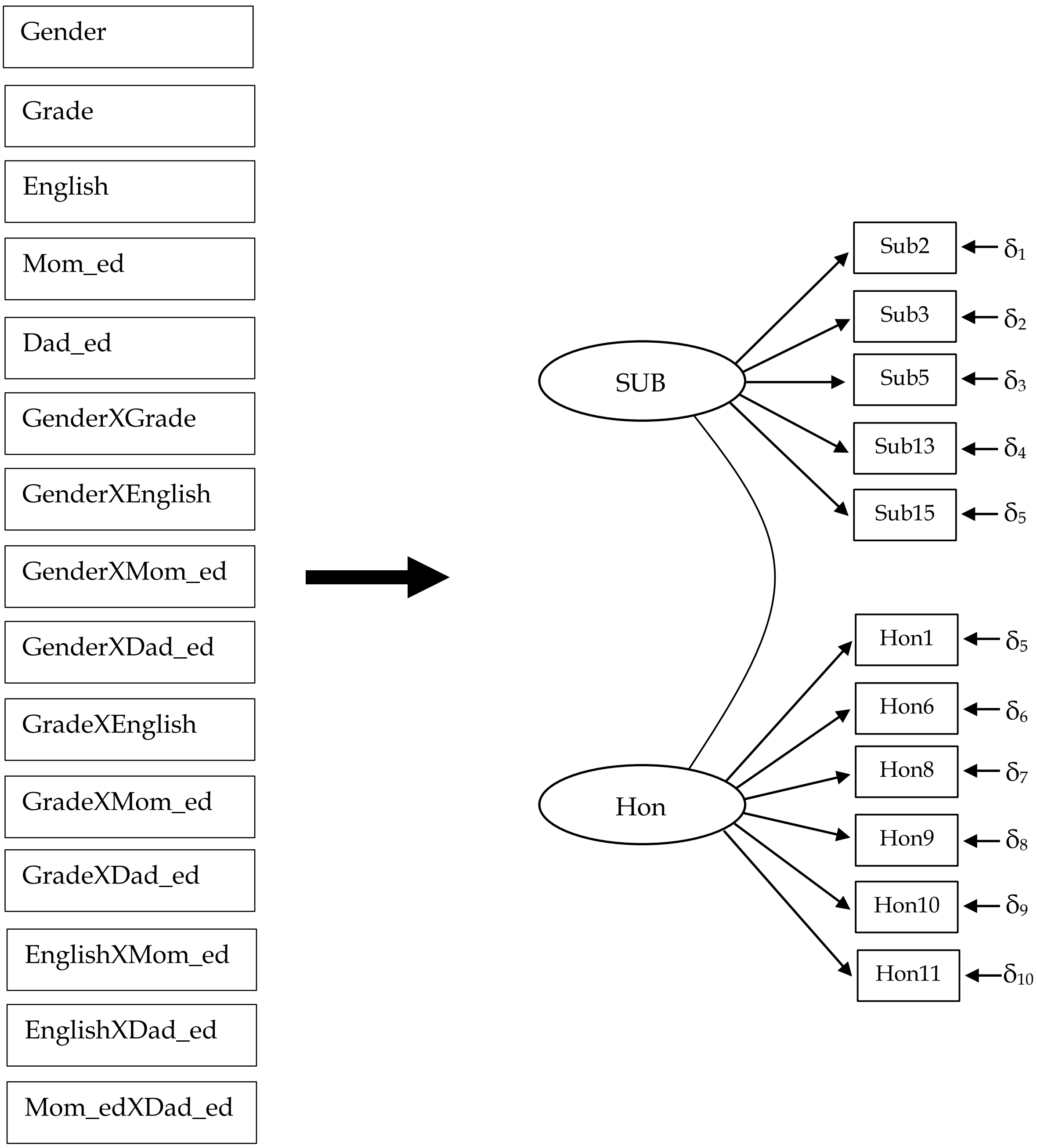

Figure 5.3. An example of a multiple-indicators multiple-causes (MIMIC) model. Grade = grade-level; English = English proficiency; Mom_ed = mother's educational attainment; Dad_ed = father's educational attainment; X indicates two-way interaction terms.

5.6 Modeling approach: Phase 2

5.6.1 Structural modeling

Once the validity of constructs in the measurement model was established with CFA-based analyses, the hypothesized model proceeded to the structural phase of the modeling process (Anderson & Gerbing, 1988; Mueller & Hancock, 2010). Structural models combine the measurement components of factor analysis, as shown in Figure 5.4, with inter-factor regression equations, or beta paths. The use of variance-covariance matrices allows simultaneous estimation of numerous inter-factor regression hypotheses among latent factors, such as psychological variables.

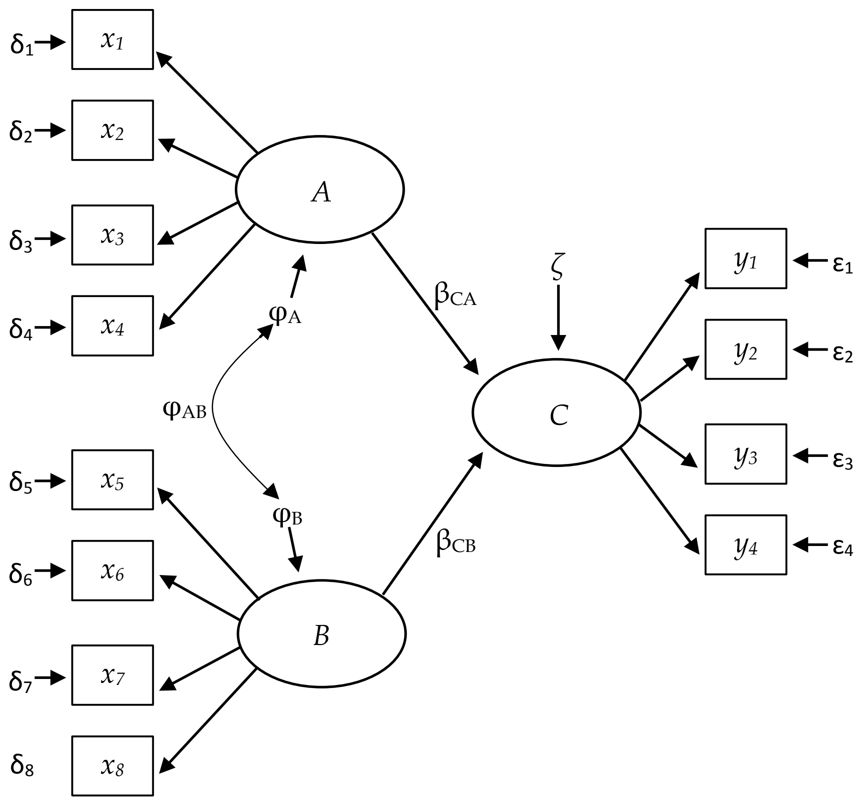

Figure 5.4. An example of a simple structural model with inter-factor regression hypotheses modeled as beta paths (β_{CA}, β_{CB})

The measurement portion of the model in Figure 5.4 consists of two exogenous latent variables, *A* and *B*, measured by four observed variables each (x_1 – x_8), and one endogenous latent variable, *C*, also measured by four observed variables (y_1 – y_4). Error terms (δ and ε, respectively) are calculated for all observed variables, as well as for the exogenous and endogenous latent variables (φ and ζ, respectively). The structural part of this model, which includes the beta paths from *A* and *B* to *C* is, therefore, purged of measurement error (Muthén, 2002). Differences in model fit due to the insertion of structural paths among constructs are, therefore, attributable to how well such paths fit the data, such that good model fit provides strong support for the hypothesized structural relations.

The model in Figure 5.4 can also be written as a set of fourteen linear regression equations. Twelve of these equations, of the form $Obs_i = \lambda_i(\text{Latent}) + Error_i$, describe the measurement portion of the model, where *Latent* represents the latent variables (*A, B*, or *C*); *Obs* represents the observed indicator variables (x's or y's); λ represents the observed variable's loadings onto their respective factors; and *Error* represents the error, or 'residual', terms of the observed indicator variables (Coote, 2011).

5.6.2 Longitudinal modeling

Cross-sectional models, as illustrated in Figure 5.4, above, offer snapshots of construct measures and structural associations at specific points in time, but afford little opportunity to explore processes of change or to isolate potential causal agents (MacCallum & Austin, 2000; Martin, 2011). Modeling data at multiple time points, or longitudinally, can, by contrast, illustrate growth processes, help isolate causal agents, and provide a more justifiable basis for prescriptive statements (Farrel, 1994; MacCallum & Austin, 2000; Martin, 2011).

Repeated-measures longitudinal models fall into two broad categories. Growth curve models, the first category, are used to explore linear and nonlinear change trajectories in variables over time. Autoregressive models, the second category, involve regressing constructs at later time points on their counterpart measures at earlier time points (McArdle & Aber, 1990). This provides a more rigorous test of the predictive validity of a cross-sectional model at later times by removing prior variance as a function of autoregressive paths (MacCallum & Austin, 2000). Removing prior variance, as may emanate from underlying personological variables such as personality structures and self-beliefs not otherwise included in the model, better isolates the strength of effects that is unique to the latter time point. Autoregressive models also afford the opportunity to test whether autoregressive paths are salient over predictive effects operating within later time points.

The salience of an effect refers to its predominance over other effects in predicting an outcome variable. While strong causal claims are principally the domain of replicable experimental manipulation, finding that longitudinal effects on a construct measured at later time points are stronger than, or 'salient over', effects operating within the later time point, or *vice versa*, suggests they have temporal precedence, which may also evince causal agency (Pearl, 2000), and justify prescriptive statements (Martin, 2011). Longitudinal modeling of phenomena in complex, real-world settings can, as such, complement experimental methodologies for investigating causal hypotheses.

The model of cheating behavior hypothesized in the present study reflects resurgent emphasis in integrity literature in the effects of relational factors on academic integrity. To this end, an autoregressive, repeated-measures longitudinal design, of the general form presented in Figure 5.5, afforded the opportunity to test the relative effects of contextual and personological variables in relation to academic integrity.

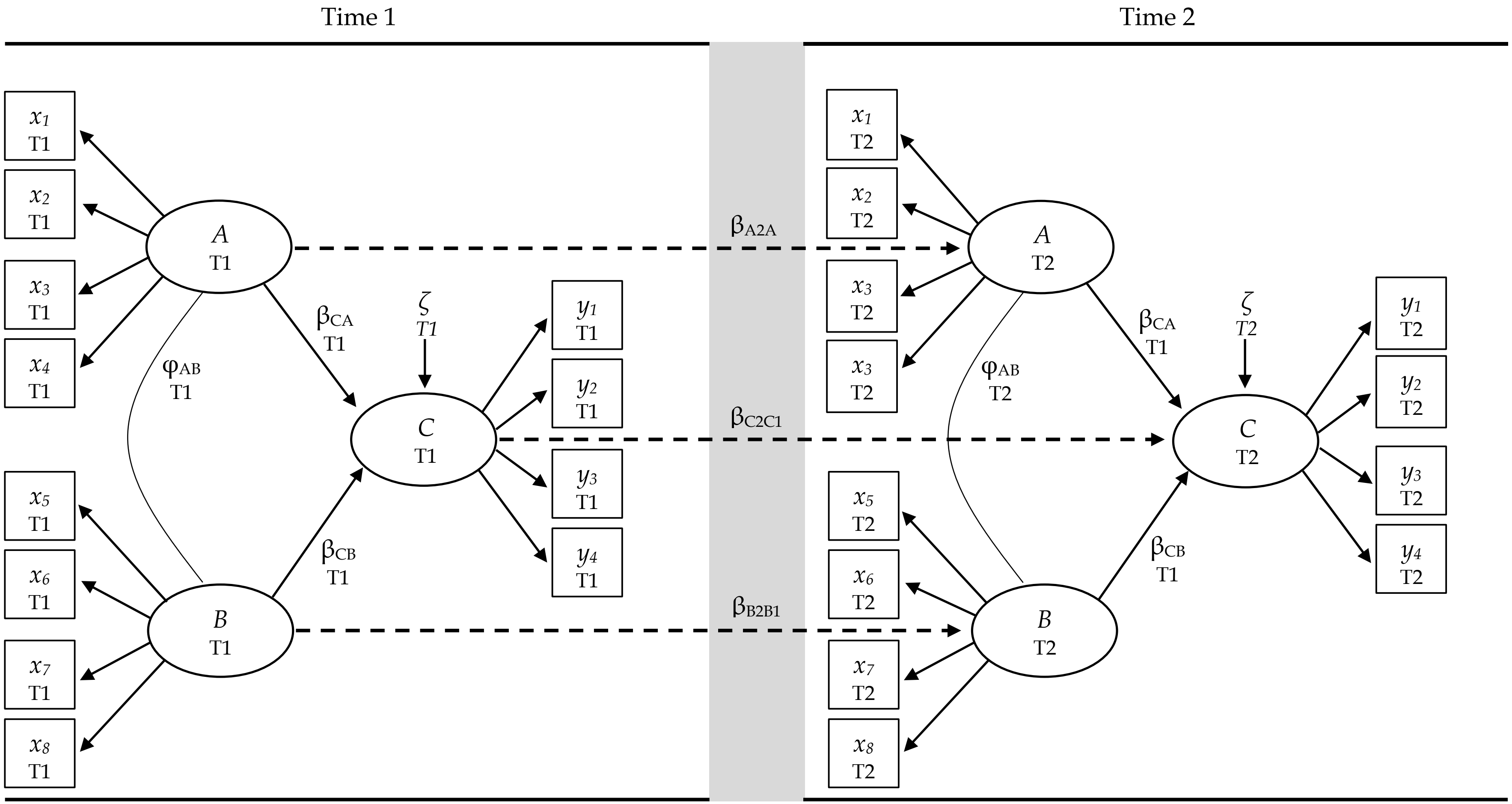

Figure 5.5. An example of a simple longitudinal structural model, in which variance in latent constructs at Time 1 (T1) is accounted for at Time 2 (T2). Prior variance paths; Hypothesized paths.

5.7 Sample size

An important consideration in structural equation modeling research is the adequacy of sample size, which tends in the present study towards the smaller end of the acceptable range, both in absolute terms (N_{time1} = 493; N_{time2} = 297; $N_{\text{Long.}}$ = 225), and in terms of the complexity of the structural models tested, which ranges from 178 and 399 free model parameters. Low sample sizes tend to negatively bias model fit indices, and can cast doubt on the trustworthiness of model results (Herzog, Boomsma, & Reinecke, 2007; Jackson, 2003; Kline, 2011). Two means of addressing the potential negative effects of small sample size are used herein. Firstly, a small sample-size adjustment to fit indices will be conducted on more complex models using the *Swain-R, version 1.2* software package (Boomsma & Herzog, 2009, 2013). Secondly, more complex models will be estimated with single-variable weighted composite scores computed according to the methodology developed by Holmes-Smith and Rowe (1994; Holmes-Smith, 2012: Course notes)

5.7.1 Recommended sample size

Recommendations for appropriate sample size are split between those that emphasize *N:q* ratios, or the ratios of sample size (*N*) to free model parameters (*q*), and those that emphasize the size of the overall sample. Kline (2011) asserts that *N:q* ratios should be the principal focus for analyses that use the maximum likelihood (ML) estimator. The present study uses a version of ML with robust standard errors (MLR) (Chou, Bentler & Satorra, 1991) for models estimated with observed indicator variables, and ML for models estimated with normalized composite variables. *N:q* ratio was, therefore, an important consideration throughout the program of empirical analysis. Recommendations for appropriate *N:q* ratios range from 20:1 (Costello & Osborne, 2005), to 10:1 (Raykov & Marcoulides, 2006), to 5:1 (Bentler & Chou, 1987), whereas recommendations for appropriate overall sample size range from 100 (Anderson & Gerbing, 1988; Ding, Velicer & Harlow, 1995), to 5,000 (Hu, Bentler &

Kano, 1992). Falling short of such recommendations may cast doubt on the "trustworthiness of results" (Kline, 2011, p. 12).

5.7.2 Swain R small sample correction

The fit of complex structural equation models tends to be penalized when small samples are used (Herzog, Boomsma, & Reinecke, 2007). The bias caused by small sample size is corrected in the present study with the *Swain-R, version 1.2* software package (Boomsma & Herzog, 2013). The Swain small sample correction uses degrees of freedom, sample size, number of observed variables, and χ^2 values of both the specified and baseline models, to generate a Swain correction factor (*SCF*) between 0-1. The Swain correction factor is then used to correct the model chi-squared, and associated fit statistics, such as *CFI*, *TLI* and *RMSEA*.

Small sample corrections will be applied to models such as multivariate confirmatory factor analyses and structural models estimated with observed indicator variables. The *SCF*, which represents the amount by which the chi-squared value is adjusted, will be reported whenever a Swain correction is applied.

5.7.3 Normalized weighted composite scores

Normalized weighted composite scores were employed extensively in the present study in order to fit complex models with comparatively small samples. Such composite scores sum the values of observed indicators used to measure latent factors. They are calculated, in the present study, from factor score coefficients that reflect the unique proportional contribution, or 'weight', of each observed indicator to the overall variance of a latent factor (Holmes-Smith & Rowe, 1994; Holmes-Smith, 2012; Rowe & Hill, 1998).

Normal equivalent deviates. Composite scores were then transformed to normal equivalent deviates (NEDs) in SPSS, based on the recommendations of Rowe (2002, 2004). NEDs retain the rank order of scores, standard deviations, and means of 'raw' composite factors, while ensuring that they are measured on a common metric, which improves the comparability of effect sizes (Rowe, 2002, 2004; Dolmans & Ginns, 2005). The syntax for generating NEDs in SPSS is provided in Appendix A.

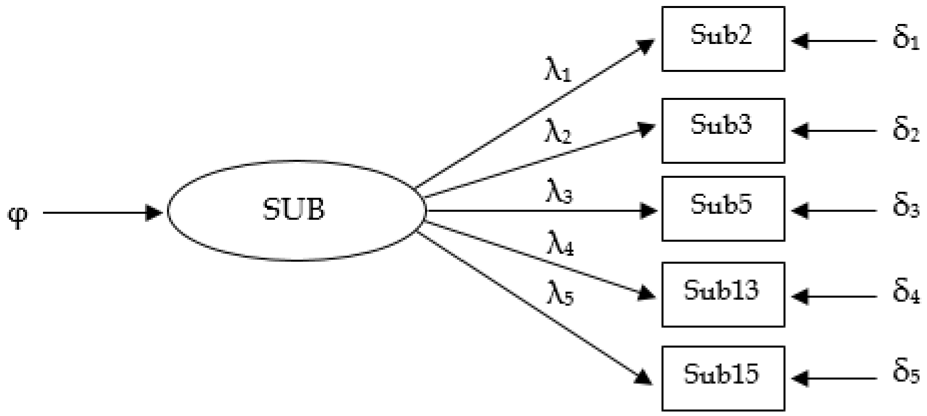

Figure 5.6. The congeneric model for Subject self-concept at Time 1, as a basis for computing weighted composite scores.

Weighted composite score computation. Figure 5.6 presents a congeneric CFA model of the latent factor *Subject self-concept* as a basis for computing weighted composite scores. The factor at Time 1 includes five observed variables, for which the raw factor score coefficients are .156, .249, .102, .360 and .219, respectively. Composite scores for *Subject self-concept* are calculated by first dividing each factor score by the sum of all five factor coefficients, which standardizes them to a scale of 1 (creating scores of .144, .229, .094, .331, and .202 respectively). Each standardized factor coefficient is then multiplied by the corresponding score in an individual's response set, according to the formula: *Subject self-concept* = (Sub2 * .144) + (Sub3 * .229) + (Sub5 * .094) + (Sub13 * .331) + (Sub15 * .202) (Holmes-Smith, 2012; Holmes-Smith & Rowe, 1994). The factor loading for each composite score is then computed as the square root of the Rho reliability statistic for the corresponding measure, and each error variance is computed as 1 – Rho, for that measure.

Rho reliability is a 'weighted' index, in that it takes into account the proportional contributions on individual scale items, as discussed in further detail in section 5.8.7. Fixing the factor loading and error variance with the M*plus* syntax shown in Appendix A expresses the latent variance of each factor in a simplified format, such that models of greater complexity can be fitted with smaller samples (Holmes-Smith & Rowe, 1994).

Composite scores reduce the number of free parameters that must be estimated in a given model, which allows more complex models to be estimated with smaller samples. They were relied upon in the present study, especially when estimating models whose *N:q* ratios would otherwise have been less than 1. While models will, wherever possible, be estimated with observed indicator variables, composite score model results will also be reported for the sake of comparison. Comparing the results of these two methods of model estimation is intended to help approximate, with respect to data in the present study, the extent to which models estimated exclusively with composite scores, such as the longitudinal structural model, would be different if estimated with observed indicator variables. These differences were summarized in terms of range of differences ($R_{\Delta\beta}$), mean difference ($M_{\Delta|\beta|}$), and absolute mean difference ($M_{|\Delta\beta|}$) in beta coefficient magnitude. Differences in levels of significance were also noted.

5.8 Measures of model fit

The fit of structural equation models is generally indicated by multiple 'fit indices' that, often derived from χ^2 values, approximate particular aspects of model fit such as the size of residual terms (e.g. *RMSEA* and *SRMR*), and the fit of the specified model as compared to that of a null model in which covariances between variables are constrained to zero (e.g. *CFI* and *TLI*). χ^2 tests, by contrast, the *exact fit* of a specified model to a sample of data. χ^2 is generally considered an overly-sensitive test of model fit especially when sample size exceeds

200 (Bollen, 1989; Schumacker & Lomax, 2010). Because large samples are generally desirable in structural modeling analyses, model fit is commonly judged with approximate fit indices.

Fit indices are generally expressed in terms of value scales, such as scales of 0 – 1 (e.g. *CFI* and *SRMR*) or 0 - ∞ (e.g. *RMSEA*), and do not entail tests of statistical significance. The prevalence of fit indices in SEM literature raises the question, therefore, of what values should be taken to indicate acceptable model fit. The need for standards that can discern between well- and poor-fitting models has given rise to so-called 'golden rules' or 'rules of thumb', according to which the acceptability of model fit can be quickly determined. Golden rules have been criticized, however, for encouraging a dogmatic approach to model assessment that too often leads researches and reviewers to ignore issues of model-complexity, sample size, and the specific purposes that various analyses are intended to serve (Hu & Bentler, 1999; Kline, 2011; Markland, 2007; Marsh, Hau, & Wen, 2004; Nye & Drasgow, 2011). Kline (2011) exhorts researchers to "*refrain from blindly applying threshold values* [italics in original]" (p. 199). Marsh et al. (2004) argue likewise that researchers should assess model fit "in relation to the specific details of their research" and beware the problem of false-negative results that "a single cutoff value for each index that generalizes across different sample sizes (*N*) and different situations" has the potential to cause (pp. 321 – 322).

In short, different types of model analysis often require different fit criteria. For the Pilot Study reported in Chapter Four, fit criteria were loosened considerably in comparison to the threshold values recommended by scholars such as Schumacker and Lomax (2010). Relaxing standards served the purpose of identifying fit that was *truly poor* among relatively simple congeneric models, fitted with a small sample (*N* = 96). Fit criteria tend to be less trustworthy when smaller samples are used (Sivo et al., 2006; Chen et al., 2008; Kline, 2011). Attempting to identify fit that is truly poor in a small sample, i.e. that falls far enough outside of desirable limits as to indicate problems with high certainty, demands different criteria than attempting to determine whether the fit of a model is *truly good*, in a large sample, as should probably be stricter.

Two broad types of model are analyzed in the 'Main Study': congeneric models and multivariate models. This distinction is noted because fit indices are often sensitive to model complexity, especially in relation to sample size, as expressed by the *N:q* ratio (Herzog, Boomsma, & Reincke, 2007). It makes little sense to apply the same fit index thresholds to both simple congeneric models and complex multivariate models, in the same sample. Two different sets of fit criteria were, therefore, followed throughout the Main Study (see Table 5.2). These criteria adopt common standards for assessing model fit, while appropriately reflecting large differences in model complexity.

5.8.1 χ^2 test of model fit

The χ^2 statistic indicates whether a model demonstrates exact fit, i.e. equality between the covariance matrix implied by the model ($\mathbf{\Sigma}$) and the sample covariance matrix ($\mathbf{S}$) (Kline, 2011). A χ^2 statistic that is significant at the $p < .05$ level indicates that fit between these two matrices is significantly *in*-exact, whereas a non-significant χ^2 indicates exact fit. The χ^2 statistic is most appropriate as an indicator of fit for relatively simple models of normally distributed data. For models of data with varying degrees of normality, however, χ^2 statistics are unlikely to indicate exact fit because "the underlying distribution is not χ^2 distributed" (Byrne, 2012, p. 69). A non-significant χ^2 statistic was, therefore, a desideratum for the fit of single-factor and multivariate models estimated with normalized, single-indicator composite scores in the present study. Significant χ^2 statistics for these models may indicate substantial model misfit when coinciding with other inadequate fit indices.

The significance of χ^2 statistics was considered, by contrast, an inappropriate desideratum for multivariate models estimated with observed indicator variables. Such models are unlikely to achieve non-significant χ^2 statistics because they entail, cumulatively, large amounts of non-normality, and are estimated with samples that are usually larger than 200 (Byrne, 2012).

Table 5.2

Desiderata for model fit and reliability

		Single-factor Congeneric models	Multivariate models
p of χ^2	$\geq$	.05	N/A
CFI	$\geq$	.95	.90
TLI	$\geq$	.90	.90
RMSEA	$\leq$	.08	.08
Lower bound *CI* of *RMSEA*	$\geq$	.05	.05
Upper bound *CI* of *RMSEA*	$\leq$	.10	.10
pclose	$\geq$	N/A	.50
SRMR	$\leq$	.08	.08
Factor loadings	$\geq$	.300	.300
Reliability (coefficient Rho)	$\geq$	.70	N/A

5.8.2 Comparative Fit Index (*CFI*)

CFI is a χ^2-derived measure of approximate fit that indicates the relative improvement of a specified research model over a statistical baseline model, on a scale of 0 – 1 (Kline, 2011). The widely cited desideratum of *CFI* $\geq$ .95, which traces back to seminal recommendations by Hu and Bentler (1999), has been questioned more recently (Marsh, Hau & Wen, 2004; Fan & Sivo, 2005). The *CFI* $\geq$ .95 threshold appears to be most appropriate for less complex models of data that is more normal. The *CFI* $\geq$ .95 recommendation of Hu and Bentler (1999) was based, for instance, on analyses of measurement models that incorporated just 15 observed indicator variables across three latent factors, with an average of 35 free model parameters. The models Hu and Bentler (1999) referred to as 'complex' differed from models referred to as 'simple', moreover, by the inclusion of three cross-loadings. These multivariate models are not, as such, especially comparable to multivariate models analyzed in the present study, which tended to have high model complexity (approximately 180+ free model parameters) and comparatively small sample sizes. Marsh, Hau, & Wen (2004) argue that a *CFI* value of .90 may often be a more appropriate fit threshold, especially for models of greater complexity estimated with smaller samples. The desideratum of *CFI* $\geq$ .90 was, therefore, adopted for multivariate analyses in the present study, in contrast the abovementioned desideratum of .95 for one-factor congeneric models (see Table 5.2).

5.8.3 Tucker-Lewis Index (*TLI*)

TLI is, like *CFI*, a χ^2-derived measure of approximate fit. *TLI* and *CFI* differ, however, in two important ways. Firstly, *TLI* values are non-normed, meaning that they can exceed 1.0, whereas *CFI* values cannot. Secondly, *TLI* punishes models that lack parsimony, especially as caused by inclusion of "parameters that contribute minimally to the improvement in model fit" (Byrne, 2012, p. 71). A number of models in the present study are, in fact, expected to lack parsimony, including large congeneric models, such as the eight-indicator model for *Good teaching*, and complex multivariate models that are designed to test dozens of hypotheses simultaneously. The minimum value of *TLI* $\geq$ .90, proposed by Schumacker and Lomax (2010), was thus applied in the present analysis to both one-factor and multivariate models.

5.8.4 Root mean square error of approximation (*RMSEA*)

RMSEA is a χ^2-derived approximate fit index for 'badness-of-fit', in that larger RMSEA values reflect higher degrees of difference between the model-implied covariance matrix and the sample covariance matrix (Kline, 2011). The desired value for *RMSEA* point estimates was set at a lower level for analyses of congeneric (.08) and multivariate models (.06) in the Main Study than in the Pilot Study (.10). More complex models with more degrees of freedom and larger sample sizes tend to generate lower *RMSEA* values (Chen, Curran, Bollen, et al., 2008) because the denominator of *RMSEA* is the product of model degrees of freedom and sample size, minus one (*df* (*N* – 1)). It is also necessary, therefore, to maintain different threshold criteria for one-factor congeneric models *versus* multivariate models, in order to reflect differences in their complexity when, at a given time point, sample size is constant. While scholars have recently argued against calculating *RMSEA* for models with few degrees of freedom (Kenny, Kaniskan & McCoach, 2011), *RMSEA* remains a conventional and widely recognized test of fit. A more forgiving desired value of .08 was, therefore, adopted for one-factor models, whereas the stricter desideratum of .06 was applied to multivariate models (see Table 5.2).

***RMSEA* confidence intervals**. As discussed in Chapter Four (see section 4.4.2), confidence intervals (*CI*s) are important indicators of the precision of *RMSEA*; wide confidence intervals tend to indicate low precision and *vice versa*. *RMSEA* confidence intervals are, like *RMSEA* itself, affected by both sample size and model complexity (Byrne, 2012). Differences in complexity between the congeneric *vs*. multi-factor measurement model may bring about corresponding differences in the breadth of *CI*s, resulting in smaller lower-bound *CI* estimates. The cutoff threshold for lower-bound 90% *CI*s may, moreover, be held to a stricter standard, i.e. lower value, than the *RMSEA* statistic.

The upper-bound confidence interval for *RMSEA* can be considered a test of the hypothesis that the model does not fit, or the ‘poor-fit hypothesis‘, whereas the lower-bound confidence interval tests the hypothesis that the model does fit, or the ‘good-fit hypothesis’ (Kline, 2011). 90% upper-bound *CI* values of less than .10 reject the hypothesis of poor fit, whereas 90% lower-bound *CI* values greater than .05 reject the hypothesis of good fit. These values conform to widely accepted conventions in structural equation modeling (Kline, 2011). It is important to note, however, that 90% confidence intervals may simultaneously reject neither of these fit hypotheses. The lower-bound *CI* may, in other words, be lower than .05 while the upper-bound *CI* is also greater than .01, for the same model. Such results suggest significant amounts of sampling error. When accompanied by an acceptable point-estimate of *RMSEA* $\leq$ .08, excessive *RMSEA* confidence intervals do not reject the model.

***RMSEA* probability of close fit**. The 'probability of close fit' (*pclose*), indicates the probability that *RMSEA* $\leq$ .05, and is considered an important indicator of whether an *RMSEA* estimate should be accepted (Kenny, 2014a). The threshold for the *RMSEA* statistic is, as explained above, set to .08 for simple, congeneric models, because *RMSEA* estimates tend to be inflated for models with fewer degrees of freedom. The question of whether *RMSEA* $\leq$ .05 is relevant, therefore, only for the full measurement model. As indicated in Table 5.2, a *pclose* threshold of $\geq$ .5, which indicates a 50% or greater probability that *RMSEA* is less than or equal to .05, will be applied to multivariate models.

5.8.5 Standardized Root mean square residual (*SRMR*)

SRMR provides a summary measure of variance-covariance matrix residuals for a given model, and has been shown to be sensitive to both sample size and skew (Nye & Drasgow, 2011). Analyses conducted by Hu and Bentler (1999) suggest that acceptable model fit is indicated by an *SRMR* value of $\leq$ .08. A follow-up study conducted by Marsh et al. (2004) found, similarly, that mis-specified three-factor models with no cross-loadings (i.e. 'simple' models) were rejected 100% of the time by a cutoff of *SRMR* $\leq$ .08. *SRMR* is, moreover, inversely sensitive to sample size. Marsh et al. (2004) found, for instance, that the *SRMR* value for a correctly-specified model increased from 0.0 to 0.047 when the sample size decreased from N = 250 to N = 150. A threshold of *SRMR* $\leq$ .08 was thus adopted both for one-factor congeneric models and for multivariate models.

5.8.6 Factor loadings

Factor loading values express the contribution made by observed indicator variables to the variance of the latent variables they were designed to measure. The minimum threshold for factor loadings in the present study of .300 is equivalent to a cutoff for variance explained of 9% (Banalos & Finney, 2010). Observed variables that explain less than nine percent of the variance of their target factor are, in essence, extraneous to the measure, and may be a significant source of poor fit.

5.8.7 Measure reliability (Rho)

Acceptable scale reliability is an especially important criterion for estimating multivariate models with weighted composite scores (Holmes-Smith & Rowe 1994; Raykov, 2009). As recommend by Holmes-Smith & Rowe (1994), reliability was estimated with a unit-weighted, or 'maximized', reliability coefficient, Rho (ρ), with syntax for M*plus* devised by Raykov (2004, 2009). Unlike alpha reliability (Cronbach, 1951), which assumes that the items of a given scale are tau-equivalent (Raykov, 1997), or "that the components measure the same underlying latent dimension with the same units of measurement" (Raykov, 2004, p. 301), coefficient Rho accounts for differences in the relative contributions of individual scale components (Raykov, 2009).

Rho reliability coefficients were used, as shown in Appendix A, to directly calculate the factor loadings and error variances of composite scores. These factor loadings and error variances were then fixed, using M*plus* syntax also given in Appendix A, in order to represent true latent variables in multivariate analyses. A threshold of .70 was, therefore, a key desideratum for the reliability in the present study, following the oft-cited guideline that reliability scores of .70 are "adequate" (Kline, 2011, p. 70), in that 30% or less of the variance represented is due to random error. While researchers frequently include constructs with reliability scores as low as .60, it was considered desirable to either modify or drop measures from the hypothesized model that exhibited reliability lower than .70.

5.9 *Post-hoc* modification

Structural equation modeling is a confirmatory statistical methodology, albeit with recognized exploratory applications (Asparouhov & Muthén, 2009; Byrne, 2012; Jöreskog, 1993; Muthén & Muthén, 1998-2010). One type of exploratory application includes *post-hoc* modifications to the measurement and/or structural elements of a model that was otherwise hypothesized *a priori*. Because measures in this study were used on a uniquely diverse population of students at international secondary schools located in 9 different countries, the likelihood of needing such modifications was anticipated.

Post-hoc modifications are often made by researchers in order to improve model fit. They also tend to cast doubt on the validity of model results, especially when they lack a strong theoretical rationale, and when they fail to be cross-validated with multiple data sets. It was often preferable, therefore, to drop factors from the measurement model that demonstrated unacceptably poor congeneric fit, unless modifications could be made from a strong theoretical or methodological basis, and could be cross-validated on multiple data sets.

Post-hoc modifications to the structural elements of hypothesized models also tend to be looked upon with suspicion (Byrne, 2012). Such modifications may, however, recognize empirical realities that were not, and perhaps could not have been, anticipated *a priori*. Such modifications were also generally avoided in the present study, unless they enjoyed a strong theoretical rationale and could be cross-validated in multiple data sets.

CHAPTER 6

CROSS-SECTIONAL ANALYSES OF TIME ONE DATA

Analyses of Time 1 data built upon the results of the Pilot Study, wherein factors demonstrating poor fit or multicollinearity were either modified, dropped from the study, or fit into higher-order factor structures. These changes resulted in a revised hypothesized model, or 'Model 2'. Model 2 was assessed for purposes of the Main Study, beginning with the congeneric fit of individual factors and basic descriptive statistics. The retained set of factors proceeded to multivariate measurement model analyses. A lack of factorial invariance across gender groups that was detected in the measurement model prompted separate structural model analyses of Model 2 for males and females, prior to estimating Model 2 for the sample as a whole.

Low *N:q* ratios were addressed in two ways. Firstly, the small sample size correction developed by Herzog and Boomsma (2009), introduced in section 5.7.2, was applied to multivariate models estimated with all observed indicator variables. Secondly, each multivariate model was re-estimated with weighted composite scores (see section 5.7.3), which improved *N:q* ratios by reducing the number of free model parameters (q). The results of composite score models were compared to their counterparts estimated with observed indicators variables, as a means of cross-validating these two methodologies.

6.1 Psychometric analysis of one-factor congeneric models

Congeneric model fit is an important prerequisite for combining the variance contributed to a latent variable by multiple observed indicator variables used to measure it, into a single weighted composite score (Holmes-Smith & Rowe, 1994; Holmes-Smith, 2012). Significant problems with congeneric fit and reliability identified at Time 1 resulted in measures being either dropped from the study or, where appropriate, modified using pilot data. No *post-hoc* modifications were made based on Time 1 modification indices. This was done to avoid damaging the generalizability of model results. Modifications to psychometric measures that are based on a single data set run a higher risk of being idiosyncratic to that data set (Bandalos & Finney, 2010).

While eight congeneric models in Table 6.1 fell short of meeting all fit desiderata established in Chapter Five (see Table 5.2), two of these models (*Self-reported cheating* and *Surface learning strategies*) missed only with respect to the upper 90% confidence interval of *RMSEA* ($CI_{upper} \leq .10$). This statistic for *Self-reported cheating* (.166) violated the established threshold of .10, thus supporting the hypothesis that the model would not fit, or the 'poor-fit hypothesis' (Kline, 2011). This result was inconclusive, however, because the lower 90% confidence interval for *RMSEA* ($CI_{lower} = .013$) was acceptably small, which simultaneously supported the 'good-fit hypothesis'. The point-estimate for *RMSEA* fell, moreover, precisely on the threshold for acceptable fit (.080). A similar pattern for *RMSEA* confidence intervals and point estimates was observed with respect to *Surface learning strategies* (*RMSEA* = .080, $CI_{lower} =.030$; $CI_{upper} = .140$), where confidence intervals simultaneously supported good- and poor-fit hypotheses, and the point-estimate fell precisely on the established threshold. The support provided by the *RMSEA* confidence intervals for contradictory fit hypotheses, together with the fact that both *RMSEA* point-estimates met minimum criteria, suggested that these two constructs were borderline, but adequate.

Table 6.1

Congeneric model results: Time 1 (N = 493)

	Congeneric CFA										
					RMSEA						
Scale (# items)	χ^2	*p*	*df*	Loading range	Value	Low *CI*	High *CI*	*CFI*	*TLI*	*SRMR*	Rho
Subject self-concept (5)	2.51	.775	5	.715-.845	.000	.000	.042	1.00	1.01	.006	.91
Honesty-trust. self-concept (6)	10.55	.308	9	.414-.885	.019	.000	.056	1.00	1.00	.018	.82
Performance structure (4)	2.97	.226	2	.547-.844	.031	.000	.100	1.00	.99	.015	.74
Appropriate workload (3)	.166	.684	1	.417-827	.000	.000	.089	1.00	1.02	.010	.62
Good teaching (8)	54.21	.000	20	.410-.756	.059	.040	.078	.97	.95	.032	.86
Usefulness of curriculum (4)	.069	.966	2	.702-.905	.000	.000	.000	1.00	1.01	.001	.89
Transparency of assessment (6)	72.46	.000	9	.592-.682	.120	.095	.146	.88	.81	.057	.81
Authenticity of assessment (7)	65.71	.000	14	.464-.740	.087	.066	.108	.93	.89	.048	.81
Peer norms (5)	6.25	.282	5	.534-.713	.023	.000	.070	1.00	.99	.015	.78
Experience of classroom rules (5)	34.88	.000	5	.450-.696	.110	.077	.146	.92	.83	.044	.75
Surface learning strategies (4)	8.35	.015	2	.324-.878	.080	.030	.140	.98	.93	.025	.71
Deep learning strategies (6)	38.80	.000	9	.601-.735	.082	.057	.109	.95	.91	.038	.82
Justifiability of cheating (3)	.065	.798	1	.552-.811	.000	.000	.076	1.00	1.02	.003	.73
Self-reported cheating (3)	4.15	.042	1	.763-.895	.080	.013	.166	.99	.96	.060	.87

Note. χ^2 = chi-squared; *p* = significance level; *df* = degrees of freedom; *CI* = confidence interval; Rho = Rho reliability coefficient; highlights = index threshold violations.

χ^2 values for both *Self-reported cheating* and *Surface learning strategies*, as well as for *Good teaching*, were significant at the $p < .05$ level, which indicated a lack of exact fit. χ^2 is not, on its own, an appropriate basis for rejecting models, however, especially those with non-normal distributions and where sample sizes exceed $N = 200$ (Bollen, 1989; Schumacker & Lomax, 2010; Byrne, 2012). Significant χ^2 statistics and excessive *RMSEA* confidence intervals were approached in the present study as potential indicators of poor fit that may lead to model rejection only in concert with other inadequate fit indices.

Self-reported cheating, *Surface learning strategies*, and *Good teaching* will, for the reasons given above, not be modified or dropped from the study. The remaining five congeneric models in Table 6.1, for which fit was more problematic (*Appropriate workload, Experience of classroom rules*, *Deep learning strategies*, *Transparency of assessment*, and *Authenticity of assessment*), are discussed below.

Appropriate workload. This factor was retained on a tentative basis for analysis in the Main Study, in light of multiple problems with its congeneric model identified during the Pilot Study. These problems necessitated removing two of the measure's original five items due to low factor loadings, and fixing the residual variance of item Apwkld52 to .00001 in order to address a Heywood case. Factor structure was additionally imbalanced by a factor loading of λ = 1.00 for item Apwkld52 in comparison to loadings of λ = .405 and λ = .460 for items Apwkld30 and Apwkld35, and scale reliability was low (.67).

A similarly imbalanced pattern of factor loadings emerged for the congeneric model for *Appropriate workload* at Time 1 of the Main Study (λ = .834, λ = .421, λ = .491, respectively), in addition to a scale reliability estimate (.62) that violated the desired threshold of .70. While the fit of the model to the data was excellent ($\chi^2(1)$ = .136, p = .684; *RMSEA* = .000, CI_{lower} = .000, *SRMR* = .089; *CFI* = 1.00; *TLI* = 1.02; *SRMR* = .010), such low reliability is problematic because factor loadings and error variances for composite scores, used extensively throughout this study, are calculated directly from reliability estimates. In view of these problems, the measure *Appropriate workload* was dropped from the hypothesized model.

Experience of classroom rules. The congeneric model for *Experience of classroom rules* failed the *RMSEA* test of close-fit, with respect to the point-estimate and both 90% confidence intervals (*RMSEA* = .110, CIs = .077 - .146). These *RMSEA* estimates were significantly worse than in the Pilot Study (*RMSEA* = .096, CIs = .000 - .156). The model also had a large and highly significant chi-squared statistic ($\chi^2(5)$ = 34.88, p = .000) and failed to achieve acceptable levels for *CFI* (.92) or *TLI* (.83). *Experience of classroom rules* was dropped from the hypothesized structural model due to these observations.

Deep learning strategies. The congeneric model for *Deep learning strategies* failed the *RMSEA* test of close-fit in all respects (*RMSEA* = .082, CIs = .057 - .109), and had a large and highly significant chi-squared statistic ($\chi^2(9)$ = 38.80, p = .000). The same measure achieved good fit in the Pilot Study only after being modified by the removal of item Deep80, which appears to have been an example of a non-replicable *post-hoc* modification. The poor fit of this model at Time 1 was unlikely to be remedied in a generalizable manner through additional *post-hoc* modifications. *Deep learning strategies* was, therefore, dropped from the hypothesized structural model.

Assessment factors. Congeneric models for both the *Transparency* and *Authenticity of assessment* failed the *RMSEA* test of close fit and fell short of cutoff thresholds for *CFI* and *TLI*. These findings, in addition to large and highly-significant chi-squared statistics for both factors, made it impossible to retain them in the study without modification.

Assessment, the principal venue for academic cheating, is theorized to play an important role in how learning contexts affect cheating. Of the three measures of assessment experience included in the original questionnaire, *Appropriate Assessment* (Wilson et al., 1997; Marsh, Ginns, Morin et al., 2011), was dropped during the Pilot Study due to psychometric dysfunction. *Transparency* and *Authenticity of assessment*, taken from the *Perceptions of Assessment Task Inventory* (Dorman & Knightley, 2006), performed well in the Pilot Study, but not at Time 1. Both measures had large and highly significant χ^2 statistics ($\chi^2(9)$ = 72.46, p=.000

and $\chi^2(14) = 65.71$, p=.000, respectively), and poor approximate fit statistics (*Transparency*: *RMSEA* = .120, *CI*s = .095 - .146; *CFI* = .88; *TLI* = .81. *Authenticity*: *RMSEA* = .087, *CI*s = .066 - .108; *CFI* = .93; *TLI* = .89).

Because these measures derive from the same secondary-level instrument, and pertain, by design, to dimensions of the same broad phenomenon, assessment experience, the possibility of using Pilot Study data to create a hybrid of the two measures was explored. Deriving a hybrid measure using pilot data, and cross-validating it with Time 1 data, ran less risk of capitalizing on the idiosyncrasies of either respective data set.

6.1.1 *Assessment quality*: Integration and cross-validation of measures

Measurement of students' perceptions of assessment was a highly desirable, if not indispensable, dimension of the second-order factor '*Teacher quality*'. Assessment is seen to both guide teaching and learning processes (Black & Wiliam, 1998; Taras, 2010), and to affect whether students cheat (Berliner, 2011; Galloway, 2012; Sisti, 2007). *Transparency* and *Authenticity of assessment* could not be retained in the hypothesized model in their original form due to the level of psychometric dysfunction observed with respect to their congeneric models (see Table 6.1). A new assessment measure that combined items from both scales was, therefore, created using pilot data, and cross-validated with Time 1 data. *Transparency* and *Authenticity of assessment* were multicollinear in the Pilot Study (r = .785) and at Time 1 (r = .901), suggesting that they already captured much of the same assessment-related variance. A measure composed of items from these measures was likely to retain a very similar operational definition.

Creating a hybrid measure for *Assessment quality*. While factor structure can be explored either by exploratory or confirmatory factor analytical techniques, confirmatory analyses are more appropriate in cases where considerable research has already been conducted on a factor's structure (Bandalos & Finney, 2010). Because the measures of *Authenticity* and *Transparency of assessment* have received considerable prior attention as measures on the *Perceptions of Assessment Tasks Inventory* (PATI) developed by Dorman and Knightley (2006; see also Alkharusi, Aldahafri, Alnabhani, & Alkalbani, 2013), both exploratory and confirmatory methods were used to identify possible combinations of items from the two assessment scales that could serve as a hybrid measure.

EFA was conducted on all thirteen (13) items of the two scales, combined, using Pearson correlations in the program *Factor version 7.0* (Lorenzo-Sevo & Ferrando, 2007) to see what coherent structures would emerge. A factor was identified with a large Eigenvalue (5.509) that comprised six items with loadings greater than .400: Trans32, Trans45, Trans63, Tans66, Auth44, and Auth78. This factor rendered marginal fit in CFA, using pilot data ($\chi^2(9) = 44.896$, p=.000; *CFI* = .94; *RMSEA* = .090, CI_{lower} = .000; *SRMR* = .040).

Running all thirteen items in CFA using M*plus version 7.11* (Muthén & Muthén, 2013) also produced a borderline factor structure ($\chi^2(65) = 104.66$; *CFI* = .90; *RMSEA* = .080, CI_{lower} = .050; *SRMR* = .069). Eliminating the three lowest-loading items of each constituent scale rendered a seven-item structure that included three *Transparency of Assessment* items (Trans28, Trans32, Trans66), and four *Authenticity of Assessment* items (Auth44, Auth60, Auth71, Auth78). The fit of this seven-item measure was the best of the three integrated scales tested with pilot data ($\chi^2(14) = 23.025$, p=.060; *RMSEA*=.082, CI_{lower}=.000, *SRMR*=.048; *CFI*=.95), keeping in mind that $\leq$ .01 was the threshold for *RMSEA* in the Pilot Study. Scale reliability was also good (Rho = .88). The content validity of this seven-item measure, dubbed *Assessment quality*, was next carefully assessed.

Theoretical considerations. Items on the integrated factor appeared to reflect, overall, the degree of congruence, or discrepancy, between what students believe they should be learning in a given class, and what they see emphasized on assessment tasks (see Table 6.2). Assessment tasks that miss the important substantive purposes of learning may appear both non-transparent and inauthentic.

Table 6.2

Integrated factor: Assessment quality, Pilot Study data (N = 96)

	CFA			
	Factor loadings (λ)			
Item	**Est.**	**S.E.**	***p***	**Item wording**
Auth78	.692	.076	.000	In my Science class, graded assignments examine my ability to answer important questions.
Auth71	.643	.077	.000	In my Science class, graded assignments test my ability to use what I've learned.
Auth60	.719	.096	.000	In my Science class, graded assignments are useful.
Auth44	.798	.043	.000	In my Science class, graded assignments check my understanding of topics.
Trans28	.676	.071	.000	I understand the purpose of graded assignments in my Science class.
Trans32	.740	.060	.000	I understand what is needed in all graded assignments in my Science class.
Trans66	.776	.058	.000	I know in advance HOW I will be graded in my Science class.

The original scale for *Transparency of assessment*, provided in Table 6.3, below, appears to reflect four aspects of graded assignments, including two substantive aspects: (1) understanding the purpose of assignments (Trans28), (2) understanding what is needed to succeed (Trans22, Trans32, and Trans45); and two administrative aspects: (3) understanding how assignments are graded (Trans66), and (4) knowing when graded assignments will be given (Trans63). The first two of these aspects are substantive, in that they pertain to a student's understanding of the nature of assignments, including what is required to succeed at them. The third and fourth aspects are administrative, in that they pertain to grading policy and scheduling.

The three items of *Transparency of assessment* included in the integrated factor (see bold items in Table 6.3, below) reflect three dimensions of the original scale, identified above: (1) the purposes of graded assignments (Trans28), (2) what is needed to succeed (Trans32), and (3) how assignments will be graded (Trans66).

Table 6.3

Transparency of assessment, Pilot Study data (N = 96)

	CFA			
	Factor loadings (λ)			
Item	**Est.**	**S.E.**	***p***	**Item wording**
Trans22	.460	.112	.000	I know what is needed to successfully accomplish graded assignments in my Science class.
Trans28	**.581**	**.059**	**.000**	**I understand the purpose of graded assignments in my Science class.**
Trans32	**.831**	**.053**	**.000**	**I understand what is needed in all graded assignments in my Science class.**
Trans45	.678	.080	.000	I am told in advance WHAT science topics and information I will be graded on in my Science class.
Trans63	.659	.092	.000	I am told in advance WHEN I will be graded in my Science class.
Trans66	**.759**	**.060**	**.000**	**I know in advance HOW I will be graded in my Science class.**

Note: Bold items are included in the integrated factor.

Whether students understand the purpose of a graded assignment (Trans28) is inherently related to the deeper issue of whether they understand what is needed, or what it requires of them (Trans32), in that learning and ability are often assessed in terms of a student's ability to meet requirements such as demonstrating, remembering, or solving. The question of how, or according to what criteria, assessments will be graded (Trans66) appears to extend student concern over what academic requirements they face in assessment situations to the administrative issue of how success will be judged. Discrepancies between what students think they should be learning in a given class, and what they see the teacher emphasizing on assessment tasks, may cause students to feel that the purposes of the class are less transparent. This might be summarized as whether the 'nature of success' is transparent.

It is not surprising that, when combined with Dorman and Knightley's (2006) measure for *Authenticity of assessment*, which emphasizes the relevance of assessment tasks to learners' lives and interests, items from the *Transparency* scale pertaining to substance and purpose would be elevated. It is also unsurprising that the fourth aspect of transparency identified above with respect to the original scale, which pertains to the relatively non-substantive issue of scheduling (Trans63), fell out of this factor structure.

Four of the seven items in the measure *Assessment quality* come from *Authenticity of assessment* (bold items in Table 6.4, above). Items of the original *Authenticity of assessment* scale measure two broad aspects of the value of graded assignments to learners: (1) the real-world value of the material they cover (Auth29, Auth37, Auth60), and (2) their effectiveness at assessing student achievement (Auth44, Auth60, Auth71, Auth78). Firstly, the value of the material covered by assignments is conceptualized in terms of its (A) meaningfulness (Auth29), (B) pertinence to the real world (Auth37 and Auth48), and (C) usefulness (Auth60). Secondly, the effectiveness of assignments at assessing student achievement is expressed in terms of (D) 'checking my understanding of topics' (Auth44 and Auth71), and (E) 'testing my ability to use what I've learned' (Auth60). A final item, Auth78, queries whether assignments check students' ability to answer 'important questions', which appears to split the difference between the value of the material covered by assignments and how well assignments measure real achievement (aspects 1 and 2, above).

Table 6.4

Authenticity of assessment, Pilot Study data (N = 96)

	CFA			
	Factor loadings			
Item	**Est.**	**S.E.**	***p***	**Item wording**
auth29	.629	.083	.000	In my Science class, graded assignments are meaningful.
auth37	.489	.094	.000	I find that in my Science class, graded assignments relate to the real world.
auth44	**.573**	**.099**	**.000**	**In my Science class, graded assignments check my understanding of topics.**
auth48	.484	.106	.000	I am asked to apply my learning to real world situations in my Science class.
auth60	**.678**	**.082**	**.000**	**In my Science class, graded assignments are useful.**
auth71	**.779**	**.117**	**.000**	**In my Science class, graded assignments test my ability to use what I have learned.**
auth78	**.782**	**.046**	**.000**	**In my Science class, graded assignments examine my ability to answer important questions**

Note: Bold items are included in the integrated factor.

The integrated, seven-item factor for *Assessment quality* in Table 6.2 emphasizes two key aspects of assessment. The first of these aspects, the 'usefulness' conception of an assessment's value (Auth60), is emphasized in terms of assessing student mastery (Auth44), student ability to use what is learned (Auth61), and ability to answer questions perceived by students as important (Auth66). The second aspect of assessment emphasized by the emergent factor structure relates to whether students understand the purpose of an assessment (Trans28), know what is needed to succeed at assessments (Trans32) and know how they will be graded (Trans66).

Conclusion. The integrated seven-item scale for *Assessment quality* was adopted tentatively, pending successful cross-validation with Time 1 data. The integrated scale appears to measure a broader conception of the quality of assessment than either of the scales from which it was derived. The integrated factor derived using Pilot Study data, was next cross-validated at Time 1.

Table 6.5

Cross-validation of the integrated factor for Assessment quality with Time 1 data (N = 493)

	CFA			
	Factor loadings (λ)			
Item	**Est.**	**S.E.**	***p***	**Item wording**
Auth78	.673	.036	.000	In my Science class, graded assignments examine my ability to answer important questions.
Auth71	.634	.041	.000	In my Science class, graded assignments test my ability to use what I've learned.
Auth60	.734	.030	.000	In my Science class, graded assignments are useful.
Auth44	.719	.032	.000	In my Science class, graded assignments check my understanding of topics.
Trans28	.637	.036	.000	I understand the purpose of graded assignments in my Science class.
Trans32	.612	.043	.000	I understand what is needed in all graded assignments in my Science class.
Trans66	.620	.043	.000	I know in advance HOW I will be graded in my Science class.

Cross-validation of Assessment quality at Time 1. The fit of *Assessment quality* to Time 1 data (N = 493) was acceptable. While chi-squared was significant ($\chi^2(14) = 26.639$, p=.021), tests of close and approximate fit were at appropriate levels (*RMSEA* = .043, CI_{lower} = .016, *CFI* = .98; *TLI* = .97; *SRMR* = .029), and reliability was good (Rho = .84) (see Table 6.5). *Assessment quality* was used, as such, to measure student experiences of assessment in the Main Study in place of measures for *Transparency of assessment* and *Authenticity of assessment*.

6.2 Basic descriptive statistics

Table 6.6 reports congeneric model fit and basic descriptive statistics for all latent factors retained in the Main Study. Factor means, standard deviations, skewness and kurtosis reported in the right-most columns of the table were derived in SPSS *version 21* from weighted composite scores. Composite scores were calculated from the factor score coefficients associated with each congeneric model, using the methodology of Holmes-Smith and Rowe (1994). The acronyms, definitions and valences of all factors are provided in Appendix E.

All ten first-order factors retained in the hypothesized model were well within the recommended criteria for normality, of 7.0 for kurtosis and 2.0 for skewness (Curran, West & Finch, 1997). Appropriate levels of normality were also observed at the item level, as can be seen in the comparison of the item-level descriptive statistics at Time 1 with those of the Pilot Study (see Appendix F). The congeneric models reported in Table 6.6 also demonstrate acceptable scale reliability, and satisfactory fit to Time 1 data.

Table 6.6

Round One congeneric model results, Time 1 (N = 493)

	Congeneric CFA																
						RMSEA											
Scale (# items)	χ^2	*p*	*df*	Loading range	Mean Loading	value	Low *CI*	High *CI*	*CFI*	*TLI*	*SRMR*	Rho	Mean	*SD*	Skew	Kurt-osis
Subject self-concept (5)	2.51	.775	5	.715-.845	.812	.000	.000	.042	1.00	1.01	.006	.91	2.48	.929	.364	-.420
Honesty-trust. self-concept (6)	10.6	.308	9	.414-.885	.651	.019	.000	.056	1.00	1.00	.018	.82	2.01	.658	.840	1.25
Performance structure (4)	2.97	.226	2	.547-.844	.638	.031	.000	.100	1.00	1.00	.015	.74	3.33	1.00	-.148	-.733
Good teaching (8)	54.2	.000	20	.410-.756	.654	.059	.040	.078	.97	1.00	.032	.86	2.47	.802	.769	.636
Usefulness of curriculum (4)	.069	.966	2	.702-.905	.820	.000	.000	.000	1.00	1.01	.001	.89	2.50	.970	.531	-.171
Assessment quality (7)	26.6	.021	14	.612-.734	.661	.043	.016	.067	.98	.97	.982	.85	2.19	.658	.506	.573
Peer norms (5)	6.25	.282	5	.534-.713	.645	.023	.000	.070	1.00	.99	.015	.78	3.56	.888	-.405	-.376
Surface learning strategies (4)	8.35	.015	2	.324-.878	.611	.080	.030	.140	.98	.93	.025	.71	3.58	.968	.287	.719
Justifiability of cheating (3)	.065	.798	1	.552-.811	.693	.000	.000	.076	1.00	1.02	.003	.73	4.05	.913	.875	.226
Self-reported cheating (3)	4.15	.042	1	.763-.895	.826	.080	.013	.166	.99	.96	.060	.87	4.30	.926	-1.26	.715

6.3 Higher-order confirmatory factor analysis (CFA)

CFA was conducted on the set of constructs presented in Table 6.6, which constituted the central measurement model. Ten of the constructs in Table 6.6 are first-order factors, whereas one, *Teacher quality*, is a second-order factor. CFA of measurement models that include higher-order factors is also referred to as 'higher-order factor analysis' (HCFA), which signals the presence of beta paths between the higher order and first order constructs (Gray, 1997),

While the measurement model achieved adequate fit ($\chi^2(1089) = 1850$; *RMSEA* = .038, *CIs* = .035 - .041, *pclose*= 1.00; *TLI* = .91; *CFI* = .92; *SRMR* = .054; *N:q* = 2.8), complex structural equation models like this one, with 49 observed variables and 178 free model parameters, tend to be penalized when tested against smaller samples (Herzog, Boomsma, & Reinecke, 2007). The size of the present sample *N* = 493, affords an *N:q* ratio of just 2.8 subjects per free parameter. Using the *Swain-R* small sample correction function (Boomsma & Herzog, 2009, 2013) a more accurate assessment of fit was reached ($\chi^2(1089) = 1783$; *RMSEA* = .036, *CIs* = .033 - .039, *pclose* = 1.00; *TLI* = .92; *CFI* = .93; *SRMR* = .054; *N:q* = 2.8; Swain correction factor (*SCF*) = .964).

All observed indicator variables appeared to be well-explained within the Time 1 measurement model by the latent factors they were designed to measure. The correlations between these latent factors accounted adequately for any inter-relatedness between the 49 observed indicator variables in the model, such that no cross-loadings or inter-factor residual co-variances were necessary to achieve good fit. Strong within-model construct validity was demonstrated both by the higher-order factor, *Teacher quality*, which explains 67% of variance in *Good teaching* and 84% of the variance in *Assessment quality*, and by first-order factors in the model, with mean factor loadings within a range of .611-.867 (see Table 6.7).

Table 6.7

Factor loadings of the measurement model HCFA, Time 1 (N = 493)

Main Study, Time 1 (*N* = 493)	Scale (# items)	Loading range	Mean loading
Person	Subject self-concept (5)	.714 - .847	.812
	Honesty-trust. self-concept (6)	.417 - .871	.653
Context	Performance structure (4)	.565 - .791	.642
	Usefulness of curriculum (4)	.720 - .896	.823
	TEACHER	**.819 - .918**	**.867**
	Good teaching (8)	.401 - .759	.653
	Assessment quality (7)	.619 - .727	.662
	Peer norms (5)	.526 - .726	.642
Moral obligation	Justifiability of cheating (3)	.526 - .798	.687
Behavior	Surface learning strategies (4)	.354 - .808	.621
	Self-reported cheating (3)	.762 - .877	.826

Note. Measurement model fit: $\chi^2(1089) = 1783$; *RMSEA* = .036, *CIs* = .033 - .039, *pclose* = 1.00; *TLI* = .92; *CFI* = .93; *SRMR* = .054; Swain correction factor (*SCF*) = .96.

6.3.1 Correlation analysis

All correlations between factors in the multivariate measurement model (see Table 6.8) conform to the direction (sign) anticipated by the hypothetical model. The magnitudes of correlation coefficients in Table 6.8 are also consistent with those calculated in the Pilot Study (see Table 4.17). Time 1 correlation coefficients were larger, on average, than those in the Pilot Study by $M_{\Delta|r|}$ = .027, with a mean absolute difference of $M_{|\Delta r|}$ = .120, across a range of $R_{|\Delta r|}$ = .014-.352.

A key similarity between correlation matrices at Time 1 and the Pilot Study was the presence of correlations exceeding .700 between *Good teaching* and measures of assessment, as well as between *Justifiability of cheating* and *Self-reported cheating*. The strong association between *Good teaching* and *Assessment quality* in Table 6.6 (r = .752) was close in magnitude to correlations in the Pilot Study between *Good teaching* and both *Transparency of assessment* (r = .785) and *Authenticity of assessment* (r = .765). The factors *Good teaching*, *Transparency of assessment*, *Authenticity of assessment*, and *Experience of school rules* were found, in the Pilot Study, to fit within a second-order factor structure for *Teacher quality*. The breadth of *Teacher quality* was reduced to a 2-factor composition at Time 1, by the loss of *Experience of classroom rules* and the integration of *Authenticity* and *Transparency of assessment*.

Higher-order factor analysis. Retaining only two first-order factors in the structure of *Teacher quality* posed a challenge to assessing its second-order fit, in that factor models cannot be identified with only two constituent variables. A factor model is 'not identified' when it does not include "enough constraints on the model and data to obtain" its unique parameter estimates, such as chi-squared statistics (Schumacker & Lomax, 2010, p. 57). The target coefficient (Marsh & Hocevar, 2003) that was used to test the structural validity of *Teacher quality* in the Pilot Study (see Section 4.6.1) could not, therefore, be used because the chi-squared values needed to calculate it could not be generated. Two alternative means of assessing the fit of *Teacher quality* were employed instead. Firstly, the multivariate measurement model was tested both with, and without, imposing the second-order structure on *Good teaching* and *Assessment quality*. This comparison found no difference (Δ = .000) in *CFI*, *TLI*, or *RMSEA*.

Secondly, *Teacher quality* was regressed on 'maternal educational attainment', a covariate with small, non-significant correlations with *Assessment quality* (β = -.042) and *Good teaching* (β = -.010), and that explained negligible variance in *Teacher quality* (β = .052; R^2 =

.1%). Regressing *Teacher quality* on maternal educational attainment used the latter in a manner similar to an 'instrumental variable', which may, in multiple regression, help "identifying models that cannot be estimated" (Kenny, 2014b). This model of *Teacher quality* met desirable criteria for the fit of multivariate analyses ($\chi^2(102) = 211$; *RMSEA* = .047, *CI*s = .038 - .055, *pclose* = .728; *TLI* = .94; *CFI* = .95; *SRMR* = .041; *N:q* = 10).

The correlation between *Justifiability of cheating* and *Self-reported cheating* (r = .766) is larger in Table 6.8 than it was in the Pilot Study (r = .709). A second-order factor structure for these two variables does not, however, appear to be appropriate. Judgment, as measured by *Justifiability of cheating* is categorically distinct from behavior such as cheating. It is unlikely that these factors represent the same underlying factor, but rather that the likelihood of cheating behavior is heavily influenced by whether cheating is viewed as justifiable. *Justifiability of cheating* is better modeled, therefore, as a predictor of *Self-reported cheating* than as the reflector variable of a second-order structure held to explain both. The large amount of statistical overlap between these two variables appears to support the key overarching hypothesis that students are more likely to cheat when they feel less moral obligation to be honest.

Table 6.8

Higher-order CFA correlation matrix, Time 1 (N = 493)

	SUB	HON	PERF	GTEACH	CURUSE	ASSESS	PEER	SURF	CHJUST	CHEAT
SUB	1									
HON	.173**	1								
PERF	-.022	-.077	1							
GTEACH	.389***	.177*	-.093	1						
CURUSE	.492***	.138**	.006	.583***	1					
ASSESS	.437***	.199***	-.104	.752***	.653***	1				
PEER	-.137*	-.218***	.202**	-.332***	-.253***	-.372***	1			
SURF	-.308***	-.275***	.328***	-.289***	-.299***	-.325***	.377***	1		
CHJUST	-.151*	-.379***	.366***	-.392***	-.328***	-.439***	.600***	.574***	1	
CHEAT	-.295***	-.466***	.253***	-.274***	-.235***	-.307***	.492***	.524***	.766***	1
TEACHER	.475***	.217***	-.114	.819***	.712***	.918***	-.405***	-.353***	-.478***	-.334***

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; *p < .05, **p < .01, ***p < .001.

6.4 Invariance of the central measurement model

The body of research on academic integrity reviewed for the purpose of this study suggests that cheating may differ among two sub-groups within secondary school populations: gender (e.g. Finn & Frone, 2004; Galloway, 2012), and grade-level (e.g. Anderman & Midgley, 2004; Galloway, 2012). After establishing that the measurement model fit the data of each group in question, three invariance models were tested (Meade et al., 2008): (1) the *configural model*, for which factor structure was held invariant, (2) the *metric model*, for which factor loadings were additionally held invariant, and (3) the *scalar model*, for which observed variable intercepts were additionally held invariant. Results of these models are presented in Table 6.9. The equivalence of factor variance, an important criterion for representing factors with composite scores, was also examined (see 'Inv. Model 4' and 'Inv. Model 2b' in Table 6.9).

Small sample correction. The degree of factorial invariance was measured in terms of the change in *CFI vis-à-vis* the configural model, when each new constraint was added (Meade et al., 2008; Kline, 2011). Model fit was, however, negatively affected by low *N:q* ratios in the present analysis (Herzog, Boomsma, & Reinecke, 2007). It is crucial to note, moreover, that the *Swain R small sample size correction* devised by Herzog and Boomsma (2009) does not work for multi-group models (A. Boomsma, personal communication, 25 March 2014). Unbiased estimates of model fit indices could not be calculated for invariance models for this reason.

An invariance test model, often referred to in the singular, actually entails testing models for multiple groups simultaneously. This increases the complexity of the analysis while also effectively employing smaller (group) samples. Sub-groups under consideration at Time 1 were small enough (201 males, 292 females, 216 Grade Eight students, and 277 Grade Nine students) to seriously bias *CFI* and chi-squared for a complex model. These sample sizes approached the lower limit (~200) of what should be used to test complex models (Barrett, 2007). Low *N:q* ratios are, as demonstrated by the application of small sample corrections to multivariate analyses throughout the present work, likely to cause considerable bias to *CFI* estimates (see also Sivo et al., 2006). When the measurement model was tested on the sample of male data (N = 201), for example (see section 6.7), the corrected *CFI* estimate was .93, as compared to the uncorrected estimate of .89, a difference of ΔCFI = .04. The correction of bias in *CFI* estimates would likely be ΔCFI = .04 or greater for invariance models, as well, due to the larger number of free model parameters (q), and resulting lower *N:q* ratios. Following arguments that fit index thresholds should be appropriate to "the specific details" of a given analysis (Marsh et al., 2004, p. 321; see also Hu and Bentler, 1999; Marsh, Hau, & Wen, 2004; Markland, 2007; Kline, 2011; Nye & Drasgow, 2011), the *CFI* threshold for multivariate models, given as .90 in Chapter Five (see Table 5.2), was lowered for invariance models to CFI $\geq$ .86, a difference of ΔCFI = -.04. A CFI value of .86 for an invariance model would, with group sample sizes close to 200 and as many as 320 free model parameters, almost certainly have exceeded .90 if a small sample correction had been possible. Lowering the *CFI* threshold for invariance models does not, moreover, change the fact that non-invariance is judged as a decline in *CFI* more than .01.

*Results***.** The fit of baseline models for grade-level and gender, reported in Table 6.9, and that of invariance test models, for which no small sample correction could be performed, was acceptable. An overall decline of ΔCFI = -.006 in the fit of grade-level models indicated an acceptable level of measurement invariance across these groups. Gender non-invariance with respect to the y-intercepts of observed variables was, however, indicated by a decrement in model fit of ΔCFI = -.014 between models 1 and 3, which exceeded the threshold of ΔCFI < -.01. This meant that the y-intercepts of observed variables, or the y-values when $x = 0$ for their respective linear equations, differed between gender groups (Byrne, 2012; Cheung & Rensvold, 2002). Since dispersions of factor data, i.e. factor variances, were equivalent between genders, differences in the *y*-intercepts of observed variables, appeared to reflect underlying differences in mean values, as would cause observed variable vectors to shift up or down.

Table 6.9

Invariance of the measurement model across grade-level and gender, Time 1

	Grade-level invariance						Gender invariance				
	χ^2	*df*	*RMSEA*	*CFI*	*SRMR*		χ^2	*df*	*RMSEA*	*CFI*	*SRMR*
Baseline models											
Grade Eight	1602	1089	.041	.90	.063	Female	1560	1089	.039	.92	.061
Grade Nine	1416	1089	.037	.93	.066	Male	1385	1089	.037	.93	.068
Inv. model 1						**Inv. model 1**					
FS	3260	2177	.045	.890	.064	FS	3181	2177	.043	.896	.064
Inv. model 2						**Inv. model 2**					
Model 1 + FL	3308	2217	.045	.890	.066	Model 1 + FL	3231	2217	.043	.895	.066
Inv. model 3						**Inv. model 3**					
Model 2 + VI	3410	2266	.045	.884	.068	Model 2 + VI	3405	2266	.045	.882	.071
Inv. model 4						**Inv. model 4**					
Model 3 + FV	3425	2277	.045	.884	.073	Model 3 + FV	3435	2277	.045	.880	.083
Inv. Model 2b						**Inv. Model 2b**					
Model 2 + FV	3324	2228	.045	.889	.071	Model 2 + FV	3259	2228	.043	.894	.079

Note. FS = Factor structure (configural invariance), FL = Factor loadings (metric invariance), VI = observed variable intercepts (scalar invariance), FV = Factor variances.

Constraining latent factor variance, or the dispersion of the data encompassed by a latent factor (Field, 2009), across genders in Model 4, increased the total difference in *CFI* to ΔCFI = -.016, reflecting a difference of just ΔCFI = -.002 over Model 3. Removing the constraint on observed variable intercepts, while retaining the constraint on factor variances, resulted in Model 2b. This final model was labeled '2b' because factor variances were constrained in addition to the constraints imposed in Model 2 (i.e. factor configuration and factor loadings). Model 2b resulted in a decrement of ΔCFI = -.001 over Model 2, and ΔCFI = -.002 over the configural model (Model 1). This strongly endorsed the equivalence of factor variance and loadings across genders. Equivalent factor loadings and variances indicate that item weightings are consistent across groups within a data set, such that factors employed in covariance modeling may be appropriately represented as weighted composites.

The scalar non-invariance detected across gender groups indicates differences in the *y*-intercepts of observed variables. Under the condition of equivalent factor variances, differences in the *y*-intercepts of observed variables are likely to reflect underlying differences in those variables' mean values, where higher means should shift variable vectors upward, and *vice versa*. Group differences are explored next at the level of observed variables, or 'items'.

6.4.1 Differential item functioning analysis.

Differential item functioning (DIF) analysis (Grayson et al., 2000; Wang & Wang, 2012) was used to investigate the source(s) of scalar non-invariance in the multivariate measurement model, at the item level. DIF analysis was conducted by regressing all indicators on grouping variables for gender and grade-level (1 = male, 2 = female; 1 = Grade Eight, 2 = Grade Nine). While the measurement model has already been shown to be acceptably invariant across grade-level, this grouping was included for purposes of comparison.

The fit of the DIF model was good ($\chi^2(1089) = 1773$; *RMSEA* = .036, *CI*s = .033 - .039, *pclose* = 1.00; *TLI* = .92; *CFI* = .93; *SRMR* = .052; *N:q* = 1.7; *SCF* = .961). Significant grade-level differences, reported as beta coefficients, were identified in 11 of the 49 observed indicator items of the measurement model (see Table 6.10). Nine were highly significant ($p < .01$). The mean absolute magnitude of these differences was $M_{|\beta|} = .136$, with an absolute value range of $R_{|\beta|}$ = .096 - .176. Grade-level differences were most concentrated in the *Peer norms* scale, with three significant effects occurring there (.161, -.141, and -.140). Five significant grade-level differences were dispersed across the measures of other classroom context measures, within a range of .096 - .176. Three significant grade-level differences were observed in measures at the right side of the structural model, including one item of the *Surface learning strategies* scale (-.144), and two items of the *Justifiability of cheating* scale (-.096 and -.148).

Gender differences achieved significance in 21 of the 49 items of the measurement model, or approximately twice as many as grade-level. Fifteen were highly significant at $p < .01$. The mean absolute magnitude of these differences was $M_{|\beta|}$= .146, with a range of absolute value range $R_{|\beta|}$ = .088 - .269. Gender differences were most concentrated at the left side of the model, with the four largest magnitudes ($\beta > .200$) occurring on the *Subject self-concept* scale. The six-item scale for *Honesty-trustworthiness self-concept* entailed, by contrast, two small significant gender differences (-.110, -.158). Nine additional significant item-level gender differences were observed in measures of learning context variables (*Performance goal structure*, *Good teaching*, *Assessment quality*, and *Peer norms*), within a range of .088 - .174. Five small but significant gender differences were also observed at the right of the hypothesized structural model, in all three items of the *Justifiability of cheating* scale (-.149, .162, .162), and two of the three items of the *Self-reported cheating* scale (.143, .133).

Results of DIF analysis revealed a dispersion of low-level gender differences in the measurement model. Among the 21 gender differences that were statistically significant, magnitudes were generally small, which is of greater interest than whether they were significant. Significance is largely a function of statistical power, or whether non-zero effects can be detected (MacCallum, Browne, & Cai, 2006). If the sample size, and thus the statistical power, is increased sufficiently in any study, Kline (2011) points out, “virtually all effects that are not nil will be statistically significant” (p. 13). Small statistical differences can, even when significant, be expected to exert proportionately minor effects on the interpretability of model results.

While gender differences were found to be excessive on a cumulative basis, as indicated by the change in *CFI* between the configural and scalar models (-.014; see Table 6.9), their effects at the item-level were found to be small and well dispersed. Six of the 21 affected items, or approximately one-third, occurred in measures for *Subject self-concept* and *Honesty-trustworthiness self-concept*, including all four of the largest differences ($\beta > .200$). While non-invariance is of concern at any point in a structural model, the effects of these two constructs on the outcome variables, *Surface learning strategies* and *Self-reported cheating*, were mediated multiple times in the hypothesized model, and had, as such, less direct impact on outcome variables, in terms of variance explained, than constructs further to the right.

While the scalar non-invariance detected between genders in the foregoing two analyses was relatively low and dispersed in the measurement model, its presence advocated further examination. A conservative approach was adopted in which the hypothesized structural model was estimated with each gender's specific data, prior to being estimated with data from the full, 'co-ed' sample.

Table 6.10

Differential item functioning analysis for gender and grade-level

Item	Gender	Grade	Item	Gender	Grade	Item	Gender	Grade	Item	Gender	Grade
SUB2	.269***	.004	PERF74	.152**	.029	CURUSE64	.061	.057	SURF87	.078	.022
SUB3	.117*	-.027	PERF75	.148**	.007	TRANS28	.082	.018	SURF88	-.043	-.069
SUB5	.218***	-.057	GTEACH18	.098*	.060	TRANS32	.063	.127**	SURF91	.023	-.144**
SUB13	.258***	-.052	GTEACH33	.100*	.078	TRANS66	.067	.021	SURF97	.007	-.061
SUB15	.230***	.032	GTEACH39	.127**	-.038	AUTH44	.005	.071	CHJUST79	-.149**	-.096*
HON_1	-.110*	-.034	GTEACH50	.088*	.068	AUTH60	.040	.035	CHJUST86	.162***	-.148**
HON6	-.026	.064	GTEACH62	-.001	.176***	AUTH71	.031	.043	CHJUST99	.162***	-.050
HON8	-.040	-.055	GTEACH67	.040	.032	AUTH78	.030	.096*	CHEAT84	.143***	-.062
HON9	-.022	-.021	GTEACH68	-.055	.072	PEER24	.043	.163***	CHEAT92	.133***	-.054
HON10	-.158***	-.026	GTEACH77	-.026	.125**	PEER31	.163***	-.055	CHEAT95	.082	-.083
HON11	-.077	-.008	CURUSE19	.127**	.138**	PEER40	.076	-.141**			
PERF61	.078	.000	CURUSE53	.052	.080	PEER58	.174***	-.140**			
PERF69	.081	.047	CURUSE56	.005	.060	PEER65	.074	-.088			

Note. (N = 493) SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; *p < .05, **p < .01, ***p < .001.

6.5 MIMIC modeling: Demographic effects (N = 422)

A 'multiple-indicators multiple-causes' (MIMIC) analysis was conducted to investigate the mean-level effects of gender, grade-level, self-rated English proficiency, maternal educational attainment, paternal educational attainment, and all two-way interactions between these variables.

Model fit. Of the 493 usable questionnaires received at Time 1, 71 included incomplete demographic information, leaving a sample size of N = 422 for the purpose of MIMIC analysis. A single MIMIC model estimated with all indicators achieved borderline fit (χ^2(1689) = 2544; *RMSEA* = .035, *CI*s = .032 - .037, *pclose* = 1.00; *TLI* = .89; *CFI* = .90; *SRMR* = .049; *N:q* = 1.3; *SCF* = .942), with an *N:q* ratio of just 1.3. Lower *N:q* ratios are generally associated with lower statistical power, or a lower likelihood of detecting parameters that are different than zero. To increase statistical power, the MIMIC analysis was conducted as two separate models. The first of these included the five covariates listed above, (χ^2(1289) = 2036; *RMSEA* = .037, *CI*s = .034 - .040, *pclose* = 1.00; *TLI* = .90; *CFI* = .91; *SRMR* = .054; *N:q* = 1.8; *SCF* = .953), and the second included all two-way interactions between these variables (χ^2(1489) = 2266; *RMSEA* = .035, *CI*s = .032 - .038, *pclose* = 1.00; *TLI* = .90; *CFI* = .91; *SRMR* = .051; *N:q* = 1.5; *SCF* = .948). Both models achieved acceptable fit. However, with 230 (*N:q* = 1.8) and 275 free parameters (*N:q* = 1.5), respectively, statistical power remained a concern. Results of these analyses are presented in Appendix G.

To maximize statistical power, both MIMIC models were re-estimated with composite scores calculated according to the methodology described by Holmes-Smith and Rowe (1994), and normalized in accordance with Rowe's (2002, 2004) recommendations. The first of these models included the five covariates listed above (χ^2(12) = 16; *RMSEA* = .029, *CI*s = .000 - .061, *pclose* = .842; *TLI* = .97; *CFI* = 1.00; *SRMR* = .010; *N:q* = 4.1); and the second included all two-way interactions of these variables (χ^2(17) = 23; *RMSEA* = .028, *CI*s = .000 - .056, *pclose* = .897;

TLI = .96; *CFI* = 1.00; *SRMR* = .010; *N:q* = 2.9). Both models demonstrated excellent fit. *N:q* ratios were substantially improved by the reduction of free model parameters to 103 (*N:q* = 4.1), and 148 (*N:q* = 2.9), respectively. The results of the MIMIC models estimated with composite scores are reported in Tables 6.11a and 6.11b. The degree of difference between the models estimated with composite scores and those estimated with observed indicator variables is briefly summarized next.

Beta coefficients in the composite score model were larger, on average, by $M_{\Delta|\beta|}$ = .0005, with a mean absolute difference of $M_{|\Delta\beta|}$ = .015, and an absolute difference range of $R_{|\Delta\beta|}$ = .000 - .054. All effects that achieved significance in the MIMIC model estimated with observed indicator variables remained significant in the model estimated with composite scores. Two additional differences in *Subject self-concept*, related to (1) English language proficiency (β = .102, p = .033) and (2) the two-way interaction between gender and grade-level (β = .114, p = .043), achieved significance only in the composite score model.

Results. Gender predicted five significant mean differences among the additional eight first-order factors, and one second-order factor in the measurement model. The largest of these differences occurred with respect to *Subject self-concept* (Science), which was higher for male respondents (β = .252, p < .001). A two-way interaction between gender and grade-level in relation to the same factor (β = .114, p < .05) suggests that male respondents in Grade Eight had higher *Subject self-concept* in Science class than female respondents in Grade Nine. Males were also more likely to perceive a *Performance goal structure* (β = .123, p < .05) and *Peer norms* favorable to cheating (β = .218, p < .001), to believe in the *Justifiability of cheating* (β = .182, p < .01), and to engage in more *Self-reported cheating* (β = .136, p < .05), all in the context of Science class.

Grade-level predicted five significant mean differences among factors in the measurement model. Grade Eight respondents reported higher (more favorable) mean scores for perceptions of *Teacher quality* ($\beta = .151$, $p < .01$) and *Usefulness of curriculum* ($\beta = .141$, $p < .01$) than Grade Nine respondents. Grade Eight respondents were also less likely to perceive *Peer norms* favorable to cheating ($\beta = -.227$, $p < .001$), to believe in the *Justifiability of cheating* ($\beta = -.172$, $p < .01$), and to use *Surface learning strategies* ($\beta = -.149$, $p < .01$).

English language proficiency predicted three significant mean differences. Respondents who indicated being more proficient with English had higher mean scores for *Subject self-concept* ($\beta = .102$, $p < .05$), whereas respondents with lower self-rated English proficiency were more likely to perceive a *Performance goal structure* in Science class ($\beta = -.110$, $p < .05$), and to use *Surface learning strategies* ($\beta = -.149$, $p < .01$). The effect of the two-way interaction of gender and English on both *Surface learning strategies* ($\beta = -.215$, $p < .05$) and perceptions of a *Performance goal structure* ($\beta = -.218$, $p < .05$) indicated, additionally, that each was more prevalent among male respondents with lower self-rated English proficiency.

Parental educational attainment also predicted mean differences in latent constructs in the MIMIC model. *Subject self-concept* was predicted positively by both maternal ($\beta = .108$, $p < .05$) and paternal educational attainment ($\beta = .111$, $p < .05$). The positive effect of higher maternal educational attainment on *Subject self-concept* also appeared to improve among respondents who indicated better English language proficiency ($\beta = .123$, $p < .05$). This may, however, be more common among female respondents, since paternal educational attainment was associated among male respondents with more use of *Surface learning strategies* ($\beta = -.226$, $p < .05$). The positive effect of higher paternal educational attainment on *Subject self-concept*, appears, by contrast, to have been especially beneficial among Grade Eight respondents ($\beta = .202$, $p < .05$). Respondents who indicated that both parents were more educated were, moreover, less likely to report perceiving a *Performance goal structure* in Science class ($\beta = -.163$, $p < .05$), and more likely to report perceiving the class curriculum as useful ($\beta = .138$, $p < .05$).

Among these results, gender and grade-level emerge, again, as important variables in the Time 1 data set. Together, they accounted for ten of the fifteen one-way effects, and six of the eight two-way effects detected in this analysis. Male respondents were, in particular, more likely to report that cheating was justifiable, that their peers viewed cheating as acceptable, and that they had cheated in Science class that year. Grade Nine respondents indicated, similarly, that cheating was more acceptable among their peers and more justifiable in Science class, than did their Grade Eight counterparts. Grade Nine respondents also reported more prevalent use of surface learning strategies. The fact that these mean differences occur in constructs at the right side of the hypothesized structural model, including both outcome variables, and, in several cases, exceeding $\beta = .200$, advocates for including them as control variables in the hypothesized structural model (see Figure 6.1).

Table 6.11a

MIMIC results: standardized beta coefficients for covariates, Time 1 (N = 422)

	Gen	Gra	Eng	Mom	Dad
Person					
Subject self-concept	.252***	-.001	.102*	.108*	.111*
Honesty-trust. self-concept	-.033	-.028	.063	.021	.024
Learning context					
Performance goal structure	.123*	.020	-.110*	.025	-.084
Teacher	.078	.151**	-.031	-.001	.033
Usefulness of curriculum	.056	.141**	.043	-.005	-.026
Peer cheating norms	.218***	-.227***	.028	-.018	.029
Moral obligation					
Justifiability of cheating	.182**	-.172**	-.033	.113	-.102
Behavior					
Surface learning strategies	.008	-.149**	-.149**	.034	-.080
Self-reported cheating	.136*	-.085	-.041	.040	-.057

Note. Model fit: $\chi^2(12) = 16$; *RMSEA* = .029, *CI*s = .000 - .061, *pclose* = .842; *TLI* = .97; *CFI* = 1.00; *SRMR* = .010; $N{:}q$ = 4.1. Gen = Gender, Gra = Grade-level, Eng = English proficiency, Mom = Maternal educational attainment, Dad = Paternal educational attainment, *p < .05, **p < .01, ***p < .000.

Table 6.11b

MIMIC results: standardized beta coefficients for two-way interaction variables, Time 1 (N = 422)

	GenXGra	GenXEng	GenXMom	GenXDad	GraXEng	GraXMom	GraXDad	EngXMom	EngXDad	MomXDad
Person										
Subject self-concept	.114*	.136	.100	-.090	-.043	-.038	.202*	.123*	.016	.020
Honesty-trust. self-concept	-.054	-.076	.085	-.060	.138	-.049	.063	-.038	-.037	.000
Learning context										
Performance goal structure	.117	-.218*	.114	-.103	.075	-.097	.119	.092	.042	-.163*
Teacher	.083	.083	.071	-.092	-.109	-.093	.129	.029	.064	.056
Usefulness of curriculum	.049	.045	.015	-.167	-.018	-.075	.118	.028	.081	.138*
Peer cheating norms	.033	-.094	-.060	.033	.109	.061	.018	.018	.007	-.059
Moral obligation										
Justifiability of cheating	.039	-.155	.117	.094	.134	-.035	-.124	.039	.006	-.067
Behavior										
Surface learning strategies	-.024	-.215*	.000	-.226*	.143	.088	.100	-.044	.052	-.071
Self-reported cheating	.025	-.144	.015	-.070	.146	.020	.021	.024	-.005	-.041

Note. Model fit $\chi^2(17)$ = 23; *RMSEA* = .028, *CIs* = .000 - .056, *pclose* = .897; *TLI* = .96; *CFI* = 1.00; *SRMR* = .010; *N:q* = 2.9. Gen = Gender, Gra = Grade-level, Eng = English proficiency, Mom = Maternal educational attainment, Dad = Paternal educational attainment. *p < .05, **p < .01, ***p < .001.

6.6 Revised hypothesized structural model (Model 3)

Following the hypothesis that relational variables in classrooms, such as teacher quality, may be thought of as aspects of a teaching-learning contract, the original hypothesized model for this project, Model 1, included seventeen factors, of which ten were intended to measure students' perceptions of various aspects of the learning environment. The effects of these relational variables on academic integrity were hypothesized to be mediated by moral obligation, measured as the justifiability of cheating and perceptions of whether the workload in a given class was appropriate. Measures of deep and surface learning strategy use were also included, along with cheating, as outcome variables. Including these learning strategy variables entertained the proposition that deep learning strategies entail greater integrity, in terms of meaningful learning, than surface strategies, and that surface strategies could be grouped together with cheating under the rubric 'disintegrity' (Miller et al., 2011).

*Model 2***.** Three of the above-mentioned constructs were dropped from the hypothesized model during the Pilot Study reported in Chapter Four due to multicollinearity that exceeded r = .850 with *Good Teaching* (see section 4.7), including *Appropriate assessment*, *Mastery goal structure*, and *Clear goals and standards*. Four additional measures of learning context, *Good teaching, Authenticity of assessment, Transparency of assessment*, and *Experiences of classroom rules* that demonstrated more moderate levels of multicollinearity (r = .750 – 800), were combined into the second-order factor structure *Teacher quality*. The remaining set of fourteen first-order factors and one second-order factor were designated 'Model 2' (see Figure 4.3) and proceeded to Time 1 of the Main Study.

Model 3. Analysis of one-factor congeneric models, conducted as a preliminary step in the modeling approach at Time 1 of the Main Study, identified significant psychometric dysfunction in five of the remaining fourteen first-order factors. *Appropriate workload*, *Experience of classroom rules* and *Deep learning strategies* were dropped from the study due to congeneric model misfit, whereas *Authenticity of assessment* and *Transparency of assessment* were integrated, using pilot data, into a new seven-item factor structure. These changes reduced the number of first-order factors in the hypothesized model to ten, in addition to the second-order factor *Teacher quality*. DIF and MIMIC analyses of group mean differences additionally revealed a number of prominent effects related to gender and grade-level. These covariates were, therefore, introduced as control variables to the hypothesized model. The model resulting from these changes was designated 'Model 3' (see Figure 6.1). It is important to emphasize that none of the constructs in Model 3 were modified using Time 1 data.

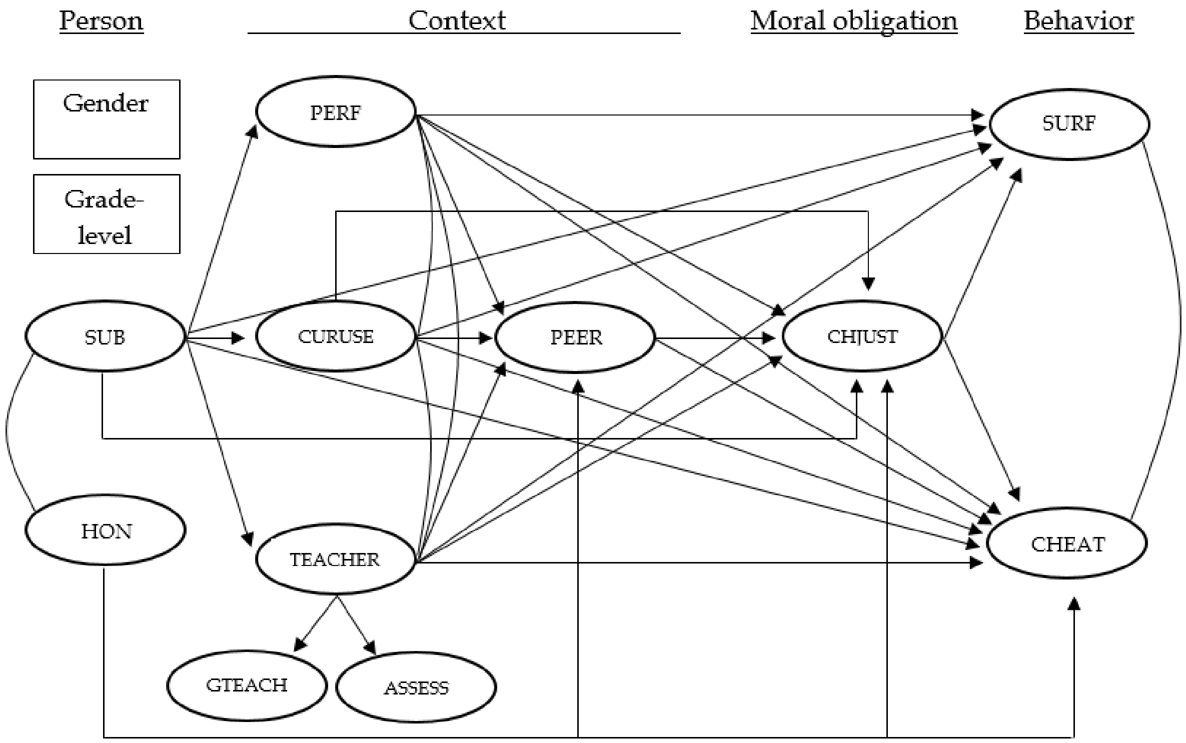

Figure 6.1. Diagram of Model 3. *Note.* SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level.

6.7 Male sample model, Time 1 (*N* = 201)

The analysis of invariance reported earlier in this chapter detected scalar non-invariance in the response patterns of male *versus* female participants. While the degree of this invariance was relatively limited in scope, it could cast doubt on the validity of the model estimated with the whole sample. To address this issue, gender-specific models were estimated in order to gauge the magnitude and import of differences caused by the scalar non-invariance. Each model was estimated initially using all observed indicator variables (see Tables 6.12 and Figure 6.2), and then re-estimated using weighted composite scores, which afforded higher *N:q* ratios. Composite score model results are presented in Table I1 of Appendix I. Differences in effect sizes estimated with composite score models *versus* those estimated with observed indicator variables are summarized below.

6.7.1 Measurement model analysis: Time 1 male sample data

Following the two-step approach to modeling applied throughout the present study, the fit and reliability of one-factor congeneric models were tested first. These models, reported in Table H1 of Appendix H were all found to have acceptable psychometric properties, with the exception of *Surface learning strategies*, for which *RMSEA* (.090) exceeded the desired threshold of .080. The model did not, however, fail the close-fit hypothesis, expressed by the lower bound confidence interval of *RMSEA*, nor did it fall short on any other fit criteria, including a non-significant χ^2 ($\chi^2(2) = 5.6$, $p = .06$).

The full multivariate measurement model was tested next against the male component of the Time 1 sample (*N* = 201), and found to demonstrate good fit ($\chi^2(1089) = 1385$; *RMSEA* = .37, *CI*s = .031 - .043, *pclose* = 1.00; *TLI* = .92; *CFI* = .93; *SRMR* = .068; *N:q* = 1.1; *SCF* = .911). The correlation matrix for this measurement model is provided in Table J1 of Appendix J.

6.7.2 Structural analysis: Time 1 male sample data

The revised hypothesized structural model, or Model 3 (see Figure 6.1), also achieved satisfactory fit to the male component of Time 1 data ($\chi^2(1135) = 1447$; *RMSEA* = .037, *CI*s = .031 - .043, *pclose* = 1.00; *TLI* = .92; *CFI* = .92; *SRMR* = .068; *N:q* = 1.1; *SCF* = .907). Effects in this model (see Figure 6.2) explained 74% of the variance in *Self-reported cheating*, 55% of the variance in *Surface learning strategies*, and 49% of the variance in *Justifiability of cheating*. Of 27 hypothesized regression paths, the fourteen that achieved statistical significance are presented in Figure 6.2. (Dashed arrows in Figure 6.2 represent effects that achieved significance when the model was estimated with the full sample.) The correlation matrix for this structural model is provided in Table J2 of Appendix J.

The pattern of effects involving class context factors was centered on *Justifiability of cheating*. Effects of class context factors on *Self-reported cheating* and *Surface learning strategies* were overwhelmingly mediated by *Justifiability of cheating*, and to a lesser extent by *Peer norms*. *Justifiability of cheating* was also the strongest predictor of both outcome variables ($\beta = .771$, $p < .001$ and $\beta = .641$, $p < .001$, respectively). All effects of context on *Self-reported cheating* were, indeed, fully mediated by *Justifiability of cheating*, whereas *Performance goal structure* was the only contextual variable to exert a direct effect on *Surface learning strategies* ($\beta = .254$, $p < .05$). The substantial zero-order correlation between *Surface learning strategies* and *Self-reported cheating* ($r = .507$; see Table J2 of Appendix J) was, moreover, fully accounted for in the model ($r = -.050$, NS) by the predictive effects of *Justifiability of cheating* and *Subject self-concept*, thus lending the first empirical support to the hypothesis that cheating and surface learning are types of disintegrity.

Usefulness of curriculum exerted, by contrast, no significant effects in the model. While its high correlation with *Teacher quality* ($r = .761$, $p < .001$) suggested that it could be incorporated with the second-order factor structure, attempts to do this were detrimental to the fit of the overall measurement model.

Table 6.12

Model 3: beta coefficients estimated with observed indicator variables, Time 1 male data (N = 201)

	Predictors							
	Grade	Sub	Hon	Perf	Curuse	Teacher	Peer	Chjust
Sub	-.010							
Hon	.014	---						
Perf	.036	-.097	---					
Curuse	.070	.522***	---	---				
Teacher	.042	.460***	---	---	---			
Peer	-.250**	---	-.139	-.037	.308	-.562*		
Chjust	-.120	.126	-.231*	.305***	-.156	-.185	.341***	
Surf	.060	-.259*	---	.254*	-.122	.291	---	.614***
Cheat	.087	-.178*	-.231**	-.100	.131	.106	.124	.771**

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level. *$p < .05$, **$p < .01$, ***$p < .001$.

Peer norms was also an important mediator in the model. The effect of *Peer norms* on *Justifiability of cheating* ($\beta = .341$, $p < .001$) fully mediated the effect of *Teacher quality*. It is likely, however, that the direct path from *Teacher quality* to *Justifiability of cheating* ($\beta = .185$, $p = .353$) is non-zero and would have achieved statistical significance in a larger sample of males, as it would later be observed to do in the model estimated with the full, co-ed sample (N = 493) (see Figure 6.4).

The effects of personological variables *Honesty-trustworthiness self-concept* and *Subject self-concept* formed a pattern within the model that clearly differed from that of class context variables. While the effects of class context on disintegrity were overwhelmingly mediated by *Justifiability of cheating*, personological variables exerted direct effects on both measures of disintegrity, and on intervening measures of class context and moral obligation. The effect of *Honesty-trustworthiness self-concept* on *Self-reported cheating* (β = -.231, p < .01) was, for instance, nearly identical to its effect on *Justifiability of cheating* (β = -.231, p < .05). The effects of *Subject self-concept* on *Surface learning strategies* (β = -.259, p < .05) and *Self-reported cheating* (β = -.178, p < .05) were, by contrast, largely unmediated, despite its effect on *Teacher quality* (β = -.460, p < .001), which may have carried through in small part to the outcome variables by way of *Peer norms* and *Justifiability of cheating*. The only significant effect of grade-level in the model was on *Peer norms* (β = -.250, p < .01), suggesting that male respondents in Grade Nine were more likely to believe that cheating in Science class was acceptable among their peers.

Weighted composite score estimation. Estimating the hypothesized model with all observed indicator variables (188 free model parameters) with a sample of 201 male subjects achieves an *N:q* ratio of just 1.1. Using composite scores to estimate the model, instead, trebled the *N:q* ratio to 3.3 by reducing the number of free parameters to 61. The CFA for the composite score model yielded excellent fit to the male data sample ($\chi^2(7)$ = 14.6, p = .04; *RMSEA* = .074, *CI*s = .014 - .127, *pclose* = .198; *TLI* = .92; *CFI* = .99; *SRMR* = .019; *N:q* = 3.5). The hypothesized structural model was also an excellent fit ($\chi^2(14)$ = 18.4, p = .19; *RMSEA* = .039, *CI*s = .000 - .084, *pclose* = .601; *TLI* = .97; *CFI* = .99; *SRMR* = .022; *N:q* = 3.3), explaining 76% of the variance in *Self-reported cheating*, 58% of the variance in *Surface learning strategies*, and 54% of the variance in *Justifiability of cheating*, an increase over the model estimated with all observed indicators of $\Delta R^2 = 2\%$, $\Delta R^2 = 3\%$, and $\Delta R^2 = 5\%$, respectively. The mean effect size of the model was larger when estimated with composite scores by $M_{\Delta\beta} = .024$, with a mean absolute difference of $M_{|\Delta\beta|} = .05$, and absolute difference range of $R_{|\Delta\beta|} = .000 - .129$. All regression paths that were significant in the model estimated with observed indicator variables remained significant when estimated with composite scores, with the exception of the path from *Subject self-concept* to *Self-reported cheating* ($\beta = .146$, $p = .143$). This path may have failed to achieve significance in the composite score model due to the increased magnitudes of nearby paths, such as that from *Justifiability of cheating* to *Self-reported cheating* ($\beta = .811$, $p < .001$). All structural regression coefficients for the composite score model for male respondents at Time 1 are presented in Table I1 of Appendix I.

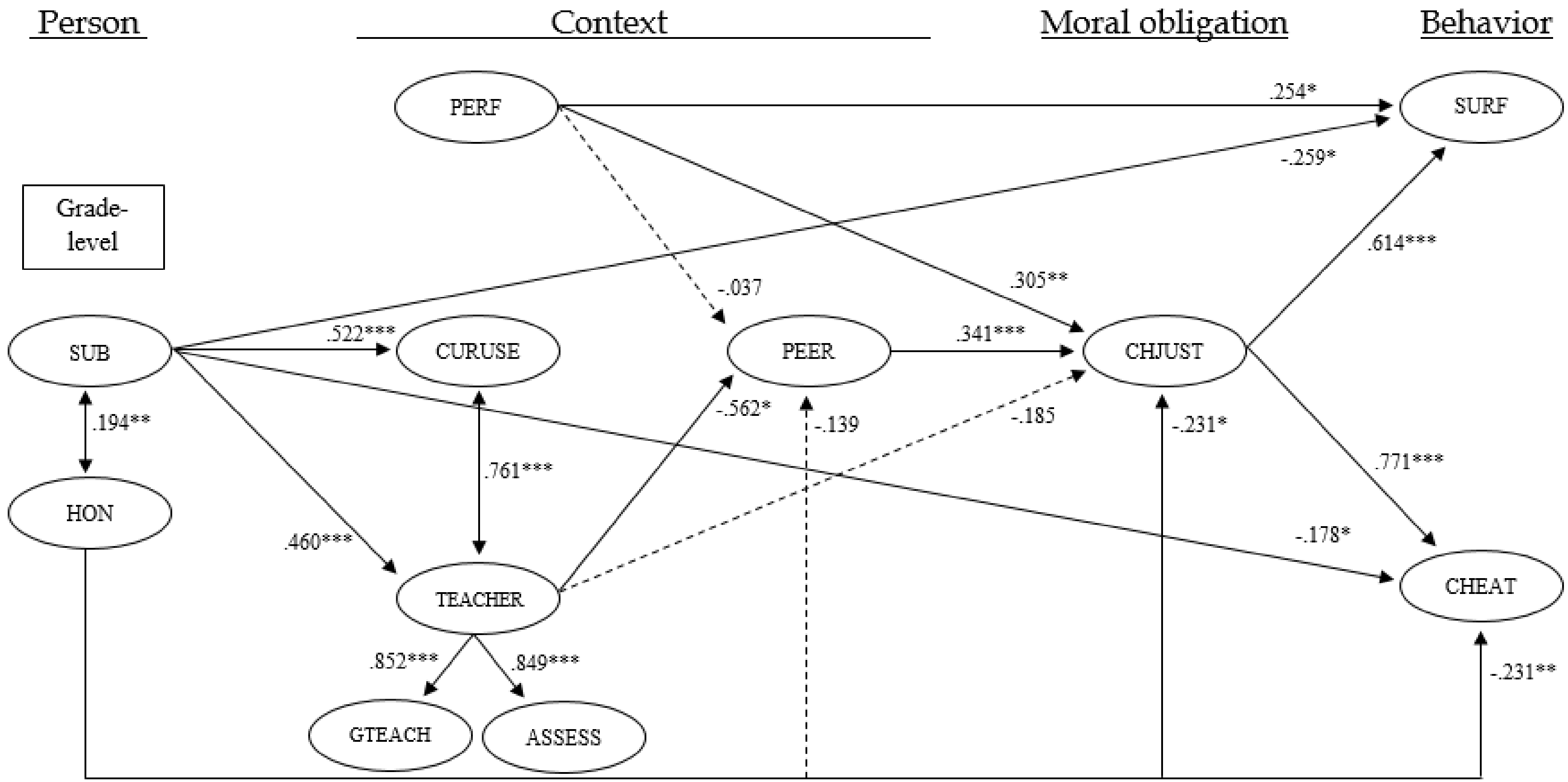

Figure 6.2. Male sample results for Model 3, Time 1 (N = 201). $\chi^2(1135) = 1447$; *RMSEA* = .037, *CI*s = .031 - .043, *pclose* = 1.00; *TLI* = .92; *CFI* = .92; *SRMR* = .068; $N{:}q$ = 1.1; *SCF* = .907. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level. **p* < .05, ***p* < .01, ****p* < .001. - - - - paths are significant in the model estimated for the co-ed sample (Figure 6.4).

6.8 Female sample model, Time 1 (N = 292)

The hypothesized structural model was estimated next with the female component of the Time 1 data set (N = 292). Results are compared to those of the model tested against the male data set.

6.8.1 Measurement model analysis: Time 1 female sample data

Analysis of one-factor congeneric models yielded acceptable psychometric properties for all factors, reported in Table H2 of Appendix H. Scale reliabilities also fell within an acceptable range, with the exception of *Justifiability of cheating* (Rho = .67), meaning that, among female respondents, more than 30% of the variance captured by this measure was due to random error (Kline, 2011). This violation of the established threshold of Rho $\geq$ .70 is somewhat detrimental to the construct validity of *Justifiability of cheating*, in that Rho values are used to fix the factor loadings and error variances of latent composite scores. The deficit is relatively small (-.03), however, and random error is explicitly estimated in structural equation modeling, in the form of error terms (Kline, 2011).

The multivariate measurement model was estimated next using observed indicator variables, and found to have satisfactory fit ($\chi^2(1089)$ = .1560; *RMSEA* = .039, *CI*s = .034 - .043, *pclose* = 1.00; *TLI* = .91; *CFI* = .92; *SRMR* = .061; *N:q* = 1.6; *SCF* = .939). The correlation matrix for the measurement model is provided in Table K1 of Appendix K.

6.8.2 Structural analysis: Time 1 female sample data

Structural Model 3 also demonstrated satisfactory fit to the female data set ($\chi^2(1135)$ = 1663; *RMSEA* = .040, *CI*s = .036 - .048, *pclose* = 1.00; *TLI* = .91; *CFI* = .91; *SRMR* = .066; *N:q* = 1.6; *SCF* = .940) explaining 63% of the variance in *Self-reported cheating*, 36% of the variance in *Surface learning strategies*, and 62% of the variance in *Justifiability of cheating*. The correlation matrix for the structural model is provided in Table K2 of Appendix K.

As in the male data set, *Justifiability of cheating* was the strongest predictor of both outcome variables, *Surface learning strategies* (β = .404, p < .001) and *Self-reported cheating* (β = .732, p < .001). *Justifiability of cheating* also fully mediated the effects of learning context on both outcome variables, such that, unlike the male sample model, even the direct effect of *Performance goal structure* on *Surface learning strategies* was non-significant. The substantial zero-order correlation between *Surface learning strategies* and *Self-reported cheating* (r = .532; see Table K2 of Appendix K) was also, again, non-significant in the female sample model (r = .188, NS), due principally to the variance they shared through *Justifiability of cheating*.

While the effect of *Peer norms* on *Justifiability of cheating* (β = .461, p < .001) was substantially stronger in the female data set than in the male data set ($\Delta\beta$ = .120), the mediating role it played in the effect of *Teacher quality* on *Justifiability of cheating* was, among females, only partial. The effect of *Teacher quality* on *Peer norms* (β = -.423, p < .001) was, additionally, weaker among female respondents ($\Delta\beta$ = .139).

As observed in the male sample model, *Subject self-concept* demonstrated a pattern of effects among females that was distinct from that of class context variables. Its effects on both disintegrity measures were largely unmediated in each gender-specific model. Its effect on *Self-reported cheating* (β = -.334, p < .001) was, however, nearly twice as strong among females as among males, whereas its effect on *Surface learning strategies* (β = .176, p < .05) was commensurate to that observed among males.

The effect of *Honesty-trustworthiness self-concept* on *Justifiability of cheating* (β = -.262, p < .001) was also commensurate to that among males, whereas its effect on *Self-reported cheating* was, by contrast, non-significant among females. *Honesty-trustworthiness self-concept* did not, in fact, predict *Peer norms* in either gender-specific model, but did predict it in the co-ed model, which may reflect the lower levels of statistical power achieved with these smaller, gender-specific samples (MacCallum et al., 2006).

Usefulness of curriculum exerted no significant beta effects in either the female sample model or the male sample model. The correlation between *Usefulness of curriculum* and *Teacher quality* (r = .536, p < .001) was also substantially lower among female respondents than among male respondents, further justifying the decision not to incorporate *Usefulness of curriculum* in the latter second-order factor structure.

Table 6.13

Model 3: beta coefficients estimated with observed indicator variables female data, Time 1 (N = 292)

	Predictors							
	Grade	Sub	Hon	Perf	Curuse	Teacher	Peer	Chjust
Sub	-.043							
Hon	-.033	---						
Perf	.048	-.051	---					
Curuse	.124*	.478***	---	---				
Teacher	.181*	.468***	---	---	---			
Peer	-.072	---	.116	.244**	-.048	-.423***		
Chjust	-.008	.025	-.262***	.166*	.002	-.291*	.461***	
Surf	-.109	-.176*	---	.079	-.054	-.060	---	.404***
Cheat	-.035	-.334***	-.070	.040	.046	.104	-.104	.732***

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level. *p < .05, **p < .01, ***p < .001.

Summary of gender-specific model differences. Three potentially important areas of difference between gender-specific models were noted in the foregoing comparative analysis. Firstly, the effects of *Performance goal structure* were more oriented to *Justifiability of cheating* and *Surface learning strategies* among males, and more oriented to *Peer norms* among females. Secondly, the effect of *Teacher quality* was fully mediated by *Peer norms* among males, whereas *Teacher quality* exerted a significant direct effect on *Justifiability of cheating* among females. Thirdly, *Self-reported cheating* was predicted by both *Honesty-trustworthiness self-concept* and *Subject self-concept* among males, whereas this pattern shifted entirely in favor of *Subject self-concept* among females.

Weighted composite score estimation. Using composite scores to estimate the female sample model (N = 292) increased the $N{:}q$ ratio for the measurement and structural models to 5.0 and 4.8, respectively. The fit of the measurement model was excellent ($\chi^2(7)$ = 7.5, p = .38; *RMSEA* = .016, *CI*s = .000 - .075, *pclose* = .770; *TLI* = 1.00; *CFI* = 1.00; *SRMR* = .010; $N{:}q$ = 5.0), whereas the fit of the structural model showed some decline, but remained good ($\chi^2(14)$ = 25.7, p = .03; *RMSEA* = .053, *CI*s = .017 - .086, *pclose* = .391; *TLI* = .94; *CFI* = .99; *SRMR* = .035; $N{:}q$ = 4.8). The structural model explained 69% of the variance in *Self-reported cheating*, 47% of the variance in *Surface learning strategies*, and 66% of the variance in *Justifiability of cheating*, an increase over the model estimated with all observed indicators of ΔR^2 = 6%, ΔR^2 = 11%, and ΔR^2 = 3%, respectively. The mean effect size of the model was $M_{\Delta\beta}$ = .026 larger when estimated with composite scores, with a mean absolute difference of $M_{|\Delta\beta|}$ = .05, and absolute difference range of $R_{|\Delta\beta|}$ = .000 - .138. The largest of these differences in path magnitude ($\Delta\beta$ = .138) represented an increase in the estimated effect of *Justifiability of cheating* on *Surface learning strategies*, which appears to explain the correspondingly large increase in variance explained in the latter factor (ΔR^2 = 11%, noted above). All regression paths that were significant in the model estimated with observed indicator variables remained so when estimated with composite scores. One additional path, from *Honesty-trustworthiness self-concept* to *Peer norms* (β = -.148, p = .030) achieved significance only in the composite score model. All structural regression coefficients for the composite score model for female respondents at Time 1 are presented in Table I2 of Appendix I.

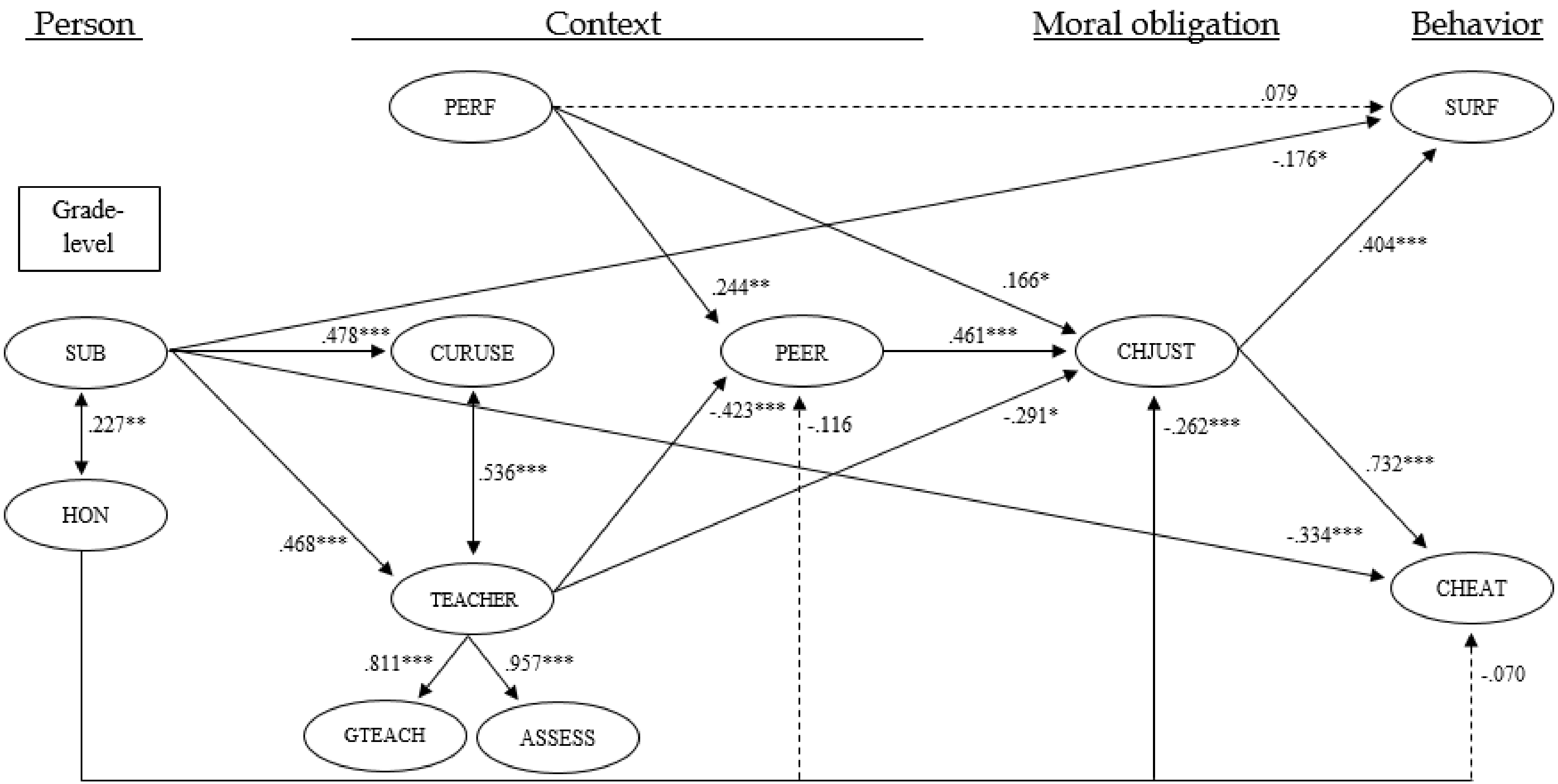

Figure 6.3. Female sample results for Model 3, Time 1 (N = 292). $\chi^2(1135)$ = 1663; *RMSEA* = .040, *CI*s = .036 - .048, *pclose* = 1.00; *TLI* = .91; *CFI* = .91; *SRMR* = .066; *N:q* = 1.6; *SCF* = .940. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level. - - - - paths are significant in the model estimated for the whole sample (Figure 6.4).

6.9 Co-ed sample structural model: Time 1 (N = 493)

The measurement model for the co-ed sample underwent detailed analysis in sections 6.1 – 6.5 of this chapter, and was found, overall, to have good validity and reliability. Analyses of multi-group factorial invariance raised a concern about potential differences between gender groups, due to scalar non-invariance. Several differences were noted in Section 6.8 between the model fitted for male respondents and that fitted for female respondents. Most of these differences reflected changes in the magnitudes of beta coefficients, however, which is more suggestive of mean-level differences than of differences in the operational definitions of the factor measures.

The fit of the structural model to the co-ed sample (N = 493) was acceptable ($\chi^2(1175) = 1962$; *RMSEA* = .037, *CI*s = .034 - .040, *pclose* = 1.00; *TLI* = .91; *CFI* = .92; *SRMR* = .056; *N:q* = 2.5; *SCF* = .962), explaining 67% of the variance in *Self-reported cheating,* 41% of the variance in *Surface learning strategies*, and 55% of the variance in *Justifiability of cheating*. Detailed output related to this model is provided in Appendix L.

As seen in both gender-specific models, two distinctive effect patterns emerged in Model 3 for the co-ed sample. The first of these patterns involved the effects of contextual variables, which were overwhelmingly mediated by *Justifiability of cheating*. The only exception to this pattern was the effect of *Performance goal structure* on *Surface learning strategies* (β = .157, $p < .05$), which also appeared in the male sample model (see Figure 6.2). The effects of *Justifiability of cheating* on *Surface learning strategies* (β = .505, $p < .001$) and *Self-reported cheating* (β = .730 $p < .001$) were again sufficient, in conjunction with the effects of *Subject self-concept*, to reduce an otherwise substantial zero-order correlation between the two outcome variables (r = .512; see Table 6.8) to non-significance (r = .079, NS; see Appendix L).

As observed in the gender-specific models, the greatest effect on *Justifiability of cheating* in the co-ed model was exerted by *Peer norms* ($\beta = .379$, $p < .001$). *Peer norms* was, in turn, most affected by *Teacher quality* ($\beta = -.396$, $p < .001$), the indirect effect of which was $\beta = -.150$, $p < .001$ (see indirect effects for the Time 1 co-ed model in Appendix M). *Teacher quality* also exerted substantial indirect effects on *Self-reported cheating* ($\beta = -.314$, $p < .001$) and *Surface learning strategies* ($\beta = -.202$, $p < .01$), by way of *Peer norms* and *Justifiability of cheating*. It is helpful to re-emphasize here that the causal language used to discuss these results is supported both by strong theory and by the results of experimental manipulations discussed in the literature review and summarized again in Chapter Five (see section 5.1).

Usefulness of curriculum failed again to exert any significant predictive effects in the model. A stepwise regression was conducted to investigate how the strength of its effects on downstream variables changed as additional constructs were added to the model. The downstream outcome variables in this analysis were: (1) *Peer norms*, (2) *Justifiability of cheating*, (3) *Surface learning strategies*, and (4) *Self-reported cheating* (see Appendix N). *Usefulness of curriculum* predicted all four variables in bivariate regression models, and tended to become non-significant when *Good teaching* was added to the equation. This pattern was seen with respect to *Peer norms, Justifiability of cheating,* and *Surface learning strategies*. In the stepwise regression used to predict *Self-reported cheating*, however, the beta path from *Usefulness of curriculum* became non-significant when *Subject self-concept* was added to the equation. The role of *Subject self-concept* as a control variable in the co-ed structural equation model (Model 3), made it less likely, however, that it was having the same sort of negating effect on *Usefulness of curriculum* in Model 3 that it had in stepwise regression, where all predictors were simply correlated. The effects of *Usefulness of curriculum* on downstream variables in Model 3 appeared, instead, to be curtailed by a large correlation with *Teacher quality* ($r = .619$), which is suggestive of complete mediation. *Teacher quality* was next positioned as a mediator for

Usefulness of curriculum by regressing it on the latter in an equivalent model (see Appendix P and section 6.9.1) (Kline, 2011), which achieved identical fit to Model 3 ($\Delta\chi^2(1) = 0$).

Table 6.14

Co-ed sample model: beta coefficients estimated with observed indicator variables, Time 1 (N = 493)

	Predictors								
	Grade	Gender	Sub	Hon	Perf	Curuse	Teacher	Peer	Chjust
Sub	-.030	.268***							
Hon	-.014	-.086	---						
Perf	.039	.180**	-.072	---					
Curuse	.103*	-.072	.514***	---	---				
Teacher	.123**	-.054	.493***	---	---	---			
Peer	-.154**	.166**	---	-.135*	.132*	.049	-.396***		
Chjust	-.057	.097*	.071	-.242***	.237***	-.064	-.249*	.379***	
Surf	-.048	-.051	-.216**	---	.157*	-.073	.077	---	.505***
Cheat	.014	.028	-.258***	-.160**	-.035	.062	.146	.057	.730***

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level; *$p < .05$, **$p < .01$, ***$p < .001$.

The second effect pattern in the co-ed model involved personological variables: the effects of *Subject self-concept* on *Surface learning strategies* ($\beta = -.216$, $p < .01$) and *Self-reported cheating* ($\beta = -.258$, $p < .001$); and of *Honesty-trustworthiness self-concept* on *Self-reported cheating* ($\beta = -.160$, $p < .01$). These effects bypassed complete mediation by *Justifiability of cheating* and *Peer norms* that was observed with respect to class context effects. The effect of *Subject self-*

concept on *Teacher quality* (β = .493, p < .001), and the effects of *Honesty-trustworthiness self-concept* on *Peer norms* (β = -.135, p < .05) and *Justifiability of cheating* (β = -.242, p < .001), were, however, transmitted to the outcome variables as significant indirect effects: *Honesty-trustworthiness self-concept* on *Self-reported cheating* (β = -.221, p < .001), and *Subject self-concept* on *Surface learning strategies* (β = -.097, p < .05) (see Appendix M).

Weighted composite score estimation. Using composite scores to estimate Model 3 with the co-ed sample (N = 493) increased the $N{:}q$ ratio for the structural model to 7. The fit of the measurement model to the co-ed data set was excellent ($\chi^2(7)$ = 13.8, p = .054; *RMSEA* = .045, *CI*s = .000 - .079, *pclose* = .552; *TLI* = .97; *CFI* = 1.00; *SRMR* = .011; $N{:}q$ = 8.5). The fit of the hypothesized structural model was also excellent ($\chi^2(15)$ = 26.8, p = .03; *RMSEA* = .040, *CI*s = .012 - .064, *pclose* = .728; *TLI* = .97; *CFI* = .99; *SRMR* = .021; $N{:}q$ = 7), explaining 72% of the variance in *Self-reported cheating*, 47% of the variance in *Surface learning strategies*, and 60% of the variance in *Justifiability of cheating*, a difference compared to the model estimated with all observed indicators of ΔR^2 = 5%, ΔR^2 = 6%, and ΔR^2 = 5%, respectively. The mean effect size of the model was $M_{\Delta\beta}$ = .001 larger when estimated with composite scores, with a mean absolute difference of $M_{|\Delta\beta|}$ = .03, and absolute difference range of $\mathrm{R}_{|\Delta\beta|}$ = .000 - .083. The overall pattern of significant regression paths in the model estimated with composite scores was very similar to that in the model estimated with observed indicator variables. One path that achieved low/borderline statistical significance in the latter model, from *Performance goal structure* to *Peer norms* (β = .132, p = .035), became non-significant when estimated with composite scores (β = .098, p = .114). The failure of this path to achieve significance in the composite score model might reflect the relative increases in magnitude of nearby paths, such as the path from *Honesty-trustworthiness self-concept* on *Peer norms* (Δβ = .015) and from *Teacher quality* to *Justifiability of cheating* (Δβ = .072). All structural regression coefficients for the composite score model for the Time 1 co-ed sample are presented in Appendix O.

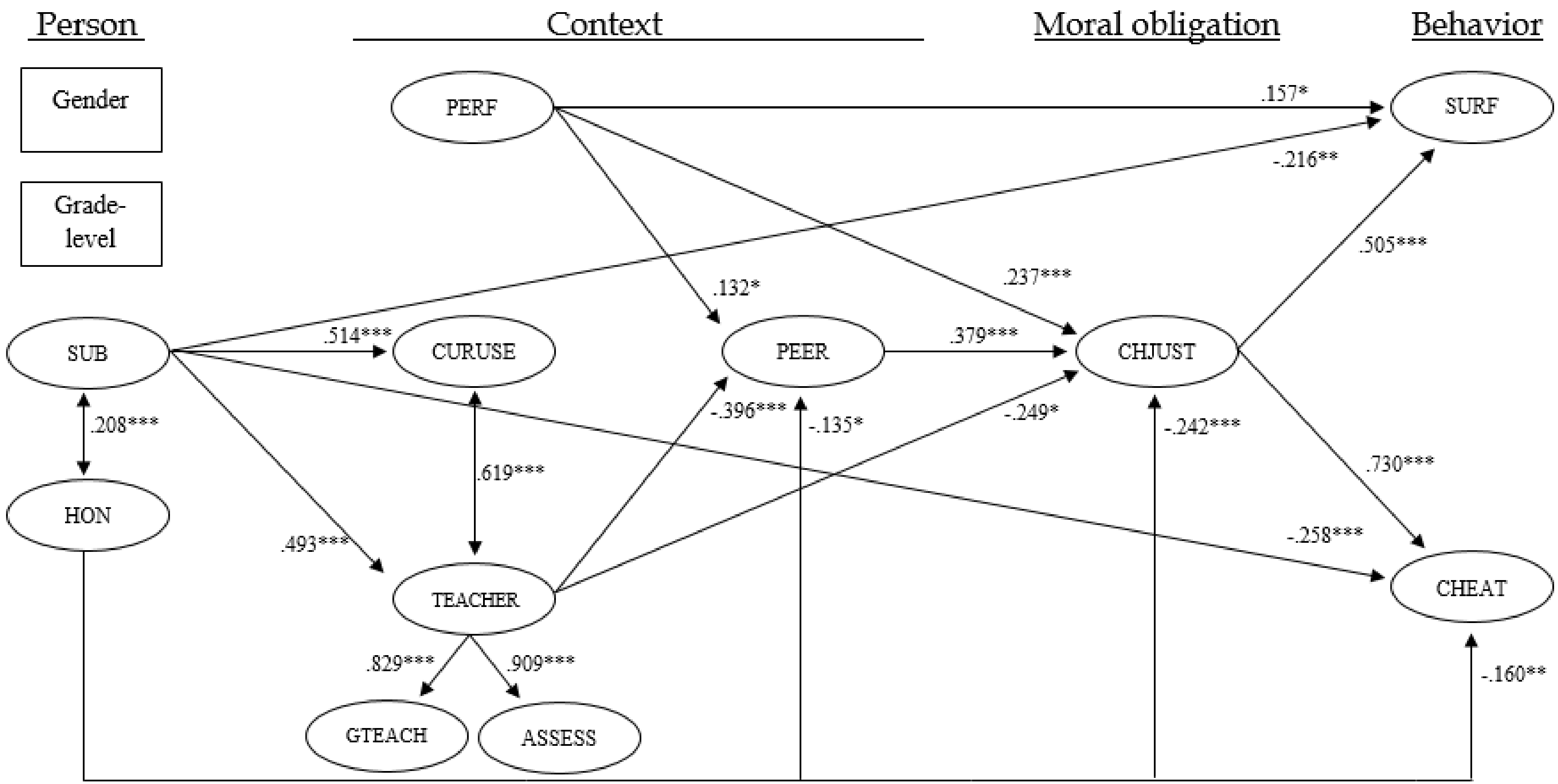

Figure 6.4. Co-ed sample results for Model 3, Time 1 (N = 493).$\chi^2(1175) = 1962$; *RMSEA* = .037, *CI*s = .034 - .040, *pclose* = 1.00; *TLI* = .91; *CFI* = .92; *SRMR* = .056; *N:q* = 2.5; *SCF* = .962. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level.

6.9.1 Equivalent models

Three 'equivalent structural models' (Kline, 2011; MacCallum & Austin, 2000) are examined below. Firstly, the absence of predictive effects exerted by *Usefulness of curriculum* in Model 3 suggests the possibility that it should be positioned as a predictor, rather than a correlate, of *Teacher quality* (Equivalent Model 1). Secondly, the decision to position *Peer norms* as a mediator between class context variables and *Justifiability of cheating* was based on the theoretical assertions of social comparison theory (Broeckelman-Post, 2008; Festinger, 1954), discussed in Chapters Two (see section 2.4.5) and Three (see section 3.4.5). Lacking a strong body of empirical research to justify this decision, however, the possibility that *Peer norms* would produce better model fit as either a correlate (Equivalent Model 2) or a predictor (Equivalent Model 3) of class context factors was tested.

Equivalent Model 1: Usefulness of curriculum as a predictor of *Teacher quality*: *Usefulness of curriculum* was positioned in Equivalent Model 1 as a predictor of *Teacher quality* (see Appendix P). The fit of this model was identical to that of Model 3 $\chi^2(1176) = 1962$; *RMSEA* = .037, *CI*s = .034 - .040, *pclose* = 1.00; *TLI* = .91; *CFI* = .92; *SRMR* = .059; *N:q* = 2.5; SCF = .962. The amounts of variance explained in *Self-reported cheating* (67%), *Justifiability of cheating* (55%), *Surface learning strategies* (41%) were also identical to Model 3. A notable difference in this equivalent model over Model 3 was the strong mediating role *Usefulness of curriculum* played between *Subject self-concept* and *Teacher quality*. While this mediation effect appears to suggest that students judge teachers largely based on whether they think their class curricula are worthwhile, no literature could be found to support the directionality of this effect. There was, moreover, no difference in χ^2 between Equivalent model 1 and Model 3.

Equivalent Model 2: Peer norms as correlate of class context. *Peer norms* was positioned in Equivalent Model 2 as a correlate of *Teacher quality*, *Usefulness of curriculum*, and *Performance goal structure*; and predicted by *Honesty-trustworthiness self-concept* (see Appendix

Q). This model fit the data well ($\chi^2(1175) = 1971$; *RMSEA* = .037, *CI*s = .034 - .040, *pclose* = 1.00; *TLI* = .91; *CFI* = .92; *SRMR* = .059; *N:q* = 2.5; SCF = .962) and explained amounts of variance in *Self-reported cheating* (67%), *Justifiability of cheating* (54%), and *Surface learning strategies* (40%) that were almost identical to Model 3, albeit with slightly less variance explained in each of latter two constructs of $\Delta R^2 = -.01$, respectively. Differences in fit, as well as path strength and significance were similarly minor. Because both models have the same degrees of freedom (1175), however, the larger χ^2 value of Equivalent Model 1 indicates a small decrement in fit over Model 3 ($\chi^2(1175) = 1962$).

Equivalent model 3: Peer norms as predictor of class context. *Peer norms* was positioned in Equivalent Model 3 as a predictor of *Teacher quality*, *Usefulness of curriculum*, and *Performance goal structure*; and as predicted by *Honesty-trustworthiness self-concept* (see Appendix R). This model fit the data well ($\chi^2(1175) = 1965$; *RMSEA* = .037, *CI*s = .034 - .042, *pclose* = 1.00; *TLI* = .91; *CFI* = .92; *SRMR* = .056; *N:q* = 2.5; SCF = .962), and explained amounts of variance in *Self-reported cheating* (67%), *Justifiability of cheating* (55%), *Surface learning strategies* (41%) that were identical to Model 3. The beta path from *Honesty-trustworthiness self-concept* to *Peer norms* increased in strength and significance from $\beta = -.131$, $p < .05$ in the hypothesized model to $\beta = -.221$, $p < .001$ in Equivalent Model 3. With the same degrees of freedom, however, the slightly larger χ^2 value associated with Equivalent Model 3 indicated that its fit was roughly equivalent to that of Model 3 ($\chi^2(1175) = 1962$).

Equivalent model 4: *Peer norms* as a correlate of *Justifiability of cheating*. While the theoretical basis for including *Peer norms* in the present study comes from the proposition in Social comparison theory that individuals' judgments are shaped by what they perceive of their peers' judgments, the possibility was entertained that when students judge cheating to be justifiable they also become more likely to believe their peers feel the same way. This hypothesis was modeled by positioning *Peer norms* as a correlate of *Justifiability of cheating* in

Equivalent Model 4. The fit of this model was identical to that of Model 3 (Figure 6.4) ($\chi^2(1175)$ = 1962; *RMSEA* = .037, *CI*s = .034 - .042, *pclose* = 1.00; *TLI* = .91; *CFI* = .92; *SRMR* = .056; *N:q* = 2.5; SCF = .962). Path coefficients observed in this model were identical to those in Figure 6.4 with the exception of inflated direct effects of *Teacher quality* and *Performance goal structure* on *Justifiability of cheating* (βs = -.400 and .287, p < .001, respectively), due to the absence of mediation by *Peer norms*. The correlation between *Peer norms* and *Justifiability of cheating* (r = .439, p < .001) was consistent with the corresponding bivariate correlation in the measurement model (r = .600, p < .001; see Table 6.8). Both models also explained the same amount of variance in *Self-reported cheating*. The variance explained in *Justifiability of cheating* was, however, considerably lower in Equivalent Model 4 (44%) than in the hypothesized PTLC model (55%). This, together with the stronger theoretical basis for positioning *Peer norms* as a mediator of the effects of class context on *Justifiability of cheating* (see Section 3.4.5), argued in favor of retaining the model in Figure 6.4.

6.10 Chapter summary

Analyses in this chapter provided the first explicit tests of two recent theoretical developments related to cheating: (1) a contractarian perspective that posits a reciprocal relationship between students' perception of the quality of a given class context and their sense of moral obligation to be honest in that context, and (2) that surface learning behaviors may be grouped together with academic cheating under the term 'disintegrity'.

Several of the measures selected to test these hypotheses were new to secondary populations and, in some cases, new to structural equation modeling. The set of fourteen first-order factors that emerged from the Pilot Study was dubbed Model 2. When this set of factors was examined at Time 1 of the Main Study, problems with congeneric model fit were addressed either by dropping measures or by using pilot data to modify them. This rendered a measurement model comprising ten first-order factors and one second-order factor. This measurement model was subjected, additionally, to multi-group invariance analysis, DIF analysis, and MIMIC analysis, by which prominent group differences were identified. Substantial mean differences observed between gender and grade-level groups prompted their inclusion as control variables in the hypothesized structural model, dubbed Model 3. No *post-hoc* modifications made solely on the basis of Time 1 data were included in Model 3.

A low level of scalar non-invariance between gender groups was found to operate across a number of measures. This prompted comparative analyses of gender-specific versions of Model 3. Prominent differences in the pattern of beta effects were noted with respect to *Performance goal structure*, *Teacher quality*, and *Subject self-concept*. Model 3 was then tested against the co-ed sample as a whole. All three of these models fit the data well, and were largely consistent with one another.

Weighted composite scores were used to estimate four models at Time 1 (MIMIC, male structural, female structural, and co-ed structural). Composite score estimation offered the opportunity to examine models with higher *N:q* ratios than estimation with observed indicator variables afforded. While differences in the effect sizes of models estimated by these two methods were consistently small, effect sizes in composite score models tended to be slightly larger than their observed variable counterparts. Such differences are consistent with improved statistical power, which is expected to accompany larger *N:q* ratios (MacCallum et al., 2006).

The hypothesized model tested in this chapter subjected a number of insights related to academic integrity gained through prior experimental research to a 'real-world' examination. The fact that the study builds largely from experimental research provides a strong basis for causal interpretations of results. Two patterns were observed in both gender-specific models and the co-ed model. The first of these entailed the complete mediation of most class context effects on outcome variables by *Justifiability of cheating*, which was, in turn, the strongest predictor of these outcomes in the model. The second pattern entailed the two personological factors, *Subject self-concept* and *Honesty-trustworthiness self-concept*, exerting direct effects on outcome variables that were, at most, partially mediated by intervening variables such as *Justifiability of cheating*. These cross-sectional findings are, taken together, consistent with the differential effects of person and situation that have often been observed in the literature, and provide support for the contractarian perspective that the moral obligation students feel to be honest tends to fluctuate positively with their assessments of the quality of a given class context.

CHAPTER 7

CROSS-SECTIONAL ANALYSES OF TIME TWO DATA

The object of Time 2 analyses was to re-evaluate the hypothesized measurement and structural models tested at Time 1, one year later. Time 2 analyses followed the overall analytical procedure followed at Time 1, including congeneric and multivariate measurement modeling followed by tests of multi-group factorial invariance. The structural model was also tested against each gender-specific data set, respectively, as well as against the co-ed sample as a whole.

7.1 Basic descriptive statistics

Means, standard deviations, distributional statistics, and reliability estimates of all congeneric factor models at Time 2 are reported in Table 7.1. Standard deviations, factor means, and Rho reliability estimates observed for Time 2 were very similar to those of Time 1. Time 2 means were larger, on average, by $M_{\Delta M}$ = .004, whereas standard deviations tended to be slightly smaller, with an average difference of $M_{\Delta SD}$ = -.022. Time 2 measures also demonstrated slightly better Rho reliability, on average, ($M_{\Delta\rho}$ = +.035), with an absolute difference range of $R_{|\Delta\rho|}$ = .01 - .14.

Time 2 data was less skewed than Time 1 data, on average ($M_{|\Delta S|}$ = .066), with an absolute difference range of $R_{|\Delta S|}$ = .031 - .392, but more kurtotic ($M_{|\Delta K|}$ = .042), with an absolute value range of $R_{|\Delta K|}$ = .013 - .826. Several specific differences in measures of skewness and kurtosis between Times 1 and 2 are worthy of note. Data for *Performance goal* *structure* and *Surface learning strategies* were more than twice as skewed at Time 2 (ΔS = .219 and ΔS = .338, respectively), whereas the data for *Good teaching* at Time 2 was approximately half as skewed (ΔS = -.392). Data for *Assessment quality*, *Peer norms*, and *Surface learning strategies* were, moreover, less than half as kurtotic at Time 2 (ΔK = -.315, ΔK = -.296, and ΔK = -.343, respectively). All measures for skewness and kurtosis at Time 2 fell within the recommended limits of kurtosis = 7.0, and skewness = 2.0 (Curran, West & Finch, 1997).

Table 7.1

Comparison of descriptive statistics from Time 1 and Time 2

	Mean		SD		Skew		Kurtosis		Rho	
	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1
Subject self-concept	2.31	2.48	.997	.929	.499	.364	-.407	-.420	.92	.91
Honesty-trust. self-concept	1.99	2.01	.742	.658	1.22	.840	2.08	1.254	.96	.82
Performance structure	3.50	3.33	.986	1.00	-.367	-.148	-.627	-.733	.79	.74
Good teaching	2.45	2.47	.821	.802	.377	.769	-.473	.636	.88	.86
Usefulness of curriculum	2.45	2.50	.974	.970	.500	.531	-.213	-.171	.92	.89
Assessment quality	2.15	2.19	.681	.658	.414	.506	.258	.573	.80	.85
Peer norms	3.70	3.56	.893	.888	-.477	-.405	-.080	-.376	.85	.78
Surface learning strategies	3.57	3.58	.940	.968	-.625	-.287	.208	-.719	.76	.71
Justifiability of cheating	4.07	4.05	.890	.913	-.785	-.875	-.117	.226	.77	.73
Self-reported cheating	4.28	4.30	1.01	.926	-1.38	-1.256	.941	.715	.86	.87

7.2 Congeneric model fit of factors in the central measurement model

Seven of the ten congeneric models presented in Table 7.2 demonstrated satisfactory fit to Time 2 data, whereas fit criteria for three models fell wide of thresholds established for the Main Study. The three misfit models were *Performance goal structure*, *Assessment quality*, and *Surface learning strategies*, of which only the latter was found to merit modification. The misfit of *Surface learning strategies* appeared to be caused by a method effect resulting from the similar wording of two items, and their close physical proximity on the questionnaire instrument (see section 7.2.3). *Surface learning strategies* was modified in order to represent this empirical reality.

7.2.1 Performance goal structure

The congeneric model for *Performance goal structure* demonstrated slight weakness with respect to two approximate fit indices, *RMSEA* and *TLI*. While the *RMSEA* point estimate (.109) and upper-bound confidence interval (.185) exceeded established thresholds of .080 and .10, respectively, its lower-bound confidence interval (.044) fell below the .050 threshold, which supported the good-fit hypothesis (Kline, 2011). The *TLI* estimate of .89 was, moreover, only slightly below the threshold of .90, and was also accompanied by a comparatively robust *CFI* estimate of .96. Together, these fit estimates appeared to indicate a degree of misfit to the data that was small and ultimately acceptable within the framework of the existing study. Modifying *Performance goal structure* or dropping it from Model 3 would have been grossly disproportionate to the degree of its misfit.

7.2.2 Assessment quality

While the congeneric model for *Assessment quality* demonstrated weakness in *RMSEA*, which was excessive across both confidence intervals (.063 and .119) and the point-estimate (.090), its performance with respect to *TLI* (.91) was acceptable, and *CFI* (.94) was only slightly lower than the threshold of .95. *Assessment quality* was, moreover, incorporated within the second-order factor structure *Teacher quality*.

Teacher quality. The two-factor model for *Teacher quality* was estimated, as at Time 1, by regressing it on the covariate 'maternal educational attainment', in order to identify the beta matrix. Maternal educational attainment had negligible correlations with both *Assessment quality* (r = -.023) and *Good teaching* (r = -.069), and explained negligible variance in the second-order factor (R^2 = .1%). As such, its role in identifying the second-order model was instrumental (Kenny, 2014b). The fit of *Teacher quality* met all desiderata for multivariate models (χ^2(103) = 200; *RMSEA* = .056, *CI*s = .045 - .068, *pclose* = .177; *TLI* = .93; *CFI* = .94; *SRMR* = .043; *N:q* = 6.2). Because *Assessment quality* functions in Model 3 as a component of *Teacher quality*, which demonstrated good overall fit, it was not modified or dropped from the study.

7.2.3 Surface learning strategies

The pattern of fit statistics for *Surface learning strategies* involved weakness in *RMSEA* across both the point estimate (.126) and confidence intervals (.062 - .201), as well as in *TLI* (.88). While these values fell only slightly wide of those for *Assessment quality*, misfit in *Surface learning strategies* was not mitigated by inclusion in a second-order factor structure.

The two largest Lagrange multiplier values to appear in the Modification Indices for *Surface learning strategies* were of equal magnitude (8.42), indicating residual covariances between items Surf97 with Surf88, and Surf91 with Surf87 (see wording in Chapter Four, section 4.5.8). The latter of these covariances, Surf91 with Surf87, also achieved the largest Lagrange multiplier value for *Surface learning strategies* at Time 1 (5.49). Examination of these items revealed both similar wording and close physical proximity on the questionnaire. Both item Surf91 (I study for Science class by skipping over parts I think the teacher will not ask questions about) and Surf97 (I study for Science class by skipping parts I do not find important), query the tendency to skip parts of study material based on assessments of whether those parts are important, either to the teacher or to themselves. The fact that these items were separated by only 4 lines on the questionnaire instrument likely increased the variance they shared for purely methodological reasons. This 'method effect' appeared to be an empirically real property of the data for this factor, and was modeled, therefore, by freeing the indicated covariance of Surf91 with Surf87, which improved most aspects of model fit to acceptable levels ($\chi^2(1) = 3.9$; *RMSEA* = .099, *CI*s = .008 - .210; *TLI* = .93; *CFI* = .98; *SRMR* = .019). While the point-estimate for *RMSEA* remained above the threshold established in Chapter Five (see Table 5.2), its lower confidence interval (.008) advocated for good fit, as did all other indices reported above.

Freeing the covariance between items Surf91 and Surf87 was also observed to improve the congeneric fit of *Surface learning strategies* at Time 1 ($\chi^2(1) = 2.3$; *RMSEA* = .052; *CI*s = .000 - .143; *TLI* = .97; *CFI* = 1.00; *SRMR* = .015), while having negligible effect on the fit of the Time 1 multivariate measurement model ($\Delta CFI = .00$; $\Delta TLI = .00$).

Table 7.2

Time 2 congeneric model results, Time 2 (N = 297)

	CFA										
					RMSEA						
Scale (# items)	χ^2	*p*	*df*	Loading range	Value	Low *CI*	High *CI*	*CFI*	*TLI*	*SRMR*	Rho
Subject self-concept (5)	3.934	.559	5	.675-.898	.000	.000	.071	1.00	1	.008	.92
Honesty-trust. self-concept (6)	20.5	.015	9	.524-.888	.066	.027	.104	.98	.97	.030	.96
Performance structure (4)	9.04	.011	2	.568-.873	.109	.044	.185	.96	.89	.029	.79
Good teaching (8)	44.9	.001	20	.324-.828	.065	.039	.090	.97	.95	.034	.88
Usefulness of curriculum (4)	.367	.834	2	.774-.927	.000	.000	.067	1.00	1.01	.003	.92
Assessment quality	47.69	.000	14	.675-.775	.090	.063	.119	.94	.91	0.04	.80
Peer norms (5)	5.027	.413	5	.617-.881	.004	.000	.081	1.00	1	.016	.85
Surface learning strategies (4)	11.415	.033	2	.349-.842	.126	.062	.201	.96	.88	.036	.76
Justifiability of cheating (3)	.000	.992	1	.693-.813	.000	.000	.000	1.00	1.03	.000	.77
Self-reported cheating (3)	.038	.8454	1	.690-.960	.000	.000	.083	1.00	1.02	.010	.86

Note. χ^2 = chi-squared; *p* = significance level; *df* = degrees of freedom; *CI* = confidence interval; Rho = Rho reliability coefficient; highlights = index threshold violations.

7.3 Multivariate higher-order confirmatory factor analysis

An examination was next conducted of the validity of the multivariate measurement model at Time 2, which included ten first-order factors (see Table 7.2) and one second-order factor (i.e. *Teacher quality*). The Time 2 measurement model also included a free covariance parameter between items Surf91 and Surf87 of the *Surface learning strategies* measure. While the initial fit of the multivariate measurement model was acceptable ($\chi^2(1088) = 1796$; *RMSEA* = .047, *CI*s = .043 - .051, *pclose* = .915; *TLI* = .89; *CFI* = .90; *SRMR* = .068), a more accurate estimate was obtained by applying Boomsma and Herzog's (2013) small sample correction ($\chi^2(1088) = 1689$; *RMSEA* = .043, *CI*s = .039 - .047, *pclose* = .998; *TLI* = .91; *CFI* = .92; *SRMR* = .068; *SCF* = .940).

Good fit of the model indicated that all factors were well-defined with respect to each other. The variances of observed indicator items were adequately accounted for by the factors they were designed to measure, without the need for cross-loading single items onto multiple factors or for freeing inter-factor residual covariances. Mean factor loadings reported in Table 7.3 fell within a range of .644 - .892, which was comparable to the range of mean loadings at Time 1 (.611 - .867). The second-order factor, *Teacher quality*, once again demonstrated excellent within-model construct validity, explaining 66% of variance in *Good teaching* and 94% of the variance in *Assessment quality*, as compared to 67% and 84%, respectively, at Time 1.

Table 7.3
Comparison of factor loadings of the measurement model HCFA, Time 2 vs Time 1

		Time 2 HCFA		Time 1 HCFA	
	Scale (# items)	Loading range	Loading mean	Loading range	Loading mean
Person	Subject self-concept (5)	.674 - .897	.839	.714 - .847	.812
	Honesty-trust. self-concept (6)	.530 - .885	.706	.417 - .871	.653
Context	Performance structure (4)	.594 -.829	.689	.565 - .791	.642
	Usefulness of curriculum (4)	.791 - .919	.867	.720 - .896	.823
	Teacher quality	**.812 - .971**	**.892**	**.819 - .918**	**.867**
	Good teaching (8)	.330 - .820	.687	.401 - .759	.653
	Assessment quality (7)	.687 - .770	.724	.619 - .727	.662
	Peer norms (5)	.628 - .888	.730	.526 - .726	.642
Moral obligation	Justifiability of cheating (3)	.666 - .755	.724	.526 - .798	.687
Behavior	Surface learning strategies (4)	.413 - .758	.644	.354 - .808	.621
	Self-reported cheating (3)	.729 - .920	.814	.762 - .877	.826

Note. Measurement model fit: $\chi^2(1088) = 1689$; *RMSEA* = .043, *CIs* = .039 - .047, *pclose* = .998; *TLI* = .91; *CFI* = .92; *SRMR* = .068; *SCF* = .940

7.3.1 Correlation analysis

The correlation matrix for the measurement model presented in Table 7.3 was examined with respect to the strength and pattern of significant correlations at Time 2, compared to Time 1. Overall, this set of correlations remained consistent with the hypothesized structural model, and strongly resembled the Time 1 correlation matrix in most respects, including the signs, strengths and significance levels of correlation coefficients. Several differences were noted, however, including a shift in the pattern of correlations related to *Honesty-trustworthiness self-concept* away from *Self-reported cheating*, and in favor of *Teacher quality*.

The largest correlation between first-order factors was, as at Time 1, between *Good teaching* and *Assessment quality* (r = .788, p < .001), whereas the correlation observed between *Justifiability of cheating* and *Self-reported cheating* is lower at Time 2 (r = .723, p < .001) than it was at Time 1 (r = .766, p < .001). Overall, the absolute correlation coefficients between the Time 1 and 2 matrices differ by a mean value of $M_{\Delta|r|}$ = .004, with an absolute mean difference of $M_{|\Delta r|}$ = .072, and an absolute mean difference range of $R_{|\Delta r|}$= .000 - .225. Within this range, three Time 2 correlations differed from their Time 1 counterparts by more than $|\Delta r|$ = .200, including those of *Honesty-trustworthiness self-concept* with both *Assessment quality* and *Teacher quality*, which were larger at Time 2, by $|\Delta r|$ = .225 and $|\Delta r|$ = .219, respectively, and the correlation of *Honesty-trustworthiness self-concept* with *Self-reported cheating*, which was smaller at Time 2 by $|\Delta r|$ = .202. This pattern of differences suggested an overall re-orientation of *Honesty-trustworthiness self-concept* at Time 2 away from *Self-reported cheating*, where the correlation weakened substantially, and in favor of the first-order components of *Teacher quality* (*Good teaching* and *Assessment quality*), where correlations were stronger. This pattern was also observed with respect to the Time 2 structural model for female respondents, as discussed in section 7.7.

Three correlation coefficients that achieved significance at Time 1 were non-significant at Time 2. Correlations of *Performance goal structure* with *Peer norms* (r = -.013, NS) and *Self-reported cheating* (r = .090, NS); and of *Subject self-concept* with *Peer norms* (-.119, NS) at Time 2 achieved significance at Time 1 (r = .202, p < .01; r = .253, p < .001; and r = -.137, p = .05, respectively). Additionally, the correlations at Time 2 of *Performance goal structure* with *Usefulness of curriculum* (r = .155, p < .05) and *Subject self-concept* (r = .192, p < .01) were non-significant at Time 1 (r = .006 and r = -.022, NS, respectively).

Table 7.4

Higher-order CFA correlation matrix, Time 2 (N = 297)

	SUB	HON	PERF	GTEACH	CURUSE	ASSESS	PEER	SURF	CHJUST	CHEAT
SUB	1.00									
HON	.277***	1.00								
PERF	.192**	-.084	1.00							
GTEACH	.386***	.354***	.103	1.00						
CURUSE	.530***	.229***	.155*	.534***	1.00					
ASSESS	.461***	.424***	.123	.788***	.638***	1.00				
PEER	-.119	-.309***	-.013	-.384***	-.360***	-.436***	1.00			
SURF	-.286***	-.252***	.223**	-.259**	-.321***	-.310***	.178*	1.00		
CHJUST	-.192**	-.257***	.359***	-.315***	-.274***	-.377***	.471***	.633***	1.00	
CHEAT	-.261***	-.264***	.090	-.311***	-.339***	-.372***	.436***	.445***	.723***	1.00
TEACHER	.475***	.436***	.127	.812***	.657***	.971***	-.473***	-.319***	-.388***	-.383***

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance structure; TEACHER = Teacher quality (ASSESS = Assessment quality; GTEACH = Good teaching); CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; *p < .05, **p < .01, ***p < .001.

7.4 Invariance of the central measurement model

The invariance of factor measures between gender and grade-level groups at Time 2 was analyzed for significant underlying diversity within the Time 2 sample. Time 1 invariance analyses detected a number of small but significant differences in item-level mean values between grade-level and gender groups. Of these two groups, only gender differences were sufficient to violate the threshold for multivariate model invariance of $\Delta CFI = |.01|$ (Cheung & Rensvold, 2002). *CFI* was observed to diminish by -.014 at Time 1 when all observed variable intercepts in the full measurement model were held invariant across male and female groups, which prompted the analysis of gender-specific structural models. Gender differences in factor loadings and factor variances at Time 1 were, in contrast to observed variable intercepts, acceptably small, which met key invariance criteria for use of composite scores for estimating structural models (Holmes-Smith, 2012).

Multi-group invariance testing of the full measurement model could not be conducted at Time 2 due to sample size limitations. The Time 2 sample size (N = 297) included 115 males and 182 females; 147 Grade Nine respondents and 150 Grade Ten respondents. These numbers fell below the recommended size of groups in multi-group analysis of N = 200 (Meade et al., 2008). Attempting to also test the measurement model on two of these groups simultaneously would have involved approximately 376 free model parameters, reducing the *N:q* ratio to less than 1.

Multi-group measurement invariance is explored at Time 2 with differential item functioning (DIF) analysis (Grayson et al., 2000; Wang & Wang, 2012). The DIF model, in which all 49 observed indicator variables were regressed on gender (1 = male, 2 = female) and grade-level (1 = Grade Eight, 2 = Grade Nine), demonstrated good fit to the Time 2 data ($\chi^2(1088) = 1684$; *RMSEA* = .043, *CI*s = .039 - .047, *pclose* = .998; *TLI* = .90; *CFI* = .92; *SRMR* =

.063; $N:q$ = 1.1; SCF = .935). The degree of non-invariance across gender and grade-level groups was similar, overall, to that observed at Time 1.

As shown in Table 7.5, gender was associated with 27 significant item-level differences, and grade-level with 18 item-level differences. While fewer significant differences across groups were observed at Time 1 (21 for gender and 11 for grade-level), the number of highly significant effects, at $p > .01$, was identical (15 and 9, respectively). The absolute mean effects of gender $M_{|\beta|}$= .170 and grade-level $M_{|\beta|}$= .156 were also only slightly higher than those observed at Time 1 ($M_{|\beta|}$ = .146 and $M_{|\beta|}$ = .136, respectively). The number of beta coefficients greater than β = .200 was also nearly the same at both time points. Six differences related to gender had coefficients greater than .200 at Time 2 *versus* four at Time 1; and 1 difference related to grade-level had a coefficient greater than .200 at Time 2 *versus* none at Time 1.

This slight increase in non-invariance, especially across gender groups, justified the decision made at Time 1 to analyze gender-specific models, while a similarly low level of invariance related to grade-level, characterized by the same number of highly significant effects within a similar range of effects, provided no appreciably greater impetus for carrying out special analyses by grade-level. The need to analyze gender-specific models was, in fact, further emphasized at Time 2 by the concentration of item-level gender differences in items used to measure factors related to cheating (*Peer norms*, *Justifiability of cheating*, and *Self-reported cheating*). As at Time 1, gender differences were also concentrated in measures of *Subject self-concept* and *Performance goal structure*. Grade-level differences were found to be concentrated, by contrast, principally in items used to measure learning context variables, especially *Good teaching* and *Assessment quality*, but were absent from factors related to cheating. The analysis of gender-specific models conducted at Time 1 was, for these reasons, conducted again at Time 2.

Table 7.5

Differential item functioning analysis for gender and grade-level

Item	Gender	Grade	Item	Gender	Grade	Item	Gender	Grade	Item	Gender	Grade
SUB2	.261***	-.121*	PERF74	.242***	-.075	CURUSE64	.098	-.140*	SURF87	.111	.080
SUB3	.151**	-.028	PERF75	.154**	-.099	TRANS28	.025	-.127*	SURF88	-.119*	.126*
SUB5	.194***	-.147**	GTEACH18	.164**	-.137*	TRANS32	.087	-.087	SURF91	.035	.058
SUB13	.261***	-.078	GTEACH33	.093	-.154**	TRANS66	.033	-.187**	SURF97	.150*	.068
SUB15	.158**	-.006	GTEACH39	.123*	-.169**	AUTH44	.051	-.090	CHJUST79	.208***	.010
HON_1	.050	-.032	GTEACH50	.085	-.133*	AUTH60	.073	-.101	CHJUST86	.133*	.023
HON6	.034	.049	GTEACH62	.080	-.159**	AUTH71	.046	-.171**	CHJUST99	.223***	.055
HON8	.081	-.056	GTEACH67	.125*	-.276***	AUTH78	.040	-.120*	CHEAT84	.175**	.025
HON9	-.024	.023	GTEACH68	-.036	-.076	PEER24	.072	.043	CHEAT92	.182**	.103
HON10	-.096	.036	GTEACH77	.132*	-.194***	PEER31	.131*	.019	CHEAT95	.107	.059
HON11	-.092	-.073	CURUSE19	.121*	-.177**	PEER40	.151**	.066			
PERF61	.202***	-.100	CURUSE53	.118*	-.144*	PEER58	.167**	.006			
PERF69	.224***	-.127*	CURUSE56	.086	.086	PEER65	.139*	.045			

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; *p < .05, **p < .01, ***p < .001.

7.5 MIMIC modeling: Demographic effects (*N* = 267)

As at Time 1, a multiple-indicators multiple-causes, or 'MIMIC' analysis, was conducted by estimating two separate models. In the first of these models, all latent factors were regressed on all covariates measured in the study, including gender, grade-level, self-rated English proficiency, maternal educational attainment, and paternal educational attainment. In the second of these models, all latent factors were regressed on all possible two-way interactions between the aforementioned covariates. It is important to note that 30 of the 297 respondents at Time 2 did not provide complete demographic information. This rendered a sample of *N* = 267 for the purposes of MIMIC analysis. Both models were estimated with composite scores to maximize *N:q* ratios. It is also important to emphasize that non-invariance between gender and grade-level groups on the measures of various factors, as reported in Table 7.5, introduces ambiguity to the interpretation of the mean differences observed in this analysis (Byrne & van de Vijver, 2010).

Model fit. While satisfactory fit was achieved when all observed variables were used to estimate both the first model, involving covariates ($\chi^2(1288) = 1865$; *RMSEA* = .041; *CI*s = .037 - .045, *pclose* = 1.00; *TLI* = .90; *CFI* = .91; *SRMR* = .064; *N:q* = 1.2; *SCF* = .925), and the second model, involving two-way interactions ($\chi^2(1488) = 2052$; *RMSEA* = .038; *CI*s = .034 - .042, *pclose* = 1.00; *TLI* = .90; *CFI* = .92; *SRMR* = .060; *N:q* = .97; *SCF* = .917), these models involved very low *N:q* ratios, 1.2 and 0.97, respectively. The second model could not, in fact, be estimated with certainty due to a non-positive definite first-order derivative matrix caused by the *N:q* ratio of less than one. The use of composite scores allowed both models to be estimated by increasing their *N:q* ratios.

Estimated with composite scores, the first MIMIC model, which included all five covariates, achieved excellent fit ($\chi^2(13) = 16$; *RMSEA* = .027; *CI*s = .000 - .070, *pclose* = .771; *TLI* = .98; *CFI* = 1.00; *SRMR* = .013; *N:q* = 2.6). The *N:q* ratio of 2.6 was approximately twice that of the same model estimated with observed indicator variables (above). Effect size differences between the two methods of estimation were minimal. The mean effect size of the composite score model was $M_{\Delta|\beta|}$ = .0002 larger, differing in absolute terms by $M_{|\Delta\beta|}$ = .012, with a range of absolute differences of $R_{|\Delta\beta|}$ = .000 - .039. All effects that achieved significance in the MIMIC model estimated with observed indicator variables remained significant in the model estimated with composite scores, with one exception. The effect of grade-level on *Performance goal structure* fell from just-significant in the model estimated with observed indicator variables (β = .130, p = .049) to non-significant in the composite score model (β = .123, p = .067).

The second MIMIC model, which included all ten two-way interaction variables, also fit the data well when estimated with composite scores ($\chi^2(18)$ = 27; *RMSEA* = .043; *CI*s = .000 - .075, *pclose* = .600; *TLI* = .91; *CFI* = .99; *SRMR* = .010; *N:q* = 1.8). The *N:q* ratio of 1.8 was, again, nearly twice that of the model estimated with observed indicator variables.

Differences between the model estimated with observed indicator variables and the composite model could not be analyzed because the *N:q* ratio was less than one in the former instance. However, differences in effect size observed between these estimation methods have been consistently small in analyses of five preceding models in the present study. Effect sizes in these five models tended to be larger when estimated with composite scores, by $\Delta\beta$ = .01, on average, with a mean absolute difference of $M_{|\Delta\beta|}$ = .03. It was reasonable to expect, therefore, that effects obtained by estimating the Time 2 MIMIC model with composite scores would differ to a similarly small extent from what its effects would have been if estimated with observed indicator variables.

*Results***.** Among the 17 significant mean differences found in the MIMIC analysis at Time 2 (see Tables 7.6a and 7.6b), more than half were related either to gender (6 effects), or the two-way interaction between English proficiency and gender (3 effects). The pattern of gender effects at Time 2 was, in fact, very similar to that at Time 1, with the single exception of the effect of gender on *Teacher quality*, which achieved significance only at Time 2 ($\beta = .141$, $p < .05$). This indicated that males tended to judge their teachers more favorably than females at Time 2. Male respondents also reported higher *Subject self-concept* ($\beta = .252$, $p < .001$), a stronger sense of *Performance goal structure* ($\beta = .123$, $p < .05$), a greater belief that *Peer norms* favored cheating ($\beta = .218$, $p < .001$), a greater likelihood to believe in the *Justifiability of cheating* ($\beta = .182$, $p < .01$), and greater engagement in *Self-reported cheating* ($\beta = .136$, $p < .05$), all within the context of Science class.

The fact that all factor-level gender differences at both time points were higher for male respondents is suggestive of a method effect in which males were simply more likely to give higher answers than females. Differences at the item level (see DIF analyses in sections 6.4.1 and 7.4) are, however, not so consistent that they can be explained as a blanket tendency among males to simply circle higher Likert values. The measure *Peer norms*, for instance, included four reversed items (Peer24, Peer31, Peer40 and Peer60; see item wording in Table 4.9), of which male respondents gave significantly lower responses at Time 1 for Peer31 (see Table 6.10), and at Time 2 for Peer31, Peer40, and Peer65 (see Table 7.5). Males also gave significantly higher answers at both time points for the only non-reverse item in the *Peer norms* measure, Peer58, which suggests that they were, in fact, more likely than their female counterparts to believe their peers to be more amenable to cheating. Mean differences detected in *Subject self-concept* were similarly corroborated by a reversed item (Sub15; see Item 2 in Section B of Appendix B), for which males gave significantly lower responses than females at both time points. Unambiguous interpretations of these mean differences cannot be made on the basis of this analysis, however, due both to gender non-invariance in the measurement model, and to the effects of the interaction between gender and English proficiency that are next discussed.

While, at Time 1, English proficiency was found to predict significant differences in *Subject self-concept* ($\beta = .102$, $p < .05$), *Performance goal structure* ($\beta = -.110$, $p < .05$), and *Surface learning strategies* ($\beta = -.149$, $p < .01$), none of these differences persisted at Time 2. Differences predicted at Time 2 by the two-way interaction between gender and English proficiency in *Surface learning strategies* ($\beta = -.167$, $p < .05$), *Justifiability of cheating* ($\beta = -.250$, $p < .01$), and *Self-reported cheating* ($\beta = -.185$, $p < .05$) suggest, instead, that these were more prevalent among male respondents whose self-rated English language proficiency was lower. Only the first of these three effects was also found at Time 1. Additionally, two-way interactions between English proficiency and both grade-level and paternal educational attainment indicate that the effect of lower self-rated English proficiency on *Self-reported cheating* was higher among respondents at Grade Ten than at Grade Nine ($\beta = -.304$, $p < .05$), and among respondents whose paternal educational attainment was lower ($\beta = -.187$, $p < .01$).

In comparison to the five significant grade-level differences found at Time 1, only the effect of grade-level on *Teacher quality* ($\beta = .178$, $p < .01$) remained significant at Time 2. This difference indicates that Grade Nine students rated their Science teachers more highly than Grade Ten students. Again, however, significant differences related to several two-way interactions variables involving grade-level, noted below, qualify interpretation of its direct effects.

The only significant difference associated with maternal educational attainment at Time 1, which was observed in *Subject self-concept* ($\beta = .108$, $p < .05$), failed to persist at Time 2. Only its effect on *Performance goal structure* achieved significance at Time 2 ($\beta = -.165$, $p < .05$), indicating that respondents whose mothers had reached a higher level of institutional education tended to be less likely to perceive performance goal structures in Science class. No two-way interactions involving maternal educational attainment achieved significance at Time 2.

The effect of paternal educational attainment was more pervasive at Time 2 than at Time 1. The only significant difference predicted by paternal educational attainment at Time 1, also in *Subject self-concept* ($\beta = .111$, $p < .05$), was stronger at Time 2 ($\beta = .196$, $p < .01$). Two additional differences predicted by paternal educational attainment at Time 2 included *Usefulness of curriculum* ($\beta = .138$, $p < .05$), and *Self-reported cheating* ($\beta = -.213$, $p < .01$). These three differences suggest, in sum, that respondents whose fathers had more formal education tended (1) to have higher self-concept in relation to Science, (2) to perceive the curriculum in Science class as more useful, and (3) to cheat less. The effect of the two-way interaction between paternal educational attainment and grade-level on *Self-reported cheating* ($\beta = -.318$, $p < .05$) suggests, however, that among students with lower parental educational attainment, cheating was more likely in Grade Ten than in Grade Nine. Respondents indicating lower paternal educational attainment were also, as noted above, more likely to report cheating when they also had lower self-rated English proficiency.

In summary, the pattern of demographic effects observed at Time 2 marked a decline, over Time 2, in the prominence of grade-level effects in addition to a slight increase in effects associated with gender and paternal educational attainment. As at Time 1, males at Time 2 tended to report higher self-concept in relation to Science class and more favorable judgments of the teacher, in conjunction with more prevalent pro-cheating attitudes and more disintegrity behavior. This conjunction of positive perceptions and dishonest attitudes and behaviors, which is surprising in the light of a large amount of published literature in addition to results in the present study that point to the opposite pattern, might reflect that males were also more likely, at Times 1 and 2, to perceive a performance goal structure. Performance goals tend to convey that good grades and favorable peer-comparisons are more important relative to actual learning, and may, as such, encourage more cheating, even among students who view a class context favorably.

Table 7.6a

MIMIC results: standardized beta coefficients for covariates, Time 2 (N = 267)

	Gen	Gra	Eng	Mom	Dad
Person					
Subject self-concept	.230***	-.078	.024	.055	.196**
Honesty-trust. self-concept	.016	.004	.076	.110	.047
Learning context					
Performance goal structure	.248***	-.123	-.087	-.165*	-.043
Teacher	.141*	-.178**	.044	-.085	.098
Usefulness of curriculum	.111	-.119	.062	-.111	.138*
Peer cheating norms	.169**	-.055	.038	.102	-.103
Moral obligation					
Justifiability of cheating	.227**	.046	-.056	.017	.150
Behavior					
Surface learning strategies	.081	.095	-.055	-.107	-.143
Self-reported cheating	.183**	.126	.041	-.006	-.213**

Note. Model fit: ($\chi^2(13) = 16$; *RMSEA* = .027; *CI*s = .000 - .070, *pclose* = .771; *TLI* = .98; *CFI* = 1.00; *SRMR* = .013). Gen = Gender, Gra = Grade-level, Eng = English proficiency, Mom = Maternal educational attainment, Dad = Paternal educational attainment, **p* < .05, ***p* < .01, ****p* < .001.

Table 7.6b

MIMIC results: standardized beta coefficients for two-way interaction variables, Time 2 (N = 267)

	GenXGra	GenXEng	GenXMom	GenXDad	GraXEng	GraXMom	GraXDad	EngXMom	EngXDad	MomXDad
Person										
Subject self-concept	-.112	-.123	-.070	.193	.167	.129	.071	-.111	.009	.052
Honesty-trust. self-concept	-.068	-.001	-.021	.057	.213	.172	-.018	.191	-.046	-.119
Learning context										
Performance goal structure	-.126	-.138	-.195	.059	.146	.043	-.092	-.110	-.061	-.001
Teacher	-.013	.075	-.150	-.014	-.039	.065	.143	.097	-.032	-.097
Usefulness of curriculum	-.082	.025	-.207	-.058	-.032	.101	.220	.080	.022	-.038
Peer cheating norms	-.093	-.086	-.067	-.084	.008	.197	-.010	.154	-.046	-.005
Moral obligation										
Justifiability of cheating	-.132	-.250**	-.075	-.031	-.156	.109	-.200	.115	-.079	.015
Behavior										
Surface learning strategies	-.069	-.167*	.057	.027	.083	-.154	-.197	-.174	-.094	.077
Self-reported cheating	-.067	-.185*	-.034	.044	-.304*	.028	-.318*	.205	-.187**	.022

Note. Model fit: ($\chi^2(18) = 27$; *RMSEA* = .043; *CIs* = .000 - .075, *pclose* = .600; *TLI* = .91; *CFI* = .99; *SRMR* = .010). Gen = Gender, Gra = Grade-level, Eng = English proficiency, Mom = Maternal educational attainment, Dad = Paternal educational attainment, **p* < .05, ***p* < .01, ****p* < .001.

7.6 Male sample structural model (N = 115)

Due to the small size of the male and female components of the Time 2 sample (N = 115 and N = 182, respectively), neither could be estimated with observed indicator variables, because of $N:q$ ratios less than 1. Both gender-specific models were estimated, for this reason, with composite scores.

7.6.1 Measurement model analysis: Time 2 male sample data

Assessments of congeneric model fit, reported in Table T1 of Appendix T were generally satisfactory for the male component of the Time 2 sample. Indications of weakness were, however, found with respect to *Good teaching* and *Assessment quality*. The model for *Good teaching* included a factor loading of λ = .109 (item gteach68), which fell below the threshold of .300. This observation was not of critical concern, however, in view of the otherwise good fit of the congeneric model for *Good teaching* $\chi^2(20)$ = 31.8, p = .046; *RMSEA* = .072, *CI*s = .010 - .117; *pclose* = .210; *TLI* = .94; *CFI* = .96; *SRMR* = .044; Rho = .86).

The fit of the congeneric model for *Assessment quality* ($\chi^2(14)$ = 30.6, p = .01; *RMSEA* = .102, *CI*s = .052 - .151; *pclose* = .045; *TLI* = .90; *CFI* = .93; *SRMR* = .046; Rho = .91) was similar to that observed in the co-ed Time 2 sample (see Table 7.2). Weakness was observed in both *RMSEA* (.102) and *CFI* (.93). These estimates were, however, still reasonably close to the desired levels of .080 and .95, respectively. *RMSEA* values as high as $\leq$ .10 have, for instance, been considered acceptable (MacCallum et al., 1996), as have *CFI* values as low as .90 (Marsh et al., 2004). These signs of weakness were also mitigated by appropriate *TLI* and *SRMR* values, robust factor loadings (R_λ = .70 - .80), and excellent reliability (.91).

The multivariate measurement model was next estimated using composite scores, and achieved excellent fit ($\chi^2(8)$ = 9.7, p = .29; *RMSEA* = .043, *CI*s = .000 - .123, *pclose* = .484; *TLI* = .98; *CFI* = 1.00; *SRMR* = .016; $N:q$ = 2). A small negative residual on *Assessment quality* (-.038), or a 'Heywood case' (Byrne, 2012), was addressed by setting its residual variance equal to .00001 with syntax prescribed by Muthén and Muthén (2014). The correlation matrix for the male measurement model is presented in Appendix U.

7.6.2 Structural analysis: Time 2 male sample data

The revised hypothetical structural model, or Model 3 (see Figure 6.1), also demonstrated excellent fit to the male data ($\chi^2(15) = 22.8$, $p = .09$; *RMSEA* = .067, *CI*s = .000 - .120, *pclose* = .278; *TLI* = .94; *CFI* = .98; *SRMR* = .042; *N:q* = 2). Variance explained at Time 2 in the two outcome variables, *Self-reported cheating* (81%) and *Surface learning strategies* (71%), was greater than that observed in the Time 1 male sample model by $\Delta R^2 = 6\%$ and $\Delta R^2 = 13\%$, respectively. The amount of variance explained at Time 2 in *Justifiability of cheating* (36%) was less, by contrast, than at Time 1 ($\Delta R^2 = -18\%$).

The pattern of structural effects for males at Time 2 retained most of the key characteristics that were observed at Time 1. The strongest effects on *Surface learning strategies* and *Self-reported cheating* were exerted by *Justifiability of cheating* ($\beta = .754$, $p < .001$ and $\beta = .776$, $p < .001$, respectively), which mediated most of the effects of class context. The role of *Peer norms* as a complete mediator of *Teacher quality* among males ($\beta = -.450$, $p < .001$) was also observed at Time 1, and *Usefulness of curriculum* failed, again, to exert any significant predictive effects in the model. This overall pattern of effects continued, therefore, to indicate that respondents who rated the quality of their Science teacher as lower were more likely (1) to believe that cheating was acceptable among their peers, and (2) to believe, in turn, that cheating in Science class was more justifiable.

The role of *Performance goal structure* at Time 1 differed in several ways among males, as compared to Time 2. *Performance goal structure* emerged as the strongest predictor of *Justifiability of cheating* among males at Time 2 ($\beta = .452$, $p = .001$), and exerted a significant direct effect on *Self-reported cheating* (β = .297, p = .05). These differences appeared, however, to be partly attributable to suppression effects, as discussed in more detail below. Another prominent difference in the male sample model at Time 2 was the significant effect of *Subject self-concept* on *Performance goal structure* (β = .264, p = .05). This parameter was non-significant in all Time 1 models.

The tendency, observed among males at Time 1, of personological variables to exert direct effects on outcome variables, thus bypassing complete mediation by *Peer norms* and *Justifiability of cheating*, was weaker among male respondents at Time 2. While the effect of *Subject self-concept* on *Surface learning strategies* remained significant at Time 2 (β = -.294, $p < .05$), its effect on *Self-reported cheating* did not (β = -.045, NS). The effect of *Honesty-trustworthiness self-concept* on *Self-reported cheating* also did not remain significant among Time 2 males (β = -.040, NS), whereas its effect on *Justifiability of cheating* did (β = -.272, $p < .05$).

Suppression effects. The effect of *Performance goal structure* on *Justifiability of cheating* (β = .452, $p < .001$) was nearly 50% larger among male respondents at Time 2 than at Time 1, as well as being nearly two-thirds larger than its corresponding bivariate correlation at Time 2 (r = .287; see Appendix U). This appears to have been the result of third variable suppression (Tzelgov & Henik, 1991). Suppressor variables increase the magnitudes of independent variable effects on dependent variables by explaining, or 'suppressing', "outcome-irrelevant variance in" the predictor (Pandey & Elliot, 2010, p. 29; see also Conger, 1974; Horst, 1941). Suppressor variables are often uncorrelated with the dependent variable, and are thus more common in models that, like Model 3, have large numbers of non-significant parameters (Pandey & Elliott, 2010). Other predictors of *Justifiability of cheating* in the Model 3 appear, in other words, to have 'suppressed' variance in *Performance goal structure* that was extraneous to its relationship with *Justifiability of cheating*, thereby inflating the regression weight. To test the hypothesis that the path coefficient in question was being suppressed, the paths from all other predictors of *Justifiability of cheating* were constrained to zero. When this was done, the effect of *Performance goal structure* fell perfectly in line with its corresponding bivariate correlation ($\beta = .286$, $p = .001$). As the beta paths from its co-predictors of *Justifiability of cheating* were then freed, one-by-one, the effect in question increased steadily back to the inflated value of $\beta = .452$. No single suppressor variable was identified.

Another substantial suppression effect was observed in the beta path from *Performance goal structure* to *Self-reported cheating* ($\beta = -.297$, $p < .05$), which was nearly twice the size of its corresponding bivariate correlation ($r = -.159$). Small suppression effects in the male sample model were also observed in the beta paths from *Subject self-concept* to *Surface learning strategies* ($\beta = -.294$, $p < .05$), and *Peer norms* to *Justifiability of cheating* ($\beta = .335$, $p < .05$), as indicated by the fact that they were larger than their corresponding bivariate correlations, $r = -.290$ and $r = .324$, respectively.

A prominent suppression effect was also observed with respect to the effect of *Teacher quality* on *Surface learning strategies* ($\beta = .344$, $p < .05$; see the dot-dashed line in Figure 7.1). This effect size was much larger than its corresponding bivariate correlations ($r = -.001$), and of opposite sign, which is indicative of negative confounding in the context of 'inconsistent mediation' (Davis, 1985). The inflation of these beta coefficients appeared, in other words, to involve a suppression effect within the context of 'inconsistent mediation', in which a mediated effect carries the opposite sign of the direct effect (Beslow & Day, 1980; Conger, 1974; MacKinnon, Krull, & Lockwood, 2000). While the beta coefficient in question could be interpreted to mean that *Teacher quality* positively predicts *Surface learning strategies* when controlling for other variables, it is not clear that this is warranted, nor that this effect is anything more than an artifact of statistical distortion in a low-power analysis.

Results of the male sample model at Time 2 were generally questionable due both to a small sample size (*N* = 115) and to the emergence of several prominent suppression effects. It is instructive, however, that key aspects of the overall pattern of results among males at Time 1 persisted at Time 2, such as the mediating roles of *Justifiability of cheating* and *Peer norms*, and the direct effects of *Subject self-concept* on contextual and outcome variables.

Table 7.7

Model 3: Standardized beta coefficients estimated with composite scores, Time 2 male data (N = 115)

	Predictors							
	Grade	Sub	Hon	Perf	Curuse	Teacher	Peer	Chjust
Sub	-.036							
Hon	.002	---						
Perf	.199*	.264*	---					
Curuse	-.113	.528***	---	---				
Teacher	-.113	.460***	---	---	---			
Peer	-.262**	---	-.122	-.174	-.167	-.470***		
Chjust	.056	-.077	-.271*	.452***	-.003	-.046	.335*	
Surf	.206*	-.294*	---	.109	-.050	† .344*	---	.754***
Cheat	.035	-.045	-.040	-.297*	-.088	.052	.189	.776***

Note. † denotes suppression in the context of inconsistent mediation, as discussed above. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level.

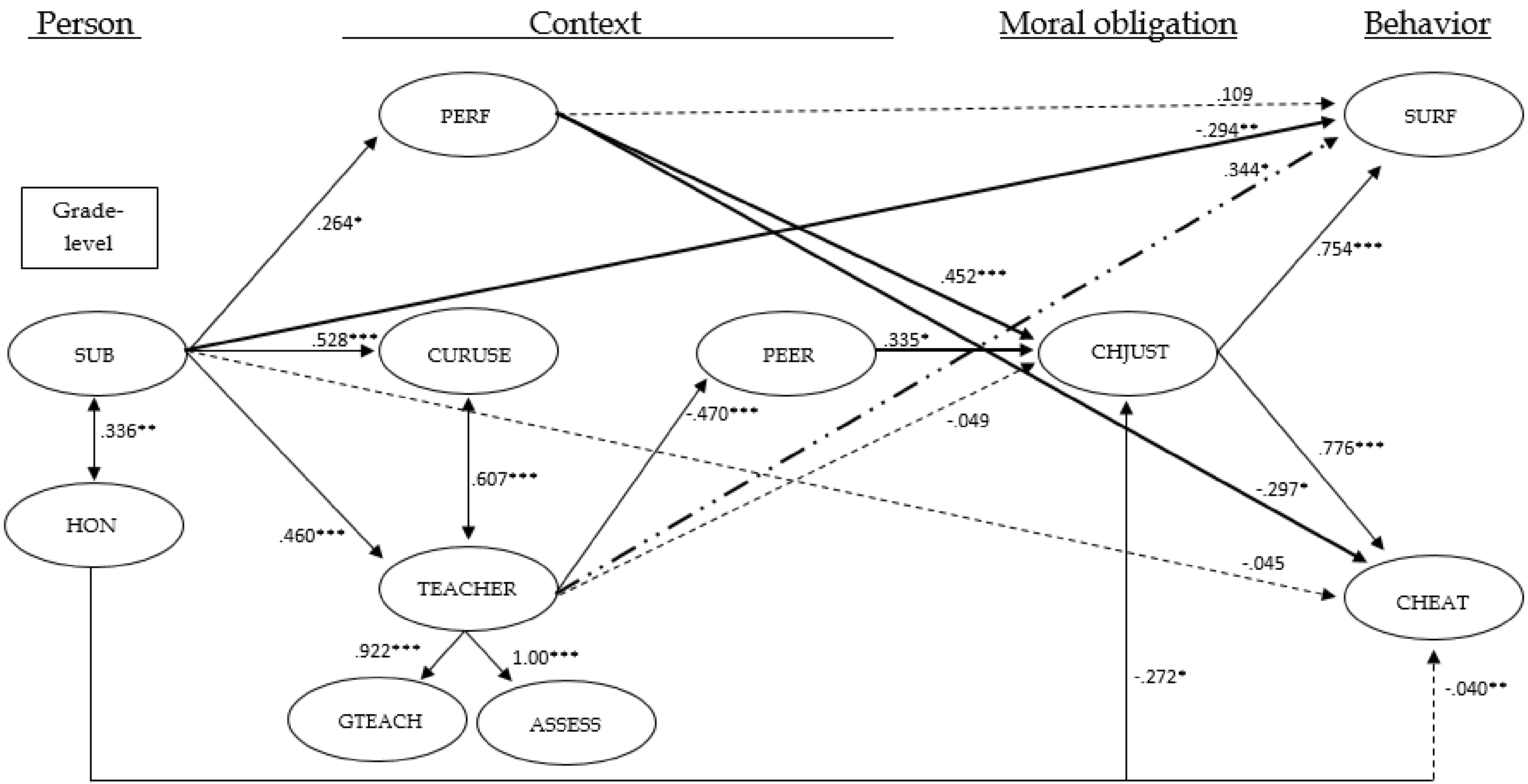

Figure 7.1. Male sample results for Model 3, Time 2 (N = 115). Male sample MODEL with COMPOSITES $\chi^2(15) = 21$; *RMSEA* = .060, *CI*s = .000 - .114, *pclose* = .351; *TLI* = .93; *CFI* = .98; *SRMR* = .050; *N:q* = 2. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; - - - - paths are significant in the male sample model estimated at Time1 (Figure 6.2) ; —··—··—·· suppression effects in the context of inconsistent mediation; Bold arrows are suppression effects.

7.7 Female sample structural model (*N* = 182)

The results of estimating Model 3 with the female component (*N* = 182) of the Time 2 sample are presented in the present section. Composite scores were used for the following analyses because estimating Model 3 with observed indicator variables involved more free parameters than there were subjects in the sample. *N:q* ratios of less than one cause non-positive definite model matrices, which invalidate results.

7.7.1 Measurement model analysis: Time 2 female sample data

Congeneric model analyses for the female component of the Time 2 sample (see Table T2 of Appendix T) identified weakness in four factors. These factors, *Performance goal structure*, *Assessment quality*, *Surface learning strategies*, and *Self-reported cheating*, all had high *RMSEA* point estimates (.165, .086, .112, and .083, respectively). The lower-bound *RMSEA* confidence intervals for the latter three fell, however, below the threshold of .050. These three factors also met desired thresholds for *CFI* and *TLI*, with the minor exception of an estimate of *CFI* = .94 for *Assessment quality*, which was counterbalanced by a satisfactory *TLI* value (.92). The fit of *Performance goal structure* was a more significant concern, as discussed below.

Self-reported cheating had, additionally, an *SRMR* value of .113, which exceeded the threshold of $\leq$ .080. An exploration of the residuals for this model identified large discrepancies in their size across indicator variables. Two of its three items, Cheat84 and Cheat95, had significantly larger residuals (δ = .546 and δ = .499, respectively) than the third item, Cheat92 (δ = .008). These differences were also reflected in the items' standard factor score coefficients (.01, .01, and .98, respectively), such that the operational meaning of *Self-reported cheating* among females at Time 2 was defined by item Cheat92 (I sometimes cheat on my Science class work, this year). Standardized factor scores for females at Time 1 were, by contrast, more equitable (Cheat84 = .21; Cheat92 = .50; Cheat95= .29).

Performance goal structure demonstrated substantial weakness, including a large *RMSEA* point estimate (.165) and confidence intervals (.083 - .261), low *TLI* (.76), and low *CFI* (.92). While this factor performed well among females at Time 1, modification indices associated with the female sample model at Time 2 indicated the need for two error covariances, with equivalent Lagrange multiplier values (12.1): Perf69 with Perf74, and Perf61 with Perf75. Exclusion of these statistical relationships was the most likely cause of the low value of *TLI* (.76), which includes a penalty for lack of parsimony (Byrne, 2012). Specifying either covariance parameter produces nearly perfect fit with a *TLI* value of 1.03. Neither modification was made, however, as they were likely idiosyncratic to the female component at Time 2, exacerbated by small sample size, and would stabilize in the larger sample.

The multivariate measurement model for female respondents at Time 2, estimated with composite scores, was excellent ($\chi^2(7) = 5.6$, $p = .58$; *RMSEA* = .000, *CI*s = .000 - .080, *pclose* = .810; *TLI* = 1.02; *CFI* = 1.00; *SRMR* = .011; *N:q* = 3.1). The correlation matrix for the measurement model, presented in Table V1 of Appendix V, was consistent with the hypothesized variable relationships in Model 3, in both strength and sign.

7.7.2 Structural analysis: Time 2 female sample data

The revised hypothesized structural model, or Model 3, also estimated with composite scores, was found to be weak on two key indices ($\chi^2(14) = 40$, $p < .001$; *RMSEA* = .101, *CI*s = .065 - .138, *pclose* = .012; *TLI* = .80; *CFI* = .95; *SRMR* = .054; *N:q* = 3). The *TLI* value of .80 as well as the *RMSEA* point-estimate of .101 and accompanying confidence intervals of .083 - .261 fell considerably wide of desired values. An exploration of modification indices for Model 3 found that the largest Lagrange multiplier among those that were consistent with the broader set of structural hypotheses in the overall model (17.7) advocated for a path from *Honesty-trustworthiness self-concept* to *Teacher quality*, which appeared to indicate that respondents who viewed themselves as more honest and trustworthy also tended to appraise their Science teachers more positively. This is consistent with correlations observed between *Honesty-trustworthiness self-concept* and *Teacher quality* among male (r = .393) and female respondents (r = .458) at Time 2 (see Table V1 of Appendix V), as well as female respondents at Time 1 (r = .328) (see Table K1 of Appendix K). Similar correlations between these two variables were also found in the co-ed samples at Time 1 (r = .217), Time 2 (r = .436), and in the Pilot Study (r = .482). A relationship between *Honesty-trustworthiness self-concept* and perceptions of teacher quality has also been noted elsewhere (Hay, 2000; Martin, Marsh, McInerney, & Green, 2006). Martin et al. (2006) found, for instance, a correlation of r = .390 between *Honesty-trustworthiness self-concept* and high school students' appraisals of their teachers.

A theoretical rationale for these observations is that they reflect cynicism, which entails believing that others are selfish and dishonest, or being "unable to take what someone says at face-value" (Mills & Keil, 2005, p. 385). One who sees him- or herself as less honest and trustworthy may, as these correlations suggest, also be less likely to trust others. Feeling doubtful about teachers would appear tantamount to holding them in lower esteem with respect to such things as subject knowledge, equitability, and work ethic, which should result in lower appraisals of their overall quality, and *vice versa*.

Cynicism is associated with factors such as stress, exhaustion, and burnout in academic settings, which have been observed to be (1) more common among females than males (Galbraith & Merrill, 2012; Simon, Carr, Mccullough, et al., 2004), and (2) to increase with age during adolescence (Galbraith & Merrill, 2012). Both of these observations from the literature on cynicism would be consistent with finding that cynicism had become more prevalent among female respondents at Time 2. This interpretation was further corroborated by Time 2 DIF analyses (see Table 7.5), which revealed no gender differences in items used to measure *Honesty-trustworthiness self-concept* or *Assessment quality*, whereas four of the eight items used to measure *Good teaching* were significantly affected by gender (R_{β} = .123 - .164).

Males gave Science teachers higher ratings with respect to feedback (Gteach18 and Gteach39), making Science interesting (Gteach67), and trying to get the best out of students (Gteach77). MIMIC analyses presented in Table 7.6a also found no mean difference between genders on *Honesty-trustworthiness self-concept* (.016, NS), but a significant mean difference in *Teacher quality* (β = .141, p < .05). This suggests, again, that male respondents tended to view their Science teachers more positively than female respondents.

7.7.3 Analysis of Model 4

The observed theoretical and empirical support for a regression path from *Honesty-trustworthiness self-concept* to *Teacher quality* led to the decision to free that parameter in the model, resulting in 'Model 4'. It must be acknowledged that freeing this parameter is a substantial *post-hoc* modification to the pattern of structural paths in the hypothesized model. Lacking the opportunity to cross-validate this modification on a separate sample, the risk that it capitalizes on idiosyncrasies of the Time 2 data set must be noted.

Model 4 was found to fit Time 2 female data well ($\chi^2(13)$ = 21, p = .07; *RMSEA* = .058, = .000 - .102, *pclose* = .344; *TLI* = .93; *CFI* = .98; *SRMR* = .030; *N:q* = 3), explaining more than half of the variance in *Self-reported cheating* (53%), *Surface learning strategies* (58%), and *Justifiability of cheating* (57%). These amounts of variance explained differed from the Time 1 female sample model by ΔR^2 = -10%, ΔR^2 = 22% and ΔR^2 = -6%, respectively. Beta coefficients for this model are presented in Table 7.8.

The effect of *Honesty-trustworthiness self-concept* on *Teacher quality* (β = .311, p < .001) was highly significant and consistent with its corresponding bivariate correlation (r = .412). The pattern of effects presented in Figure 7.2 also differed from the Time 1 female sample in that neither *Honesty-trustworthiness self-concept* nor *Subject self-concept* exerted a significant effect on either outcome variable. Both appeared, instead, to be mediated by intervening variables such as *Teacher quality* and *Justifiability of cheating*.

Table 7.8

Model 4: Standardized beta coefficients estimated with composite scores, Time 2 female data (N = 182)

	Predictors							
N = 182	Grade	Sub	Hon	Perf	Curuse	Teacher	Peer	Chjust
Sub	.016							
Hon	-.150*	---						
Perf	-.093	.011	---					
Curuse	-.102	.529***	---	---				
Teacher	-.224**	.425***	.311***	---	---			
Peer	-.054	---	-.174	.102	-.076	-.270		
Chjust	.020	-.172	.192*	.237**	.033	-.417**	.414***	
Surf	-.026	-.181	---	.129	-.048	-.346*	---	.321**
Cheat	.057	-.093	-.158	-.002	-.044	.099	-.047	-.706***

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; *$p < .05$, **$p < .01$, ***$p < .001$.

Another aspect of the female sample model at Time 2 that was not observed at Time 1 was the lack of any significant predictor for *Peer norms*. Both *Performance goal structure* and *Teacher quality* exerted unmediated, direct effects on *Justifiability of cheating* ($\beta = .237$, $p < .01$ and $\beta = -.417$, $p < .01$, respectively). So while *Peer norms*, itself, also predicted *Justifiability of cheating* with a magnitude similar to that in all previous models ($\beta = .414$, $p < .001$), it did not, unlike previous models, play a significant mediating role with respect to *Teacher quality* and *Performance goal structure*.

Teacher quality also directly predicted *Surface learning strategies* ($\beta = -.346$, $p < .05$), with a magnitude greater than that of *Justifiability of cheating* ($\beta = .321$, $p < .01$), thus indicating that its effect on cheating operated independently of moral obligation. The direct effect of *Teacher quality* on *Surface learning strategies* also supports the interpretation of the effect of *Honesty-trustworthiness self-concept* on *Teacher quality* as representing cynicism. Cynicism, which involves doubting the integrity of others (Mills & Keil, 2005), has long been held to "lead directly to surface learning" (Biggs, 1991, p. 219). The hypothesis that teacher evaluations were affected by cynicism is, therefore, consistent with the observed direct, inverse relationship between *Teacher quality* and *Surface learning strategies*. The mediating role played by *Justifiability of cheating* in previous models remained prominent, however, especially with respect to *Self-reported cheating* ($\beta = .706$, $p < .001$). All beta coefficients in the Time 2 female sample model were consistent with their corresponding bivariate correlations. No suppression effects were observed.

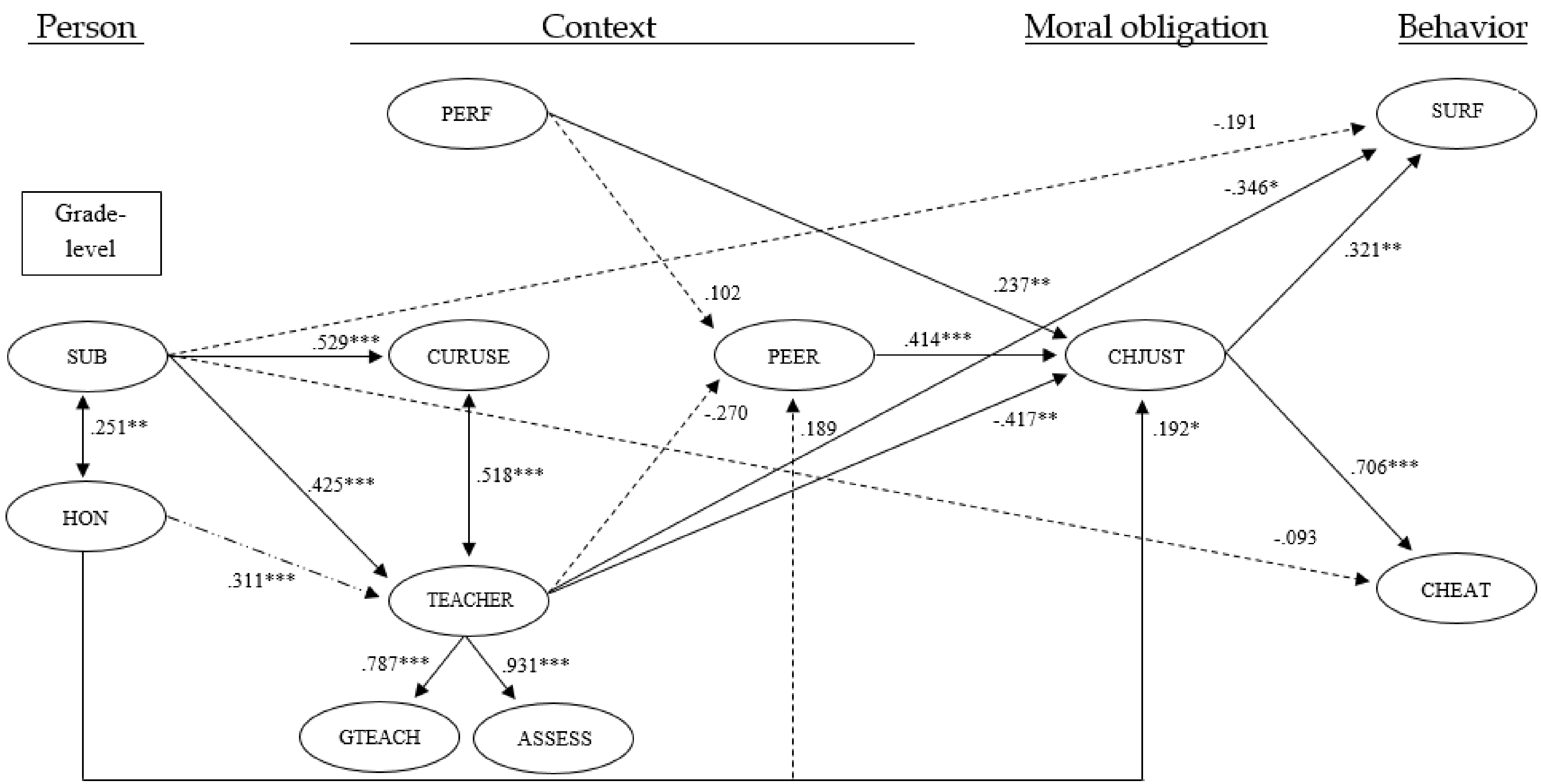

Figure 7.2. Female sample results for Model 4, Time 1 (*N* = 186). $\chi^2(13) = 21$; *RMSEA* = .058, = .000 - .102, *pclose* = .344; *TLI* = .93; *CFI* = .98; *SRMR* = .030; *N:q* = 3. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; *p < .05, **p < .01, ***p < .001. - - - - - paths were significant in the female sample model estimated at Time 1 (see Figure 6.3); —··—··— new path, added to the hypothesized model for females at Time 2.

7.8 Co-ed sample structural model: Time 2 (*N* = 297)

Detailed analyses of the measurement properties of the hypothesized model, reported in sections 7.1 – 7.5, demonstrated that the multivariate measurement model, and the individual factors it comprises, were sufficiently valid and reliable at Time 2 to proceed with analysis of the hypothesized structural model.

Results of invariance reported in section 7.4 cross-validated those conducted at Time 1 (see section 6.4), thus justifying concerns over gender invariance. While differences in the gender-specific structural models presented in sections 7.6 and 7.7 appear to be more pronounced than those observed at Time 1 (Sections 6.6 and 6.7), most key characteristics observed in Time 1 structural models remained intact, such as the prominent mediating roles of *Justifiability of cheating* and *Peer norms*. A major difference between the gender-specific models was the need, in the female sample model, to free a regression parameter from *Honesty-trustworthiness self-concept* to *Teacher quality*. The addition of this parameter resulted in Model 4 (see Figure 7.2). The results of Model 4 estimated with the full, co-ed sample for Time 2 (*N* = 297) are reported below.

Analysis of Model 4

Estimating Model 4 with observed indicator variables produced acceptable fit to the co-ed data set at Time 2 ($\chi^2(1173) = 1821$; *RMSEA* = .043, *CI*s = .039 - .047, *pclose* = .999; *TLI* = .90; *CFI* = .91; *SRMR* = .069; *N:q* = 1.5; *SCF* = .937), explaining 57% of the variance in *Self-reported cheating*, 51% of the variance in *Surface learning strategies*, and 43% of the variance in *Justifiability of cheating*, which differed from Time 1 estimates by $\Delta R^2 = -.10$, $\Delta R^2 = .10$, and $\Delta R^2 = -.04$, respectively. Beta coefficients for this model are presented in Table 7.9, and the full model output is provided in Appendix W.

The mediating role that *Justifiability of cheating* played at Time 1, between the personological and contextual variables to the left of the model, and both outcome variables to the right of the model, was prominent again at Time 2. A substantial correlation between *Surface learning strategies* and *Self-reported cheating* (r = .445) became non-significant in the model (r = -.060, NS) principally due to the variance they shared through *Justifiability of cheating*, which was the only variable to exert a significant direct effect on *Self-reported cheating*. Variance from contextual factors was overwhelmingly channeled into *Justifiability of cheating*, which exerted, in turn, large and highly significant effects on both *Surface learning strategies* (β = .586, p < .001) and *Self-reported cheating* (β = .687, p < .001). An exception to the pattern observed at Time 1 was that the significant direct effect of *Performance goal structure* on *Surface learning strategies* (β = .157, p < .05), disappeared at Time 2 (β = .094, NS). The effect of *Performance goal structure* on *Justifiability of cheating* was larger, by contrast, at Time 2 (β = .388, p < .001) than at Time 1 (β = .237, p < .001), conveying substantial indirect effects to both *Self-reported cheating* (β = .267, p < .01) and *Surface learning strategies* (β = .227, p < .01) (see Time 2 indirect effects in Appendix X).

The role of *Peer norms* at Time 1 as a partial mediator for the effects of both *Performance goal structure* and *Teacher quality* on *Justifiability of cheating* shifted, at Time 2, to one of negligible mediation of *Performance goal structure* (β = -.040, NS), and complete mediation of *Teacher quality* (β = -.412, p < .001), of which a significant indirect effect carried through to *Justifiability of cheating* (β = .121, p < .01). *Teacher quality* also exerted a substantial but non-significant direct effect on *Justifiability of cheating* (β = -.229, p = .077) that has been principally associated with female sample models in the present study, and a substantial indirect effect on *Self-reported cheating* (β = .267, p < .01), by way of *Peer norms* → *Justifiability of cheating*.

Usefulness of Curriculum was, again, predictively inert in Model 4. Stepwise regression was conducted with respect to all predictors of *Peer norms*, *Justifiability of cheating*, *Surface* *learning strategies*, and *Self-reported cheating* (see Appendix Y). Results were similar to those of the stepwise regression performed at Time 1, with the exception that at Time 2 the disappearance of effects exerted by *Usefulness of curriculum* was associated exclusively with the addition of *Teacher quality*. In all four stepwise regression models, the large correlation between these two variables (r = .546 in Figure 7.3) appeared to mute the independent effects of *Usefulness of curriculum* as if *Teacher quality* were acting as a complete mediator. As at Time 1, an equivalent model tested with *Teacher quality* regressed, as a mediator, on *Usefulness of Curriculum* (see Appendix AA; and section 7.8.1) demonstrated equivalent fit to Model 4 ($\Delta\chi^2(1) = 1$).

Personological factors predicted downstream variables principally by way of their effects on learning context factors. The only significant direct effect of self-concept on an outcome variable at Time 2 was exerted by *Subject self-concept* on *Surface learning strategies* (β = -.192, p < .05). Both self-concept variables exerted significant indirect downstream effects, including from *Subject self-concept* to *Self-reported cheating* (β = -.235, p < .001) and *Surface learning strategies* (β = -.176, p < .01), and from *Honesty-trustworthiness self-concept* to *Justifiability of cheating* (β = -.135, p < .01) and *Peer norms* (β = -.126, p < .01).

A major difference in the pattern of effects observed at Time 2 *versus* at Time 1 was the inclusion of the path from *Honesty-trustworthiness self-concept* to *Teacher quality* (β = .306, p < .001). The magnitude of this effect was consistent with that observed in the female sample model at Time 2 (β = .311, p < .001), and with the zero-order correlation between these two factors in the Time 1 measurement model (r = .217; see Table 6.8). As the only significant direct effect exerted by *Honesty-trustworthiness self-concept* at Time 2, this new path conveyed variance to *Teacher quality* that was transmitted as the abovementioned indirect effects (see Appendix X) to further downstream variables. It is worth noting that the effects of *Honesty-trustworthiness self-concept* on *Self-reported cheating* (β = -.043, NS), *Justifiability of cheating* (β = -

.001, NS), and *Peer norms* (β = -.095, NS) were also small and non-significant when the path to *Teacher quality* was constrained to zero.

Table 7.9

Model 4: Standardized beta coefficients estimated with observed variables, Time 2 co-ed data (N = 297)

	Predictors								
N = 297	Grade	Gender	Sub	Hon	Perf	Curuse	Teacher	Peer	Chjust
Sub	-.099	.241***							
Hon	.013	-.028	---						
Perf	-.136*	.257***	.113	---					
Curuse	-.103*	-.009	.523***	---	---				
Teacher	-.177**	.020	.369***	.306***	---	---			
Peer	-.156**	.238***	---	-.095	-.040	-.107	-.412***		
Chjust	.045	.149*	-.144	-.001	.388***	-.009	-.229	.295**	
Surf	.080	-.003	-.192*	---	.094	-.116	.060	---	.586***
Cheat	.064	.044	-.065	-.053	-.153	-.093	.063	.087	.687***

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; **p* < .05, ***p* < .01, ****p* < .001.

Weighted composite score estimation. To address the low *N:q* ratio of the model estimated with observed indicator variables (1.5), composite scores were used to re-estimate the model. The *N:q* ratio for the structural model estimated with composite scores improved to 4.3, and the model's fit to the co-ed data set was excellent ($\chi^2(8)$ = 13.7, *p* = .09; *RMSEA* = .049, *CI*s = .000 - .092, *pclose* = .458; *TLI* = .97; *CFI* = .99; *SRMR* = .014; *N:q* = 4.3). A Heywood case involving a small negative residual variance for *Assessment quality* (-.094) was corrected by setting the residual variance equal to .00001 (Muthen & Muthen, 2014). The fit of the hypothesized structural model was otherwise very good ($\chi^2(15) = 33.7$, $p = .004$; *RMSEA* = .065, *CI* s= .035 - .094, *pclose* = .182; *TLI* = .92; *CFI* = .98; *SRMR* = .025; *N:q* = 4), explaining 63% of the variance in *Self-reported cheating*, 56% of the variance in *Surface learning strategies*, and 43% of the variance in *Justifiability of cheating*. These estimates differed from the model estimated with all observed indicators by $\Delta R^2 = 6\%$, $\Delta R^2 = 5\%$, and $\Delta R^2 = 0\%$, respectively. The mean effect size of the model was larger by $M_{\Delta\beta} = .001$ when estimated with composite scores, with a mean absolute difference of $|\Delta\beta_M| = .02$, and absolute difference range of $|\Delta\beta| = .000$ - .098. The pattern of significant and non-significant paths in the model estimated with composite scores was identical to that estimated with observed indicator variables (see Appendix Z).

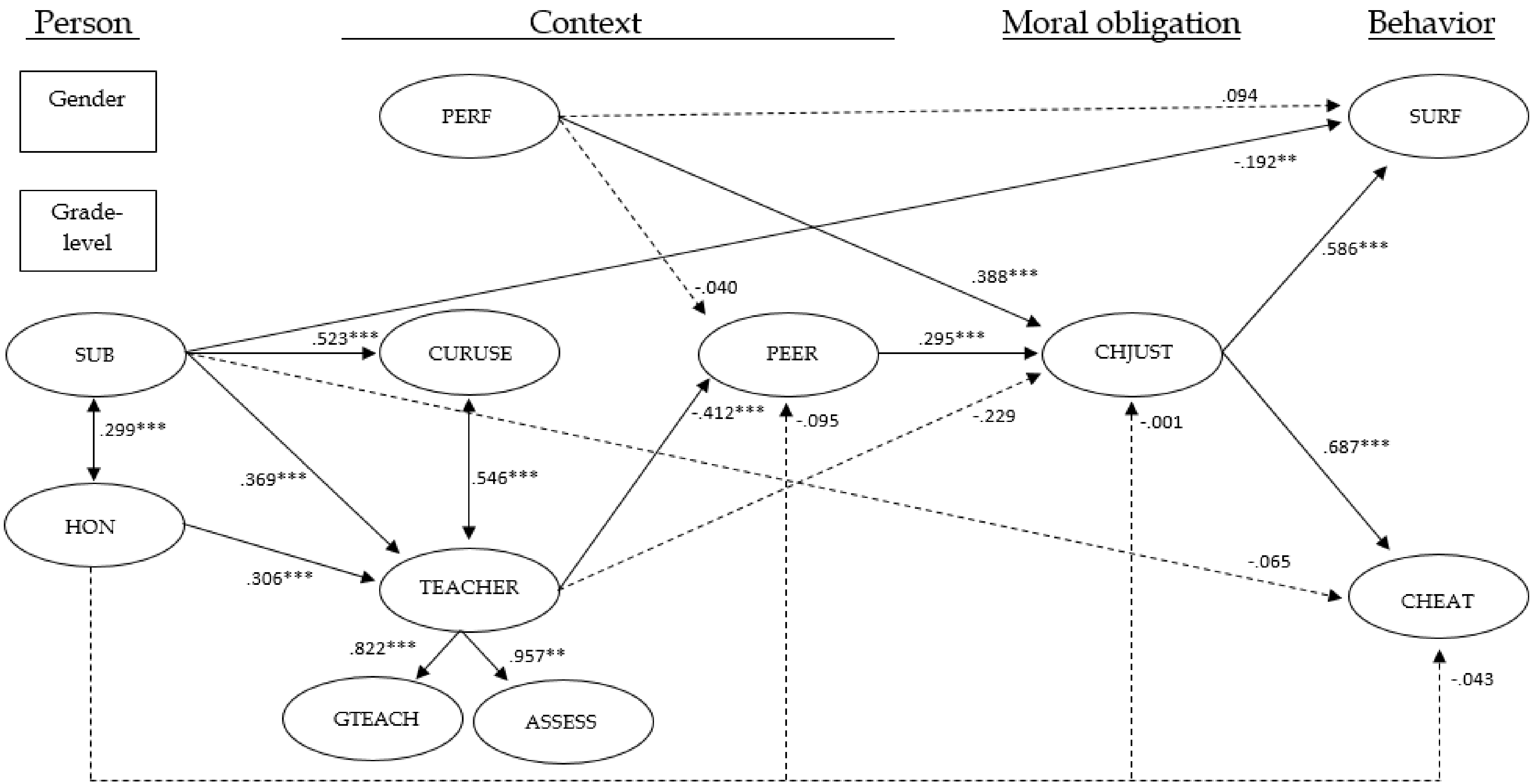

Figure 7.3. Co-ed sample results for Model 4, Time 2 (N = 297). $\chi^2(1173) = 1821$; *RMSEA* = .043, *CIs* = .039 - .047, *pclose* = .999; *TLI* = .90; *CFI* = .91; *SRMR* = .069; *N:q* = 1.5; *SCF* = .937. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; *p < .05, **p < .01, ***p < .001. - - - - paths were significant for the co-ed sample at Time 1 (see figure 6.4).

7.8.1 Equivalent models

The four equivalent models tested at Time 1 (see section 6.9.1) were re-tested at Time 2. *Usefulness of curriculum* was again positioned as a predictor of *Teacher quality* in Equivalent Model 1 (see Appendix AA), which achieved acceptable fit ($\chi^2(1174) = 1822$; *RMSEA* = .043, *CI*s = .039 - .047, *pclose* = 1.00; *TLI* = .90; *CFI* = .91; *SRMR* = .069; *N:q* = 1.5; SCF = .937) that was virtually identical to that of the hypothesized PTLC model (Model 4; see Figure 7.3) ($\Delta\chi^2(1) = 1$, NS). *Peer norms* was positioned as a correlate of class context in Equivalent Model 2 (see Appendix AB), and as a predictor of class context in Equivalent Model 3 (see Appendix AC). Both models explained the same amount of variance in *Self-reported cheating* (57%), *Justifiability of cheating* (42%), and *Surface learning strategies* (50%), but with a small decrement in variance explained in each of latter two constructs of $\Delta R^2 = -.01$, respectively, as compared to the hypothesized PTLC model. Acceptable fit was achieved both by Equivalence Model 2 ($\chi^2(1173) = 1818$; *RMSEA* = .043, *CI*s = .039 - .047, *pclose* = .999; *TLI* = .90; *CFI* = .91; *SRMR* = .068; *N:q* = 1.5; SCF = .937) and by Equivalent Model 3 ($\chi^2(1173) = 1820$; *RMSEA* = .043, *CI*s = .039 - .047, *pclose* = .999; *TLI* = .90; *CFI* = .91; *SRMR* = .068; *N:q* = 1.5; SCF = .937).

An Equivalent Model 4 was additionally tested, in which *Peer norms* was positioned as a correlate of *Justifiability of cheating*. As was the case at Time 1 (see Section 6.9.1), the most prominent difference between Equivalent Model 4 and the hypothesized PTLC model was in the amount of variance explained in *Justifiability of cheating* (37% and 43%, respectively). The models otherwise demonstrated identical fit to the data ($\chi^2(1173) = 1821$; *RMSEA* = .043, *CI*s = .039 - .047, *pclose* = .999; *TLI* = .90; *CFI* = .91; *SRMR* = .069; *N:q* = 1.5; SCF = .937), explained the same amount of variance in *Self-reported cheating* (57%), and differed structurally only with respect to the effects on *Justifiability of cheating* of *Performance goal structure* ($\beta = .377$, $p < .001$) and *Teacher quality* (β = -

.350, $p < .01$), with a correlation between *Peer norms* and *Justifiability of cheating* in Equivalent Model 4 ($r = .308$, $p < .01$) that was consistent with the corresponding bivariate correlation in the Time 2 measurement model (see Table 7.4).

Having the same degrees of freedom as the hypothesized PTLC model, Equivalent Model 2 ($\chi^2(1173) = 1818$) demonstrated the best fit of the five models considered above. The improvement in fit demonstrated by Equivalent Model 2 over the hypothesized PTLC model ($\Delta\chi^2(0) = -3$) was not, however, sufficient to justify adopting Equivalent Model 2 as the central structural model for the present study, especially in light of the fact that Equivalent Model 2 demonstrated worse fit than the hypothesized PTLC model at Time 1 (see Figure 6.4).

7.9 Chapter summary

Time 2 data ($N = 297$) was used in this chapter to analyze the factorial validity, multi-group invariance, and demographic characteristics of hypothesized PTLC measurement and structural models. Several of these models were estimated separately with both observed indicator variables and composite scores. Composite scores generally more than doubled the $N{:}q$ ratios achieved with observed indicator variables, but with minimal difference observed among effect sizes.

DIF analysis, used at Time 2 to assess multi-group invariance, cross-validated the gender differences in factorial structure observed at Time 1, especially with respect to *Subject self-concept*, *Justifiability of cheating* and *Self-reported cheating*. Gender-specific models were estimated for this reason prior to estimating the co-ed model. While hypothesized Model 3 demonstrated acceptable fit to the male data set ($N = 115$), modification indices associated with the female sample model ($N = 182$) indicated the need for a beta path from *Honesty-trustworthiness self-concept* to *Teacher quality*. Correlations between *Honesty-trustworthiness self-concept* and *Teacher quality* had been noted in earlier work, and were interpreted in the present study as reflecting cynicism, which literature cited in section 7.7.2 indicates is more common among female secondary students, and tends to increase with age during adolescence. The addition of this parameter resulted in Model 4.

The overall pattern of effects that emerged when Model 4 was tested against the co-ed data set (N = 297) was very similar to the co-ed model at Time 1 with respect to contextual effects, but notably different with respect to personological effects. The pronounced mediating roles played by *Justifiability of cheating* and *Peer norms* at Time 1, persisted at Time 2. All effects of class context on both outcome variables were, in fact, mediated by *Justifiability of cheating* in the Time 2 co-ed model, including the effect of *Performance goal structure* on *Surface learning strategies*, which was significant at Time 1. Additionally, substantial bivariate correlations between *Surface learning strategies* and *Self-reported cheating* fell to non-significance in all models at Times 1 and 2. This appeared to happen in the co-ed model at Time 2 because the two variables were sharing variance through *Justifiability of cheating*, the only significant predictor of *Self-reported cheating*, which supports the proposed grouping of cheating and surface learning under the term *disintegrity* (Miller et al., 2011)

The pattern of personological effects was, by contrast, more heavily mediated in the co-ed model at Time 2 than that at Time 1. The combined number of direct effects on downstream variables exerted by *Subject self-concept* and *Honesty-trustworthiness self-concept* fell from seven at Time 1, to three at Time 2. The effects of both of these variables in the Time 2 model appeared, moreover, to be transmitted principally by *Teacher quality* to *Peer norms*.

The three structural models examined in this chapter (male, female, and co-ed) all supported the PTLC hypothesis that contextual and personological variables affect cheating largely as a function of moral obligation. The two distinctive patterns of effects noted at Time 1, involving (1) personological variables exerting direct effects on outcome variables and (2) contextual variables being overwhelmingly mediated by *Peer norms* and *Justifiability of cheating*, appeared to consolidate at Time 2. Personological variables at Time 2 were also overwhelmingly mediated by downstream variables, including *Teacher quality*, *Peer norms*, and *Justifiability of cheating*. Finally, the fact that substantial bivariate correlations between *Self-reported cheating* and *Surface learning strategies* became non-significant in all of the models tested at Times 1 and 2 supports Miller et al.'s (2011) contention that cheating and surface learning should both be considered forms of disintegrity.

CHAPTER 8

LONGITUDINAL ANALYSES

This chapter presents the longitudinal analysis of Model 4 with data provided by $N = 225$ respondents at Time 1, that were matched, using anonymous identification codes, with data from the same respondents at Time 2. The goals of longitudinal analysis were to (1) test the consistency of hypothesized effect patterns over time, and (2) control for Time 1 variance in Time 2 measures included in Model 4. The latter of these goals, controlling for prior variance in Time 2 measures, addressed two important concerns about the validity of self-report questionnaire studies: firstly, that self-report measures of how individuals perceive contexts "are not objective… and are influenced by individual differences" (Bing, Davidson, Vitell, et al., 2012, p. 33); and secondly, that self-reports of cheating-related attitudes and behaviors may be biased by respondents presenting themselves in socially appropriate ways, even on anonymous questionnaires (Johnson & Richter, 2004; Martin, Rao, & Sloan, 2009; Miller et al., 2008; Walker, 2010). Both of these concerns are addressed by longitudinally purging extraneous within-person variance that may carry across time and context due to individual tendencies, self-beliefs, and personality factors. Longitudinal design rendered, in this manner, a closer approximation of Model 4 effects that were unique to Time 2 contexts.

8.1 Longitudinal measurement model analysis

The large number of free model parameters included in the longitudinal model (see Figure 8.1) could be estimated with N = 225 cases only by using weighted composite scores. Six models in preceding chapters estimated separately with weighted composite scores and observed indicator variables demonstrated that the two estimation methods conform closely with respect to data in the present study, in terms of effect patterns, effect magnitudes, and variance explained. These six models included MIMIC analyses at Time 1 and 2, gender-specific and co-ed models at Time 1, and the co-ed model at Time 2. When estimated with weighted composite scores, beta coefficients were found to be approximately .011 larger, overall, with an absolute mean difference of .03. The amount of variance explained in outcome variables also tended to be slightly larger in models estimated with composite scores, with average absolute differences of $|\Delta R^2|$ = 3% in *Justifiability of cheating*, $|\Delta R^2|$ = 6% in *Surface learning strategies*, and $|\Delta R^2|$ = 5% in *Self-reported cheating*. This evidence suggests that using composite scores to estimate models with data collected for the present study renders approximately equivalent effect sizes, but with higher statistical power and improved model fit.

8.1.1 Congeneric model analysis.

The majority of congeneric models presented in Table 8.1 demonstrated satisfactory fit to Time 1 and Time 2 subsets of the longitudinal data. This excluded consideration of whether upper-bound *RMSEA* confidence intervals exceeded .01, unless they were accompanied by *RMSEA* point-estimates that exceeded .08. A notably weak congeneric model at Time 2 was *Assessment quality*, which demonstrated poor approximate fit to the longitudinal sample at both Time 1 ($\chi^2(14)$ = 43; *CFI* = .92; *TLI* = .88; *RMSEA* = .096, *CI*s = .064 - .129; *SRMR* = .048; Rho = .84) and Time 2 ($\chi^2(14)$ = 57.2; *CFI* = .92; *TLI* = .88; *RMSEA* = .117, *CI*s = .086 - .149; *SRMR* = .048; Rho

= .89). Weakness in this factor was indicated with respect to Time 1 and 2 subsets of the longitudinal data by high *RMSEA* point-estimates and confidence intervals, as well as by low *CFI* and *TLI* estimates. Acceptable fit with respect to *SRMR*, and good Rho reliability at both time points suggested, however, that misfit observed in this congeneric model might be mitigated by inclusion in the second-order factor structure *Teacher quality*. The fit of the second-order factor model for *Teacher quality* to the longitudinal data set was next examined.

Teacher quality. The fit of the second-order factor *Teacher quality* to longitudinal data was assessed in two ways. Firstly, CFA results of the multivariate measurement model were compared with, and without, specifying the second-order factor. Secondly, second-order structure was estimated with Time 1 and Time 2 data, respectively, by regressing it, for the purpose of model identification, on the covariate 'maternal educational attainment'. Maternal educational attainment, which had negligible statistical associations with both *Teacher quality* (β = .061 and β = .064, NS, respectively) and its first-order components ($R_{|\beta|}$ = .047 - .064), was thus used as an instrumental variable, in order to identify the two-factor structure of the second-order model (Kenny, 2014b).

Specifying the second-order factor improved the fit of the multivariate measurement model very slightly with respect to Time 2 longitudinal data (Δ*CFI* = .001), and made no difference to fit with respect to Time 1 longitudinal data (Δ*CFI* = .000). The models that employed maternal educational attainment in order to identify *Teacher quality* also achieved satisfactory fit at Time 1 ($\chi^2(102)$ = 165; *CFI* = .94; *TLI* = .93; *RMSEA* = .053, *CI*s = .038 - .068; *SRMR* = .049) and at Time 2 ($\chi^2(103)$ = 195; *CFI* = .93; *TLI* = .92; *RMSEA* = .064, *CI*s = .050 - .078; *SRMR* = .047).

These results indicated that *Teacher quality* fit the longitudinal data. The questionable fit of *Assessment quality* to longitudinal data did not, therefore, merit *post-hoc* modification or exclusion from the model because *Teacher quality* was the measurement structure within which it functioned in Model 4.

Surface learning strategies, Peer norms, and Performance goal structure. The weak approximate fit to Time 1 data of *Peer norms* (*RMSEA* = .097; *CFI* = .93; *TLI* = .87) and *Surface learning strategies* (*RMSEA* = .112; *TLI* = .86) was offset in both cases by acceptable lower-bound *RMSEA* confidence intervals, appropriate *SRMR* values, and adequate scale reliability (see Table 8.1). In the case of *Surface learning strategies*, *CFI* was also acceptable (.95).

At Time 2, *Peer norms* fit the data well on all criteria, whereas *Surface learning strategies* had, again, a high *RMSEA* point-estimate (.088), albeit with an acceptable lower-bound confidence interval and better overall fit than it demonstrated at Time 1, including a non-significant chi-squared value ($\chi^2(2) = 5.48$, $p = .065$; *CFI* = .98; *TLI* = .94; *RMSEA* = .088, *CI*s = .000 - .180; *SRMR* = .030; Rho = .76). A nearly identical pattern of fit statistics was also observed at Time 2 for *Performance goal structure*, which had a high RMSEA point-estimate (.094), but otherwise good fit ($\chi^2(2) = 5.97$, $p = .051$; *CFI* = .98; *TLI* = .92; *RMSEA* = .094, *CI*s = .000 - .185; *SRMR* = .029; Rho = .76).

Table 8.1

Congeneric model results longitudinal sample (N = 225)

Scale (# items)	χ^2	p	df	Loading range	RMSEA value	RMSEA low CI	RMSEA High CI	CFI	TLI	SRMR	Rho
Time 1 data set											
Subject self-concept (5)	8.57	.128	5	.72-.86	.056	.000	.119	.99	.99	.016	.92
Honesty-trust. self-concept (6)	11.63	.235	9	.41-.85	.036	.000	.088	.99	.99	.025	.82
Performance structure (4)	3.93	.141	2	.48-.82	.065	.000	.162	.98	.95	025	.72
Good teaching (8)	34.28	.024	20	.42-.79	.056	.020	.088	.97	.96	.037	.87
Usefulness of curriculum (4)	2.18	.337	2	.72-.94	.020	.000	.135	1.00	1.00	.011	.91
Assessment quality (7)	43.00	.000	14	.56-.72	.096	.064	.129	.92	.88	.048	.84
Peer norms (5)	15.67	.008	5	.45-.76	.097	.045	.154	.93	.87	.035	.74
Surface learning strategies (4)	7.61	.022	2	.29-.86	.112	.036	.201	.95	.86	.034	.71
Justifiability of cheating (3)	.451	.502	1	.61-.86	.000	.000	.153	1.00	1.02	.012	.79
Self-reported cheating (3)	2.32	.128	1	.77-.91	.076	.000	.211	.99	.97	.073	.86
Time 2 data set											
Subject self-concept (5)	4.79	.442	5	.70-.92	.000	.000	.091	1.00	1.00	.011	.92
Honesty-trust. self-concept (6)	20.14	.017	9	.50-.88	.074	.030	.118	.97	.95	.036	.85
Performance structure (4)	5.97	.051	2	.53-.88	.094	.000	.185	.98	.92	.029	.79
Good teaching (8)	33.57	.030	20	.30-.83	.055	000	.086	.98	.97	.031	.88
Usefulness of curriculum (4)	.517	.772	2	.78-.94	.000	.000	.088	1.00	1.01	.003	.93
Assessment quality (7)	57.15	.000	14	.69-.78	.117	.086	.149	.92	.88	.048	.89
Peer norms (5)	3.02	.696	5	.59-.91	.000	.000	.070	1.00	1.02	.015	.85
Surface learning strategies (4)	5.48	.065	2	.35-.85	.088	.000	.180	.98	.94	.030	.76
Justifiability of cheating (3)	.009	.925	1	.65-.82	.000	.000	.061	1.00	1.05	.003	.78
Self-reported cheating (3)	.046	.830	1	.62-.96	.000	.000	.105	1.00	1.03	.015	.84

Note. χ^2 = chi-squared; p = significance level; *df* = degrees of freedom; *CI* = confidence interval; Rho = Rho reliability coefficient; highlights = index threshold violation.

8.1.2 Multivariate measurement model analysis

The twenty-factor longitudinal measurement model demonstrated excellent fit ($\chi^2(33)$ = 43.337; *CFI* = .99; *TLI* = .97 *RMSEA* = .037, *CI*s = .000 - .065, *pclose* = .745; *SRMR* = .017; *N:q* = 1.1). Bivariate correlations reported in Table 8.2 were consistent, in terms of magnitude and direction, with the correlation matrices estimated for cross-sectional data sets in Chapters Six and Seven.

Correlational analysis. Ten correlation coefficients reported in Table 8.2 could be considered excessive, at $r > .750$. The largest six of these correlations, with a range of R_r = .776 - .995, were between first-order measures of *Good teaching* and *Assessment quality*, respectively, and with the second-order factor *Teacher quality* that comprised them. These six large correlations further validated the decision to model *Good teaching* and *Assessment quality* as a single factor.

Correlations at Time 2 between *Justifiability of cheating* and both *Surface learning strategies* and *Self-reported cheating* were of nearly identical magnitude (r = .758 and r = .759, respectively). This conformed to a pattern of large statistical associations between these factors that was observed throughout the present study. It would be inappropriate to model these relationships with a higher-order factor structure, however, because of the categorical difference between psychological processes, as measured by *Justifiability of cheating* (e.g. moral judgment), and behavioral activities such as surface learning and cheating. These large correlations appear, instead, to support the hypothesis that the justifiability of cheating is a key psychological driver of disintegrity behaviors.

The final two large correlations observed in Table 8.2 were between the Time 1 measures of *Subject self-concept* and *Honesty-trustworthiness self-concept*, and their respective Time 2 counterparts (r = .796 and r = .754). These longitudinal correlations suggest that each factor in question represented a source of within-person variance that remained consistent over time. The sources of variance tapped by these measures appeared, additionally, to be distinct from one another, as indicated by their comparatively small bivariate correlations at Time 1 (r = .206, p < .01) and Time 2 (r = .212, p < .01).

Table 8.2

Higher-order CFA correlation matrix for the longitudinal model (N = 225)

	SUB1	HON1	PERF1	GTEACH1	CURUSE1	ASSESS1	PEER1	SURF1	CHJUST1	CHEAT1	TEACHER1
SUB1	-----										
HON1	.218**	-----									
PERF1	.128	-.019	-----								
GTEACH1	.405***	.287***	-.112	-----							
CURUSE1	.507***	.182*	.109	.609***	-----						
ASSESS1	.434***	.307***	-.121	.787***	.653	-----					
PEER1	-.189*	-.419***	.236***	-.468***	-.247**	-.501***	-----				
SURF1	-.356	-.374***	.271**	-.371***	-.349***	-.397***	.517***	-----			
CHJUST1	-.161*	-.434***	.428***	-.451***	-.329***	-.483***	.695***	.708***	-----		
CHEAT1	-.313***	-.539***	.236**	-.402***	-.256**	-.430***	.586***	.642***	.675***	-----	
TEACHER1	.472***	.335***	-.131	.857***	.711***	.918***	-.546***	-.433***	-.526***	-.469***	-----
SUB2	.796***	.206**	.084	.286***	.407***	.306***	-.118	-.379***	-.109	-.286***	.333***
HON2	.212**	.754***	-.010	.179**	.127	.191*	-.191	-.280**	-.206*	-.428***	.208**
PERF2	.000	-.189*	.340***	-.008	-.007	-.009	.065	.261**	.287***	.272**	-.009
GTEACH2	.260**	.208*	-.018	.325***	.274*	.348**	-.256*	-.221**	-.107	-.184*	.379***
CURUSE2	.355***	.189*	.075	.280**	.444***	.300***	-.161*	-.258**	-.125	-.189*	.326***
ASSESS2	.332***	.265***	-.023	.415***	.349***	.444***	-.179*	-.258***	-.125	-.189**	.484***
PEER2	-.142	-.238**	.073	-.296***	-.256***	-.317***	.573***	.239**	.318***	.267***	-.346***
SURF2	-.176*	-.269**	.133	-.156	-.208**	-.167**	.195*	.632***	.353***	.360***	-.182***
CHJUST2	-.199*	-.341***	.189*	-.290*	-.259**	-.310***	.413***	.506***	.561***	.527***	-.338***
CHEAT2	-.240**	-.243**	.152	-.124	-.185*	-.133*	.260**	.374***	.296***	.440***	-.145
TEACHER2	.334***	.267***	-.023	.417***	.351***	.446***	-.179*	-.283***	-.138	-.236**	.486***

Table 8.2, continued

Higher-order CFA correlation matrix for the longitudinal model

	SUB2	HON2	PERF2	GTEACH2	CURUSE2	ASSESS2	PEER2	SURF2	CHJUST2	CHEAT2
SUB2	-----									
HON2	.244**	-----								
PERF2	.121	-.096	-----							
GTEACH2	.407***	.312***	.089	-----						
CURUSE2	.579***	.221**	.152*	.522***	-----					
ASSESS2	.519***	.398***	.113	.776***	.666***	-----				
PEER2	-.184*	-.355***	-.049	-.360***	-.334	-.460***	-----			
SURF2	-.335***	-.329***	.374***	-.308***	-.382***	-.393***	.243**	-----		
CHJUST2	-.276***	-.292***	.369***	-.281***	-.335***	-.359***	.454***	.758***	-----	
CHEAT2	-.369***	-.318***	.130	-.285***	-.377***	-.363***	.460***	.587***	.759***	-----
TEACHER2	.522***	.400***	.114	.780***	.669***	.995***	-.462***	-.395***	-.360***	.365***

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating. *p < .05, **p < .01, ***p < .001.

8.2 Demographics: Longitudinal matched-samples *t*-tests

MIMIC models were used in previous chapters to identify a number of mean differences between grade-level groups. It is similarly of interest to examine how factor means changed within groups during the year between Time 1 and Time 2 data collections. Over this period, Grade Eight students matriculated to Grade Nine, marking a transition from a Middle School environment to a High School environment; and Grade Nine Students matriculated, within the high school environment, to Grade Ten. Changes in factor means over these grade-level transitions were examined with matched-samples *t*-tests instead of a MIMIC model, due to sample size restrictions (see Appendix AD).

Following the transition from Grade Eight to Grade Nine, students reported using surface learning strategies more often (t = 2.67, p = .009), having worse impressions of their Science teachers (t = -2.04, p = .044), and being more aware of performance goal structures in Science class t = -2.65, p = .009). While the mean difference in *Self-reported cheating* for this group was non-significant over the Grade 8 – 9 transition (t = 1.57, NS), it is important to note that the proportion of students who reported cheating at all in Science class during the preceding year, or the 'incidence', increased from 48% in Grade Eight to 56% in Grade Nine.

Following the transition from Grade 9 to Grade 10, students reported an overall improvement in *Subject self-concept* (t = 4.45, p < .000), *Usefulness of curriculum* (t = 2.96, p = .004), and on both constituent measures of *Teacher quality*: *Good teaching* (t = 3.92, p < .000) and *Assessment qualtiy* (t = 4.04, p < .000). While these positive changes did not accompany a significant mean difference in *Self-reported cheating* (t = -.489, NS), the incidence of cheating fell from 57% in Grade Nine to 47% in Grade Ten. This stands in contrast to the more common finding that the incidence of cheating increases across the high school years (Galloway, 2012; Miller et al., 2007).

8.3 Invariance analysis of the longitudinal measurement model

While the persistence of effect patterns over time is of principal concern to the present study, differences between Times 1 and 2 are also potentially of interest. Interpreting such differences treats Times 1 and 2 as separate groups, despite the use of matched samples, and rests, therefore, on the assumption that factor structure is invariant between them (Cheung & Rensvold, 2002). The assumption of factorial invariance was tested with respect to Time 1 and Time 2 data by treating each data set (N = 225) as if it represented a different group of people. This effectively doubled the size of the longitudinal sample to N = 450 for the purpose of invariance analysis, but only just surpassed the minimum recommended size for sub-groups, or N = 200 (Meade et al., 2008). CFA was used to make initial estimates of the fit of each cross-sectional measurement model, reported as 'baseline models' in Table 8.3. Then, as in prior chapters, an increasingly strict series of equality constraints was applied to both models, simultaneously.

The greater complexity of tests of invariance, over CFAs can be seen by comparing the number of free model parameters in baseline CFAs (186) in Table 8.3 to that in the invariance models (273 – 322). While the longitudinal sample was large enough to execute such complex analyses, it was not sufficient to obtain good approximate fit, nor was Boomsma and Herzog's (2014) small sample correction function, used in prior chapters to reduce bias associated with low $N{:}q$ ratios, viable for multi-group analyses (A. Boomsma, personal communication, 25 March 2014). As argued in section 6.4, $CFI \geq .86$ is a conservative CFI threshold for invariance analyses in the present study that involve approximately two-thirds more free model parameters than baseline CFAs, but with group sample sizes approaching 200. Lowering the CFI threshold for invariance models was done in the knowledge that (A) both baseline CFAs demonstrated acceptable fit, (B) all invariance models reported in Table 8.3 achieved acceptable values for

RMSEA and *SRMR*, and (C) invariance testing is generally less concerned with overall model fit than with changes in approximate fit associated with the imposition of equality constraints.

Table 8.3

Longitudinal invariance of the measurement model

	Longitudinal invariance				
	χ^2	*df*	*RMSEA*	*CFI*	*SRMR*
Baseline CFAs					
Time 1	1485	1088	.040	.92	.065
Time 2	1695	1088	.050	.90	.076
Inv. model 1					
FS	3456	2176	.051	.879	.071
Inv. model 2					
Model 1 + FL	3525	2215	.051	.876	.074
Inv. model 3					
Model 2 + VI	3676	2264	.053	.867	.077
Inv. model 4					
Model 3 + FV	3692	2275	.053	.866	.079
Inv. model 2b					
Model 2 + FV	3541	2226	.051	.876	.077

Note. FS = Factor structure (configural invariance), FL = Factor loadings (metric invariance), VI = observed variable intercepts (scalar invariance), FV = Factor variances.

Results. The longitudinal invariance of factor configuration (configural invariance), factor loadings (metric invariance), observed variable intercepts (scalar invariance), and factor variances are indicated by changes in *CFI* (ΔCFI) reported in Table 8.3. Holding observed variable intercepts equal across groups (Inv. Model 3) was associated with a decrement of ΔCFI = -.012 *vis-à-vis* the configural model (Inv. Model 1). This decrement exceeded the desired threshold of $\Delta CFI < |.010|$, indicating a lack of scalar invariance in the multivariate measurement model.

Inasmuch as factorial non-invariance implies that the operational meanings of factors vary between groups (Cheung & Rensvold, 2002), the scalar invariance observed in Table 8.3 implied that the interpretation of mean-level differences in measures used at Times 1 and 2 should be regarded as ambiguous (Byrne & van de Vijver, 2010; Byrne, 2012). The primary purposes of longitudinal modeling were, however, to test for the consistency of effect patterns over time, and to control for Time 1 variance in Time 2 factors. So while it had to be acknowledged that some item vectors at Time 1 appeared to intersect with the *y*-axis at different values than at Time 2, of greater concern was the observed invariance of factor loadings and factor variances.

When factor loadings (i.e. the loadings of items onto the factors they measure) were constrained to be equal in Inv. Model 2, *CFI* fell by just ΔCFI = -.003 with respect to the configural model (Inv. Model 1), which indicated that factor loadings were invariant between the two time points (metric invariance). No change in *CFI* was observed, additionally, between Inv. Model 2 and Inv. Model 2b, when equality constraints on factor variances were added to Inv. Model 2, but observed indicator intercepts were allowed to vary freely, which resulted in Inv. Model 2b. This indicated factor variances were also longitudinally equivalent.

While the small but significant degree of observed scalar non-invariance introduced ambiguity to the interpretation of differences between Time 1 and Time 2 data, multi-group metric invariance and equivalent factor variances supported the calculation and use of composite scores to estimate the longitudinal model. Differences in the longitudinal functioning of questionnaire items were next examined at the item-level with DIF analysis.

8.3.1 Differential item functioning analysis

Differential item functioning (DIF) analysis was used to explore item-level differences in how factors were functioning at Times 1 and 2, in order to identify where within the measurement model factorial non-invariance was concentrated (Wang & Wang, 2012). The DIF analysis was conducted by regressing all observed indicator variables, or items, on a grouping variable for time (Time 1 = 1, Time 2 = 2). This model demonstrated satisfactory fit ($\chi^2(1081) = 1867$; *CFI* = .92; *TLI* = .91; *RMSEA* = .040, *CI*s = .037 - .45, *pclose* = 1.00; *SRMR* = .057; *N:q* = 1.9; *SCF* = 958). Significant effects identified in this model, indicated by asterisks in Table 8.4, were concentrated principally in measures of class context. A single, item-level difference was also observed in item Surf91 ($\beta = -.096$, $p < .05$) of the *Surface learning strategies* measure. Eight significant differences were identified in total, with a mean absolute beta coefficient of $M_{|\beta|} = .136$, and an absolute range of $R_{|\beta|} = .096 - .218$. Of these differences, five were highly significant at the $p < .01$ level. The only difference to exceed .200 in magnitude was in item Gteach18 of the *Good teaching* measure. The signs associated with differences in *Good teaching* and *Assessment quality* items indicated more favorable mean scores for teachers at Time 2, than at Time 1, which comports with grade-level differences observed by the DIF analysis conducted at Time 2. No item on the *Justifiability of cheating* and *Self-reported cheating* measures appeared to differ between Times 1 and 2.

Table 8.4

Longitudinal differential item functioning analysis

Item	Time	Item	Time	Item	Time	Item	Time
SUB2	-.042	PERF74	.138**	CURUSE64	.012	SURF87	-.029
SUB3	-.062	PERF75	.063	TRANS28	-.115*	SURF88	.014
SUB5	-.096	GTEACH18	-.218***	TRANS32	-.086	SURF91	-.096*
SUB13	-.079	GTEACH33	.004	TRANS66	-.085	SURF97	-.029
SUB15	.020	GTEACH39	-.053	AUTH44	.008	CHJUST79	.010
HON1	-.003	GTEACH50	-.081	AUTH60	-.135**	CHJUST86	-.028
HON6	-.036	GTEACH62	-.102*	AUTH71	-.161***	CHJUST99	.046
HON8	-.011	GTEACH67	-.032	AUTH78	.069	CHEAT84	-.019
HON9	-.023	GTEACH68	.090	PEER24	.125**	CHEAT92	-.009
HON10	.011	GTEACH77	-.033	PEER31	.074	CHEAT95	-.008
HON11	.071	CURUSE19	-.073	PEER40	.037		
PERF61	-.010	CURUSE53	-.089	PEER58	-.069		
PERF69	.050	CURUSE56	.013	PEER65	.079		

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; **p* < .05, ***p* < .01, ****p* < .001.

The longitudinal measurement model appears, in sum, sufficiently invariant to proceed with the calculation of composite scores and the longitudinal analysis of Model 4. The level of longitudinal non-invariance observed in the measurement model (see Table 8.3) was generally low, and confined to observed variable intercepts. DIF analysis showed, moreover, that group differences in item-level means were concentrated primarily in measures of class context.

8.4 Hypothesized longitudinal model

The hypothetical longitudinal model presented in Figure 8.1 comprises an iteration of Model 4 estimated with Time 2 data that is regressed on a second iteration of Model 4 estimated with Time 1 data. Both iterations were estimated with matched data provided by the same 225 participants. The dot-dashed lines that crosscut the longitudinal model represent regression paths that control for Time 1 variance in Time 2 measures.

Model 4 includes two *post-hoc* modifications to the hypothesized structure. Firstly, a path from *Honesty-trustworthiness self-concept* was added at Time 2 based on analyses of female data in Chapter Seven (see section 7.7). This statistical relationship was also found to be significant at Time 1 and in the Pilot Study. Secondly, a residual covariance was included between items Surf91 and Surf97 of the measure *Surface learning strategies* to represent a method effect between items of similar wording and in close proximity to one another on the questionnaire. This residual covariance was included in the congeneric model for *Surface learning strategies* tested on the longitudinal sample (see Table 8.1), from which factor score coefficients were derived for the calculation of weighted composite scores. Gender and grade-level were also added to the hypothesized model as control variables at Time 1, based on MIMIC model results. These modifications were included in both iterations of Model 4 that compose the longitudinal structural model presented in Figure 8.1.

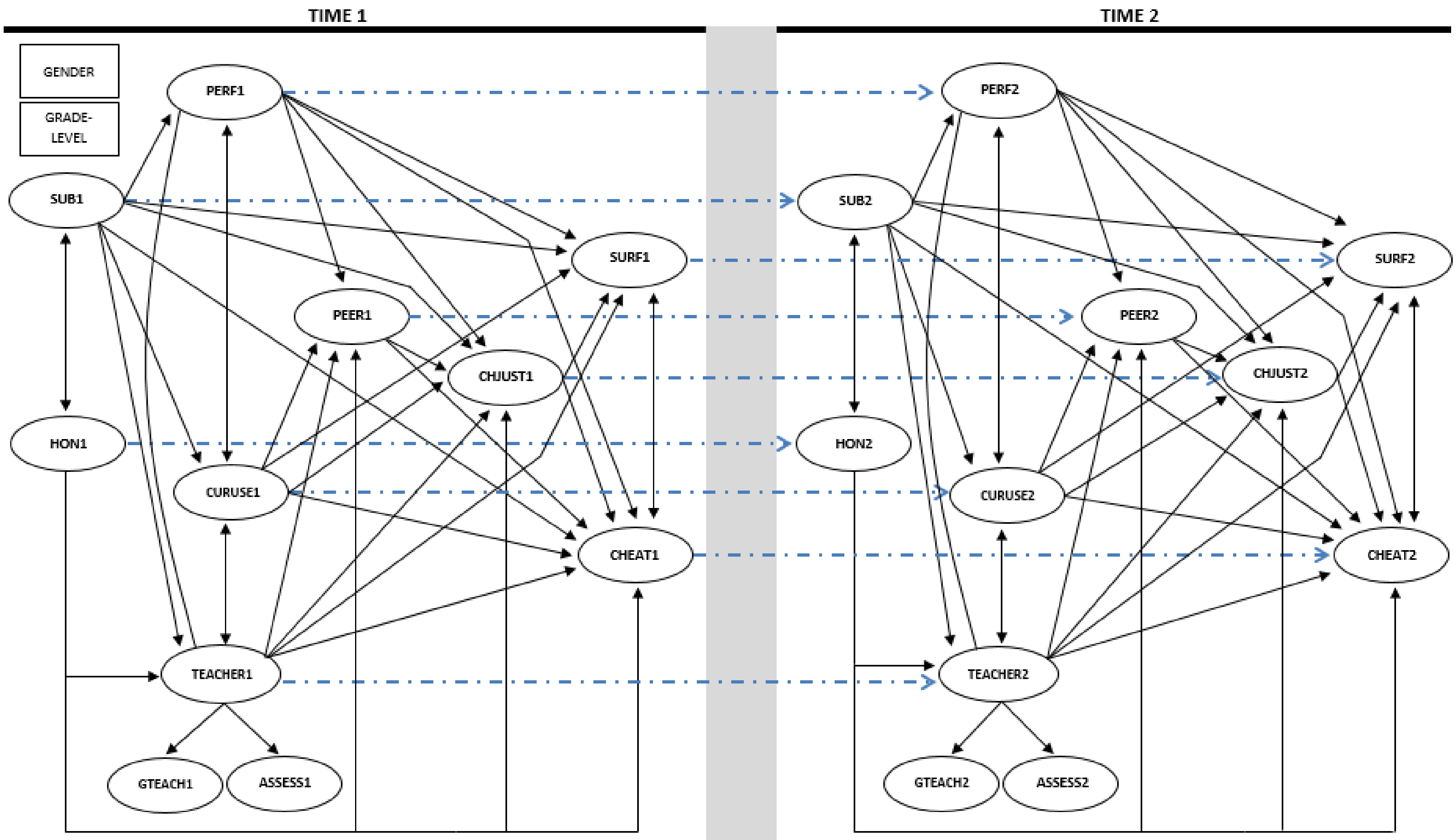

Figure 8.1. Hypothetical longitudinal model. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Abbreviations of Time 1 measures are followed by 1, and Time 2 measures by 2. · — · —· — denotes longitudinal regression paths.

8.5 Longitudinal structural model analysis (*N* = 225)

Having established the validity of the longitudinal measurement model, the longitudinal structural model was estimated using weighted composite scores. Results of this analysis are reported in Figure 8.2 and Table 8.5, and discussed in detail, below.

The longitudinal model demonstrated excellent fit to the data ($\chi^2(119) = 180$; *CFI* = .97; *TLI* = .94; *RMSEA* = .048, *CI*s = .033 - .061, *pclose* = .560; *SRMR* = .043; *N:q* = 1.5), explaining large amounts of variance in Time 2 measures of *Self-reported cheating* (61%), *Surface learning strategies* (61%), and *Justifiability of cheating* (49%). The eighteen parameters, including sixteen beta paths and two correlations, that achieved significance in the longitudinal model are presented in figure 8.2. If an effect was significant in one cross-sectional iteration but not the other, its non-significant counterpart is depicted as a dashed line. Unstandardized parameter estimates, covariance and correlation matrices, and various other longitudinal model data and descriptive statistics are provided in Appendix AE.

Patterns of significant effects in the Time 1 and Time 2 iterations of the longitudinal model closely resembled patterns observed in the corresponding co-ed cross-sectional models, hereafter 'cross-sectional models', analyzed in Chapters Six and Seven. Controlling for prior variance in Time 2 variables modestly reduced most effect sizes in comparison with the Time 2 cross-sectional model. An exception to this trend was the path from *Performance goal structure* to *Surface learning strategies* ($\beta = .187$, $p < .05$), which was significant in the longitudinal model, but non-significant in the Time 2 cross-sectional model ($\beta = -.092$, NS). The effect of *Subject self-concept* on *Surface learning strategies* became, by contrast, non-significant in the Time 2 iteration of the longitudinal model ($\beta = -.025$, NS), whereas it was significant in the Time 2 cross-sectional model ($\beta = -.192$, $p < .05$) (see Figure 7.3).

The largest autoregressive effects in the model were observed between *Subject self-concept* and *Honesty-trustworthiness self-concept* at Time 1 and their Time 2 counterparts (β = .779, and β = .742, respectively), which indicates substantial consistency in these aspects of self-perception over time and between contexts. Longitudinal effects for all other predictor variables were moderate to small, falling within a range of $R_{|\beta|}$ = .054 – .386. The smallest autoregressive effect observed in the model was for *Self-reported cheating* (β = .054, NS). This suggests that the moderate correlation (r = .440, p <.001) between its Time 1 and Time 2 measures (see Table 8.2) was accounted for by regressing the Time 2 measure on its only significant predictor, *Justifiability of cheating* (β = .639, p < .001). The effect of *Justifiability of cheating* on *Self-reported cheating* at Time 2 was, in other words, 'salient' over the effect exerted by *Self-reported cheating* at Time 1 (Martin, 2011). The effect of *Justifiability of cheating* on *Surface learning strategies* (β = .472, p < .001) at Time 2 was also salient over the longitudinal effect of *Surface learning strategies* at Time 1 (β = .237, p < .05). Substantial bivariate correlations between *Self-reported cheating* and *Surface learning strategies* at both times (r = .642, and r = .587) were, moreover, non-significant in the model (r = .215, and r = .115), due principally to the effects of *Justifiability of cheating*. Within-person differences and contextual factors at Time 2 also affected disintegrity as a function of whether students viewed it as justifiable. These observations provide strong support for the assertion that cheating and surface leaning are forms of disintegrity (Miller et al., 2011), and for the assertion that contextual factors are preeminent influences on cheating behavior as a function of moral obligation.

The longitudinal relationship between measures of *Justifiability of cheating* (β = .274, p < .05) was also exceeded, both in magnitude and significance, by the Time 2 effect of *Performance goal structure* (β = .331, p < .01). This is consistent with key assertions of achievement goal theory (Anderman et al., 1998; Anderman & Murdock, 2006), as well as with experiments that have demonstrated a positive causal relationship between extrinsic motivation and cheating (Lobel & Levanon, 1988; Mills, 1958; Shelton & Hill, 1969; Taylor & Lewit, 1966).

Teacher quality exerted large to moderate effects on *Peer norms* at Times 1 and 2 (β = -.544, $p < .001$ and β = -.357, $p < .001$, respectively), which predicted, in turn, *Justifiability of cheating* (β = .452, $p < .001$ and β = .240, $p < .05$). The causal path running from *Teacher quality* to *Peer norms* to *Justifiability of cheating* has been a consistent feature of structural models throughout the present research, with the only notable exception being the failure of *Teacher quality* to predict *Peer norms* among female respondents at Time 2 (see Figure 7.2). This causal path is also consistent with experimental findings that perceptions of teacher quality affect the justifiability of cheating (Murdock et al., 2004, 2007), and that peer norms affect cheating behavior (Gino et al., 2009; Walker et al., 1966).

Indirect effects. Indirect effects of *Teacher quality* were conveyed by *Peer norms* to *Justifiability of cheating*, both cross-sectionally, at Times 1 and 2, and longitudinally (see Appendix AF). *Teacher quality* at Time 1 exerted significant indirect effects on *Justifiability of cheating* at Time 2 via two pathways: (1) Teacher1 → Peer1 → Peer2 → Chjust2; and (2) Teacher1 → Peer1 → Chjust1 → Chjust2. *Justifiability of cheating* at Time 2 was also indirectly affected by Time 1 measures of *Peer norms*, *Performance goal structure*, *Subject self-concept*, and *Honesty-trustworthiness self-concept*. Significant indirect effects from all five of these Time 1 measures were also conveyed to *Self-reported cheating* at Time 2 by *Justifiability of cheating*.

The overall pattern of indirect effects in the longitudinal model suggests that cheating behavior at Time 1 was related to cheating behavior at Time 2 through six, longitudinal third-variable associations: (1) *Honesty-trustworthiness self-concept*, (2) *Subject self-concept*, (3) *Teacher quality*, (4) *Performance goal structure*, (5) *Peer norms*, and (6) *Justifiability of cheating*. The correlation of past cheating behavior to future cheating behavior in the present study appears, in other words, to reflect the ways students tended to perceive themselves and their Science class contexts at Time 1 that carried over to Time 2, and that affected *Self-reported cheating* in a similar manner at both times.

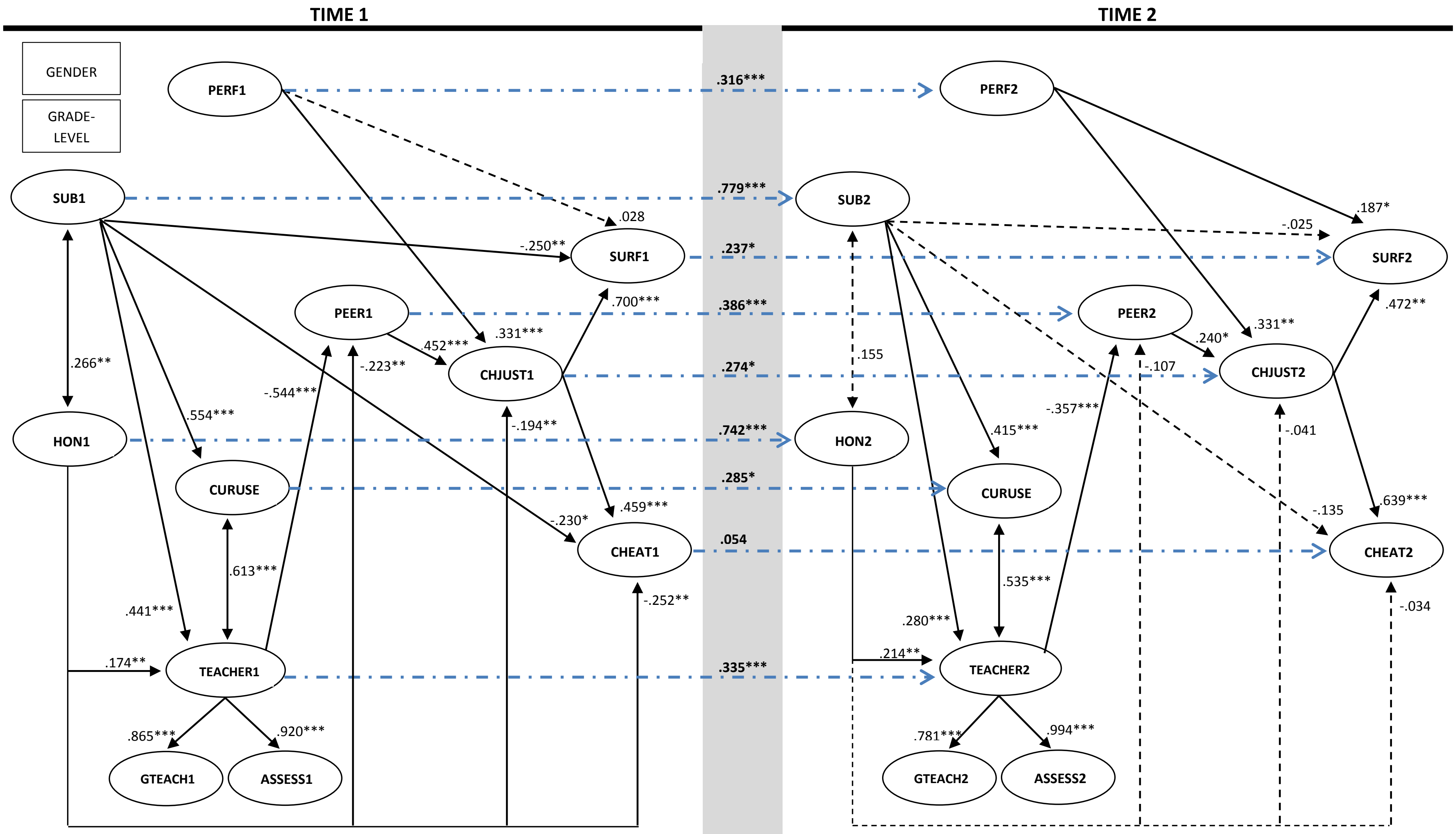

*Figure 8.2.*Longitudinal model estimated with composite variables (N = 225). $\chi^2(119)$ = 167; *CFI* = .97; *TLI* = .94 *RMSEA* = .042, *CIs* = .026 - .057, *pclose* = .799; *SRMR* = .043; $N{:}q$ = 1.5. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Abbreviations of Time 1 measures are followed by 1, and Time 2 measures by 2.

Table 8.5
Longitudinal model effects: standardized beta coefficients estimated with weighted composites

	Covariates		Time 1 predictors									Time 2 predictors						
N = 297	Grade	Gender	Sub1	Hon1	Perf1	Teacher1	Curuse1	Peer1	Chjust1	Surf1	Cheat1	Sub2	Hon2	Perf2	Teacher2	Curuse2	Peer2	Chjust2
Sub1	.014	.288***																
Hon1	-.011	-.068	---															
Perf1	.058	.100	.087	---														
Teacher1	.100	-.076	.441***	.174**	---													
Curuse1	.051	-.132*	.544***	---	---	---												
Peer1	-.072	.137	---	-.223**	.122	-.544***	.163											
Chjust1	.003	.088	.023	-.194**	.331***	-.050	-.197	.452***										
Surf1	.008	-.083	-.250**	---	.028	.107	-.057	---	.700***									
Cheat1	-.026	.088	-.230**	-.251**	.021	-.089	.133	.083	.459**	---								
Sub2	-.150**	.054	.779***															
Hon2	-.017	.050		.742***								---						
Perf2	-.100	.173*			.386***							.016	---					
Teacher2	-.193**	.136*				.335***						.280***	.214***	---				
Curuse2	-.146*	.039					.285*					.415***	---	---	---			
Peer2	-.142*	.187**						.386***				---	-.144	-.106	-.357***	-.051		
Chjust2	.072	.086							.274*			-.111	-.041	.331**	-.077	-.145	.240*	
Surf2	.082	.020								.237***		-.025	---	.187*	-.077	-.127	---	.472***
Cheat2	.074	-.001									.054	-.135	-.034	-.096	.103	-.060	.157	.639***

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating. Abbreviations of Time 1 measures are followed by 1, and of Time 2 measures by 2. Longitudinal effects are arranged as a diagonal in the lower left quadrant of the table.

8.6 Chapter summary

The matched-samples longitudinal model examined in this chapter, which comprised of two iterations of Model 4, achieved satisfactory fit using weighted composite scores. Composite scores have, in six prior models in the present study, generated effect sizes, significance levels, and R^2 estimates that were very similar to those in models estimated with observed indicator variables. Composite scores were relied upon in the longitudinal analysis because the number of free model parameters would otherwise have exceeded the number of observations, resulting in non-positive definitive model matrices.

A small but significant degree of longitudinal scalar non-invariance (in observed variable intercepts) was identified that should call into question interpretations of differences between the Time 1 and 2 iterations of the longitudinal model. The extent of this invariance was not great, however, nor was analyzing differences one of the principal purposes of the longitudinal model. Of greatest concern to the longitudinal analysis was the observed multi-group equivalence of factor loadings and variances. DIF analysis determined, additionally, that the measures most affected by non-invariance between Times 1 and 2 were the two first-order components of *Teacher quality*, i.e. *Good teaching* and *Assessment quality*. Factorial non-invariance in the longitudinal measurement model could not be attributed to changes in sample composition between Times 1 and 2, because the longitudinal sample was composed of matched observations. Longitudinal non-invariance in measures of *Teacher quality* appears, instead, to have reflected the fact that respondents were rating different individual teachers at each time point.

Control of prior variance in Time 2 measures was achieved by regressing them on their counterpart measures at Time 1. Attenuated effect sizes in the Time 2 iteration, as compared to the Time 2 cross-sectional model, reflected the improved isolation of respondents' experiences of Science class contexts from individual differences that would otherwise have contributed extraneous variance. All longitudinal effects in the model were significant, except for that of *Self-reported cheating*, which was explained almost exclusively by *Justifiability of cheating*. The longitudinal effect of *Justifiability of cheating* at Time 1 on its counterpart measure at Time 2 was, moreover, subordinate to the contextual effect of *Performance goal structure* at Time 2, and just slightly larger than the contextual effect of *Peer norms* at Time 2. These observations are consistent with theory and research that suggests cheating-related attitudes and behaviors tend to be context-specific (e.g. Murdock et al., 2001, 2004) and to become more likely in performance goal settings (Anderman, 1998; Lobel & Levanon, 1988; Mills, 1958; Murdock & Anderman, 2006; Shelton & Hill, 1969; Taylor & Lewit, 1966).

The fact that *Peer norms* completely mediated the effect of *Teacher quality* on *Justifiability of cheating* is consistent with evidence from earlier studies that cheating-related attitudes and behaviors among one's peers substantially influences one's own cheating-related attitudes and behaviors (Gino et al., 2009; Walker et al., 1966), especially inasmuch as individuals tend to formulate opinions and judgments in conformity with the opinions and judgments of their peers (Broeckleman-Post, 2008; Festinger, 1954).

The longitudinal model also provided an opportunity to test the degree of consistency between the Time 1 and Time 2 iterations. The most consistent broad characteristic of Model 4 at both time points was the pattern of contextual effects on both outcome variables. These were overwhelmingly mediated by *Justifiability of cheating*, and, to a lesser extent, by *Peer norms*. The most prominent difference observed between the Time 1 and Time 2 iterations of the longitudinal model involved the apparent consolidation of two separate and distinct effect patterns for personological factors and contextual factors, respectively, at Time 1, into a unitary pattern at Time 2, in which all effects were mediated by *Justifiability of cheating*. This change may have occurred due either to respondents' increased ages and cognitive developmental levels at Time 2 (i.e. a maturation effect), to changes they experience at higher grade-levels (i.e. a transition effect), or due to their thoughts about morality and academic integrity having been affected by taking the questionnaire at Time 1 (i.e. a Hawthorne effect) (Kline, 2009). The validity of interpretations of differences between Time 1 and 2 data is tempered, however, by the relatively small but significant degree of longitudinal scalar non-invariance detected in the measurement model.

CHAPTER 9

DISCUSSION

> An action beneficial to the welfare of society as a whole or of a fellow human being would not be considered moral if it were performed under hypnosis or under physical constraints but only if it were performed willingly, in response to values that are understood and accepted by the agent. Here lies the reason for the emphasis on judgment… Without judgment, an action, no matter how beneficial, would not be moral.
>
> —Augusto Blasi, 1980, p. 4

Results are reviewed in this chapter in relation to key hypotheses of the psychological teaching-learning contract (PTLC) perspective on disintegrity. The scholarly literatures and sub-disciplines to which such findings pertain, and their implications for educational practice, are also discussed. The PTLC model was tested longitudinally across the Grade Eight-to-Nine transition, with data from a diverse sample of students in eleven American international schools located in Eastern and Western Europe, Eastern and Western Asia, and Eastern Africa.

The present work is the first empirical study of the contractarian nature of student cheating, and the PTLC model is the first structural model to expressly reflect a dual-process conception of the psychology of academic cheating. Major findings include evidence that (1) perceived teacher quality and felt moral obligation are directly related (i.e. perceived teacher quality and the justifiability of cheating are inversely related) (2) felt moral obligation played a large role in cheating behavior, (3) the degree to which peer norms were perceived to be favorable to cheating was a key predictor of whether cheating was judged to be justifiable; (4)

'disintegrity' appeared to be an appropriate umbrella term for cheating and surface learning behaviors; and (5) the statistical relationship between past and future cheating was a function of third-variable associations.

The incidence of cheating observed among participants in the present study is reviewed below, followed by a summary of key findings related to measurement hypotheses, as pertain to psychometric modeling, multi-group factorial invariance, and demographic effects. Results are subsequently discussed in relation to key structural hypotheses in the PTLC model, followed by a discussion of the implications of these findings for educators. Prominent limitations of the present study are then discussed and suggestions made for future research in this vein.

9.1 Cheating: Definition and incidence

An effort was made to use language in cheating-related measures, including *Self-reported cheating*, *Justifiability of cheating*, and *Peer norms*, that clearly indicated to participants that the specific behaviors queried would be interpreted unequivocally as cheating (see Appendices B and C). Measures were limited, as such, to what students understood to contravene the spirit and/or letter of rules related to honorable academic conduct, thus reflecting the definition of cheating given in Chapter Three (see section 3.1). At Time 1, 53% of students admitted to having cheated during the year, as compared to 35% who admitted specifically to having cheated on tests. Proportions were nearly identical at Time 2 (54% and 34%), and in the Pilot Study (55% and 31%). This incidence of cheating in American international schools compares favorably to domestic American schools, where the reported incidence of cheating 'in any form' tends to be above 80%, and cheating on tests tends to be above 50%.

A difference was observed at Time 1 in the incidence of self-reported cheating between Grade Eight (48%) and Grade Nine (59%), that is consistent with prior findings that cheating is more prevalent at higher grade levels in secondary schools (Cizek, 1999; Galloway, 2012; Miller et al., 2007). This difference was reversed, however, at Time 2, when less cheating was reported by Grade Ten students (51%) than Grade Nine students (57%). This pattern was also observed with respect to the matched-samples longitudinal data. Students in this group who matriculated from Grade Eight to Grade Nine reported an increase in the incidence of cheating from 48% to 56%, whereas students who matriculated from Grade Nine to Grade Ten reported a decrease in cheating from 57% to 47%.

9.2 Measurement hypotheses: Findings

The PTLC model began with seventeen first-order factors (Model 1) derived from various sources in the psychometric literature. Many of these measures had not been used previously in structural equation modeling research and had, as such, not been subjected to the strict tests of construct validity entailed by congeneric and multivariate CFA. Measures associated with the *Course Experience Questionnaire* (CEQ) had also not been used previously for research on secondary school populations. CEQ measures were chosen instead for their long history of use in research on deep learning strategies and surface learning strategies, both of which were outcome variables in Model 1. Seven of the original seventeen factors of Model 1 were, over the course of analyses of both Pilot Study data and data collected for Time 1 of the Main Study, either dropped due to model misfit or multicollinearity, or combined into second-order factor structures. The specific reasons for these decisions are covered in detail in prior chapters.

9.2.1 Higher-order conceptions of class quality

Students' evaluations of class quality were initially hypothesized to consist of higher-order factors for pedagogy and assessment. This hypothesis was not supported. Assessment and pedagogy factors were highly correlated, instead, with a second-order factor for *Teacher quality*. Measures of *Clear goals and standards* and *Mastery goal structure* were found, in particular, to be extremely multicollinear with *Good teaching*, and were dropped from the study for this reason. Similar levels of multicollinearity between *Mastery goal structure* and *Teacher commitment* have previously been observed by Murdock et al. (2001, 2004), and *Clear goals and standards* has also been found to correlate highly with *Good teaching* in prior studies (Wilson et al., 1997).

Four other contextual variables sharing bivariate correlations of between r = .750 - .800 in the Pilot Study were fitted into a second-order factor structure that was identified principally with *Good teaching*. This second-order factor, dubbed *Teacher quality,* originally included a measure of student perceptions of classroom rules (*Experience of classroom rules*) and two measures of perceived assessment quality (*Authenticity* and *Transparency*), in addition to *Good teaching*. This group of four factors was reduced to a pair at Time 1 of the Main Study (*Good teaching* and *Assessment quality*), due to congeneric model misfit. It was nonetheless evident that all class quality measures, including *Mastery goal structure*, assessment measures, and CEQ measures, with the exception of *Appropriate workload*, were highly related through second-order conceptions of teacher quality.

Appropriate workload, hypothesized in the present work to measure the moral obligation to work hard, was psychometrically distinct from measures of class quality. Finding that students hold higher-order conceptions of class quality that largely exclude the issue of workload is consistent with the results of higher-order factor analyses of CEQ data in previous studies (for a review, see Wilson et al. 1997). The congeneric model for *Appropriate workload* was, however, not sufficiently valid and reliable in the present study to be included in the PTLC model tested at Time 1 (Model 3).

Three class quality factors were sufficiently valid, reliable, and distinct to retain in the final model: (1) *Good teaching* and (2) *Assessment quality*, which formed the second-order factor *Teacher quality*, and (3) *Usefulness of curriculum*. The second-order factor *Teacher quality* consistently explained between 60% and 70% of the variance in *Good teaching*, and between 75% and 99% of the variance in *Assessment quality*. This suggests that *Teacher quality* was defined in the Main Study principally in terms of *Assessment quality*, which stands in keeping with prior scholarship that argues for the crucial importance of assessment to student experience (e.g. Black & Wiliam, 1998; Hattie & Timperley, 2009; Waldrip et al., 2009).

9.2.2 Learning strategies

Measures for deep and surface learning strategies did not, on the whole, perform well in the present study. The measure *Deep learning strategies* was modified in the Pilot Study, but still failed to achieve good fit at Time 1 of the Main Study. The measure *Surface learning strategies* split into two separate factors in the Pilot Study. The more conceptually valid of these factors, which included items related to two key surface learning strategies (skipping and memorizing material), was retained for the Main Study. However, this modified measure *Surface learning strategies* additionally demonstrated misfit related to a method effect at Time 2 that was addressed by freeing an error covariance parameter between two of its items (see section 7.2.3). Surface learning behaviors have proven difficult to measure psychometrically in prior secondary-level research (Ramsden et al., 1988; Wong, Lin, & Watkins, 1996), which suggests that individual strategies may be better measured and modeled as independent factors, possibly within higher-order structures.

9.2.3 Multi-group invariance results

An important assumption of all measurement hypotheses entailed by this study was that they would be 'invariant' (also 'equivalent') across subgroups that composed the overall sample. This assumption was tested at the level of the full measurement model at Times 1 and

2 with respect to gender and grade-level groups, and was also tested in the longitudinal analysis by treating Time 1 and Time 2 data as subgroups of the longitudinal sample. Invariance of cross-sectional data was also examined at the item-level with differential item functioning (DIF) analysis (Wang & Wang, 2012).

Invariance analyses at Times 1 and 2 indicated that while factor variances and loadings were equivalent across both groups, the *y*-intercepts of observed variables differed significantly between genders. Non-invariance in observed variable *y*-intercepts is called 'scalar non-invariance' (Byrne, 2012; Kline, 2011). The discovery of scalar non-invariance in the measurement model at Time 1 prompted an investigation of item-level mean differences using DIF analysis. Gender differences observed at the item-level were small and dispersed, with concentrations in measures of *Subject self-concept*, *Peer norms*, *Justifiability of cheating,* and *Self-reported cheating*, and at Time 2 also in *Performance goal structure*. The PTLC model was, based on these observations, tested separately on gender-specific data in the course of each cross-sectional analysis.

A small deficit in scalar invariance was also observed with respect to longitudinal data. Item-level differences in longitudinal DIF analysis were small and sparse across the measurement model, and concentrated primarily in contextual factors. No significant longitudinal differences in factor structure were observed with respect to *Peer norms*, *Justifiability of cheating*, or *Self-reported cheating*.

9.2.4 Demographic effects: MIMIC models and *t*-tests

Group-level differences in factor mean scores were investigated with multiple-indicator multiple-cause (MIMIC) models. It is important to emphasize that MIMIC analysis focuses on how the mean scores of overall factors vary by group, whereas invariance analysis (discussed in the previous section) is concerned with how aspects of the underlying data structure of a latent factor varies between groups. Factor means were found to differ considerably between gender-groups and grade-level groups at Time 1, but only gender at Time 2.

Mean differences by gender. Males at Time 1 had better self-concept as Science students, perceived performance goal structures as more prevalent, believed that cheating was more acceptable among their peers, believed that cheating was more justifiable in general, and reported having cheated more extensively during the preceding year. While the fact that male scores were higher for each of these measures is suggestive of a method effect in which males were simply more likely to circle higher Likert scores, significant differences running in the opposite direction on reversed items in both the *Peer norms* and *Subject self-concept* measures, suggest that the mean differences were, in fact, real. All of these gender differences were, moreover, significant again at Time 2, in addition to a significant difference in teacher quality, which males tended to evaluate more favorably. Together, these results are also consistent with prior studies that suggest cheating is more prevalent among males in self-report-based research (Whitley, 1998), and especially when they are at lower grade levels than the females to whom they are compared (Finn & Frone, 2004; Newstead, Franklyn-Stokes, & Armstead, 1996).

Cross-sectional, between-group mean differences. While a number of grade-level differences were observed at Time 1, the only one that persisted at Time 2, in *Teacher quality*, had the opposite sign, i.e. showed the opposite pattern observed at Time 1; namely the teachers received worse evaluations from older students (Grade Nine) at Time 1, but better evaluations from older students (Grade Ten) at Time 2. Teacher quality was seen most negatively, in other words, by Grade Nine students at both times. This might reflect the fact that the transition from Grade Eight to Grade Nine also entails a transition from the Middle School environment to the High School environment, wherein students tend to face both more challenging workloads and higher pressure to achieve good grades. The 'growing pains' associated with such transitions may lead to a generally more negative outlook that would affect students' perceptions of their teachers. Students transitioning from Grade Nine to Grade Ten may, by contrast, have largely adjusted to the challenges of high school, and may therefore greet their Grade Ten teachers with a more positive outlook.

These systematic differences in how students at different grade levels tended to regard their teachers were not associated with mean differences in cheating behavior (i.e. differences in factor means for *Self-reported cheating* were not observed), except when moderated by English language ability and Paternal educational attainment, which were both associated with more cheating among Grade Nine students. They did, however, follow a pattern similar to that of the reported prevalence of cheating behavior (i.e. the percentage of students who reported having cheated in the prior year), in which cheating was observed to be more prevalent among Grade Nine students at both times. Grade Nine students reported more cheating at Time 1 (57%) than their Grade Eight counterparts (48%), and more again at Time 2 (59%) than their Grade Ten counterparts (51%). The fact that Grade Nine students tended to both have lower regard for their teachers and the tendency to cheat more than their Grade Eight and Grade Ten counterparts supports the overarching PTLC hypothesis that perceived teacher quality and cheating behavior have an inversely proportional relationship, as a function of moral obligation.

Longitudinal, within-group mean differences. Within-group mean differences in the matched-samples data (N = 225) used to estimate the longitudinal model were assessed with t-tests (see Table AD1 in Appendix AD). Differences observed to have taken place longitudinally, within matched grade-level groups during the year between Time 1 and Time 2 showed two patterns: (1) a general deterioration in perceptions and behaviors over the transition from Grade Eight to Grade Nine, but (2) an overall improvement in perceptions over the transition from Grade Nine to Grade Ten (see Appendix AD). The deterioration of measures observed over the Grade Eight-to-Nine transition included an increase in perceptions of performance goal structure and a decrease in perceived teacher quality, which accompanied an increase in the incidence of cheating, from 48% to 56%. Students who transitioned from Grade Nine to Grade Ten developed, by contrast, better opinions of their teachers and class curricula, and better self-concepts as Science students, which accompanied a decrease in the incidence of cheating from 57% to 47%. These results point to a general deterioration in students' perceptions of teacher quality and academic integrity, following their transition from Grade Eight to Grade Nine, which is consistent with findings in the only published longitudinal study of the Grade Eight to Nine transition in American Schools (Anderman & Midgley, 2004). The present study is, however, the first to observe a longitudinal within-groups improvement both in students' perceptions of teacher quality and in their academic integrity following their transition to Grade Ten.

9.3 Structural hypotheses: Findings

Structural equation models allow researchers to test multiple *a priori* structural hypotheses, simultaneously. A four-component PTLC framework was developed in Chapter Three (see Figure 3.8) to unpack the implications of the overarching PTLC hypothesis, that *the degree of moral obligation that students feel to work hard and be honest in a given class context fluctuates directly with how well they think the basic obligations of teachers and classes are met in that context*. This initial model, 'Model 1', comprised seventeen first-order factors and seventy-five hypothesized structural paths. These seventy-five structural hypotheses were summarized and justified as thirty-two more general hypotheses (see Section 3.4), in accord with wide consensus that a "model under investigation must be thoroughly justified by a synthesis of the theory thought to underlie the model" (Mueller & Hancock, 2010, p. 371). When, in the various models presented over the course of this study, hypothesized structural paths were non-significant, the hypotheses they represented were not supported by the data, and *vice versa*. Individual structural hypotheses are referenced in the discussion that follows, albeit not in an exhaustive manner, so as not to distract from consideration of the primary hypothesis under consideration: the overarching PTLC hypothesis.

The numbers of variables and structural hypotheses were, in fact, reduced over the course of subsequent analyses to a set of twenty-eight structural paths hypothesized among ten first-order factors and one second-order factor (Model 4). Of the two measures of felt moral obligation originally included in Model 1, *Appropriate workload* was dropped due to weak factor structure, which prevented a direct examination of students' felt moral obligation to work hard. The PTLC hypothesis was examined principally, therefore, in terms of the moral obligation to be honest, using the scale *Justifiability of cheating*.

The central implication of the PTLC hypothesis, that context affects integrity behavior as a function of felt moral obligation, was explicitly supported by the numerous significant indirect effects of context that were conveyed to both disintegrity variables by the perceived justifiability of cheating. The PTLC model consistently explained more than 50% of the variance in *Self-reported cheating* (see Table 9.1), and is the first model found to do so in research on secondary school students (see Chapter Two, section 2.5). This may reflect that the measure *Justifiability of cheating* captured variance from both cognitive and non-cognitive processes. The PTLC model represents student judgment of moral obligation (*Justifiability of cheating*) independently of the factors hypothesized to affect it (i.e. the potential reasons why students might feel cheating is justified), such that they are related statistically by the researcher, instead of cognitively by respondents. Cognition-intensive approaches, such as asking students to identify why they cheat, run the risk of ignoring non-cognitive processes and/or evoking a 'moral dumbfounding' effect in adolescent subjects (Bjorklund et al., 2000; Cushman et al., 2010; Sneddon, 2007).

The justifiability of cheating also explained variance in *Surface learning strategies* that was comparable to what it explained in *Self-reported cheating*, thereby supporting arguments that surface learning and cheating are both forms of academic 'disintegrity' (Miller et al., 2011). *Justifiability of cheating* was, itself, well-explained by the PTLC model, wherein its most salient predictor was *Peer norms*. The structural alignment of *Peer norms*, *Justifiability of cheating*, and disintegrity behavior (*Peer norms* → *Justifiability of cheating* → Disintegrity) formed a consistent theme in analyses of the PTLC model that was most often accompanied by strong predictive paths from *Performance goal structure* and *Teacher quality*. The directionality of each of these structural connections is supported by prior experimental research reviewed in preceding chapters, with the exception of the path from *Justifiability of cheating* to *Surface learning strategies*. This chain of structural effects bearing on *Self-reported cheating* thus supports the causal hypothesis that perceived class quality and goal structure will influence cheating behavior as a function of whether students believe cheating is justifiable. Peer norms appear, moreover, to have conveyed information about appropriate in-group behaviors and, by extension, about whether a teacher is performing well.

Table 9.1

Variance explained in Self-reported cheating, Surface learning strategies, and Justifiability of cheating

	Time 1						Time 2				Long.
	Male		Female		Co-ed		Male	Female	Co-ed		Co-ed
	Ind	Comp	Ind	Comp	Ind	Comp	Comp*	Comp	Ind	Comp	Comp
Self-reported cheating	74%	76%	63%	69%	67%	72%	81%	53%	57%	63%	61%
Surface learning strategies	55%	58%	36%	47%	41%	47%	71%	58%	51%	56%	61%
Justifiability of cheating	49%	54%	62%	66%	55%	60%	36%	57%	43%	43%	49%

Note. Ind = Observed indicator variable estimation; Comp = Weighted composite score estimation

*Results may have been enlarged due to suppression effects in the Time 2 Male Sample Model (See section 7.6.2).

9.3.1 Disintegrity

Cheating and surface learning behaviors, referred to throughout the present work as 'disintegrity' (Miller et al., 2011), share in common the pursuit of grade-credentials at the expense of meaningful learning (Marton & Säljö, 1976; Schab, 1991). Effect patterns observed in PTLC structural models analyzed in chapters Six, Seven, and Eight are consistent with a body of research related to student learning theory (Biggs & Tang, 2011), in which surface learning has been associated with a number of environmental variables related to teaching and assessment quality (Biggs, 1987, 1991; Biggs et al., 2001; Biggs & Tang, 2011; Kember & Gow, 1989; Trigwell & Prosser, 1991; Wilson & Fowler, 2005). Similar or identical factors have also been associated with cheating at the secondary level (Anderman et al., 2010; Day et al., 2011; Evans & Craig, 1990a, 1990b; Murdock & Anderman, 2006; Murdock et al., 2001, 2004, 2007; Nichols & Berliner, 2007; Sisti, 2007; Strom & Strom, 2007). While surface strategies such as rote learning do not usually violate academic codes of honor, and may even be adaptive in classes where superficial learning goals are emphasized, they lack integrity inasmuch as they produce false impressions of coherent understanding (Marton & Säljö, 1976; Miller et al., 2011).

The observed moderate-to-large bivariate correlations between *Surface learning strategies* and *Self-reported cheating*, together with the fact that the variance shared by these two constructs was explained almost exclusively by *Justifiability of cheating*, suggests that the term 'disintegrity' is used advisedly to describe both cheating and surface learning. The variables *Performance goal structure* and *Teacher quality* were observed, moreover, to convey indirect effects to each outcome variable, separately, through the medium of *Justifiability of cheating* (see Appendix AF). In the co-ed model at Time 1, for instance, *Performance goal structure* exerted substantial indirect effects on *Surface learning strategies* and *Self-reported cheating* that were conveyed by *Justifiability of cheating*. Felt moral obligation appeared, as such, to act as a conduit for the effects of perceived goal structure and teacher quality on both types of disintegrity behavior. These observations suggest that cheating and surface learning are both integrity-related behaviors that are interrelated in specific contexts through felt moral obligation.

Conventional cheating. In the same sense that surface learning may, as an aspect of academic disintegrity, have moral relevance, cheating may occupy a conventional, or rules-based, domain in the minds of students (Eisenberg, 2004). The overarching hypothesis that students may sometimes sincerely believe that acts of cheating are justifiable suggests that they view such acts as violations of 'conventional' rules, instead of moral imperatives. This follows from the fact that judging an act of cheating to be justifiable, and the rules that forbid it to be morally imperative, involves a contradiction in terms. The fact that rules may be identified with either the moral domain or the conventional domain (Turiel, 1983, 2002; 2006) implies that disintegrity may entail two types of cheating: (1) cheating as a moral violation and (2) cheating as a conventional violation. The large amount of variance explained in *Self-reported cheating* by *Justifiability of cheating*, discussed in the next section, indicates that many acts of cheating in secondary school are viewed by students as conventional violations.

Self-reported cheating. Self-reported cheating at Time 1 had no direct effect on self-reported cheating at Time 2. In the longitudinal structural model, which included autoregressive paths that represented variance explained by Time 1 measures in their Time 2 counterparts, the autoregressive path for this longitudinal relationship was non-significant. The moderate bivariate correlation between the two measures of *Self-reported cheating* ($r = .440$, $p < .001$) appears, therefore, to have been the result of associations with variance in third-variables that carried-over longitudinally. The correlation between Cheating at Times 1 and 2 appeared, in other words, rooted in how individual students tended to perceive Science class contexts at Times 1 and 2. That is not to say that cheating was driven by context, *per se*, but by *perceptions* of context that carried over from Time 1 to Time 2 due potentially to environmental similarities as well as context-free tendencies among students to perceive certain factors in certain ways. This observation suggests that direct effects of past cheating on future cheating observed in cross-sectional research (e.g. Davis & Ludvigson, 1995; Harding et al., 2012; Mayhew et al., 2009; Passow et al., 2006; Whitley, 1998) may, in fact, be artifacts of how individual students tend to perceive themselves and class contexts.

Self-reported cheating in the Time 2 iteration of the longitudinal model reflected, in particular, the degree to which students felt that cheating was justifiable. This was indicated by amounts of variance in excess of 50% that were consistently explained in *Self-reported cheating* by large predictive effects of *Justifiability of cheating*. A significant amount of variance in the Time 2 measure of *Justifiability of cheating* was explained, in turn, by its counterpart measure at Time 1, as well as by a number of indirect effects from self-concept and teaching context measures at Time 1. Many of these Time 1 variables also exerted significant indirect effects on *Self-reported cheating* at Time 2 (see Appendix AF). Past cheating (Time 1) appeared to share variance with future cheating (Time 2), therefore, as function of third-variable associations channeled through felt moral obligation.

Perceptions of self and context at Time 1, including self-concept in relation to both Science ability and honesty, cheating-related peer norms, performance goal structures, and teacher quality exerted distinguishable indirect effects on *Self-reported cheating* at Time 2, firstly as a function of the variance they explained in their counterpart measures at Time 2, and secondly in the degree to which their Time 2 counterparts carried their effects over to *Justifiability of cheating* (see Appendix AF). This suggests that while schools often direct a great deal of attention to the immediate problem of cheating acts, through honor policies, monitoring, and punishment, patterns of cheating behavior over time appeared, in the present study, to be rooted ultimately in how individual students tended to perceive themselves and their academic contexts.

Surface learning strategies. This study joins a small body of research on the empirical relationship between academic integrity and learning strategy. Research in this area to date indicates that students who employ deep learning strategies, which are oriented to the substance of learning, are less likely to cheat (Anderman et al., 1998; Bong, 2008).

Justifiability of cheating was the strongest predictor of *Surface learning strategies* in every iteration of the PTLC model, with the exception of the female sample model at Time 2. The substantial relationship between *Justifiability of cheating* and *Surface learning strategies* is interpreted here in the light of how *Justifiability of cheating* is believed to relate to *Self-reported cheating*. The surprising strength and consistency of this relationship suggests that students frequently engage in surface learning while understanding that it essentially cheats the duty to learn with integrity, i.e. to strive to understand and master course material. Instead of cheating the 'letter of the law' with respect to integrity, however, surface learning cheats its spirit.

The large effects that *Justifiability of cheating* exerted on *Surface learning strategies* in cross-sectional models were observed to convey a number of indirect effects, such as those exerted by students' perceptions of Science class contexts (see Appendix AF). This was also the case in the longitudinal model, where the Time 2 measure of *Justifiability of cheating* conveyed indirect effects of perceived context at Time 1. Thus, while the use of surface learning strategies at Time 2 was predicted by surface learning behaviors at Time 1, the otherwise large bivariate correlation between these variables (r = .632; see Table 8.2) appears to have largely reflected the effects of longitudinal continuity in third-variables, i.e. students tending to perceive themselves and their Science classes in ways that influenced whether they studied in a surface manner.

9.3.2 Justifiability of cheating

Justifiability of cheating was treated in the present study as a measure of moral obligation that, by querying the end results of students' judgment processes, i.e. the judgments they held at the time of the questionnaire, should have captured variance from both cognitive and non-cognitive mental processes (Haidt, 2001; Kahneman, 2011; Machery & Mallon, 2010), which are both associated with social contract-based judgment (Fiddick et al., 2005). The measure *Justifiability of cheating* asked only for students' final judgments of whether cheating was felt to be 'justifiable' or 'reasonable' in their Science class. Whatever cognitive and non-cognitive processes were involved in students' judgment of this issue would have come to a conclusion, even if temporary, before students could report their judgments on the questionnaire. Experimental studies indicate that individuals tend to judge social contract violations as 'cheating' (Cosmides, 1989; Cosmides & Tooby, 2013). This broader conception of cheating may be ascribed, from a student point of view, to laziness, incompetence, injustice, or irresponsibility in a teacher. Such judgments are held to occur by rapid-fire, non-cognitive processes, and often to lead to negative reciprocation by those who feel cheated (Boles et al., 2000; Gneezy, 2005; Pillulta & Murningham, 1996).

Prior psychometric studies of academic cheating have generally been situated, by contrast, within the rational-cognitive paradigm in cheating psychology. The tendency in such studies to focus on students' explicit reasons and intentions for cheating may have often obfuscated variance from non-cognitive processes. Qualitative studies of cheating typically focus, for instance, on getting students to explain *why* they cheat (e.g. Stephens & Nicholson, 2008; Zito & McQuillan, 2011), which is likely to evoke moral dumbfounding among adolescents who are not yet ready to articulate their non-cognitive moral understandings

(Bjorklund et al., 2000; Sneddon, 2007). Multivariate models derived from the theory of planned behavior (TPB), which have been applied widely in cheating research at the tertiary level, hinge, by contrast, on measures of *Intention to cheat* (Beck & Ajzen, 1991; Mayhew et al., 2009; Whitley, 1998). While TPB research has been moving in the direction of the dual-process paradigm factors for at least ten years (Harding, Carpenter, Finelli, & Passow, 2004; Harding et al., 2012; Passow et al., 2006), *Intention to cheat*, which emphasizes cognitive processes, such as planning and premeditation (Beck & Ajzen, 1991; Simkin & McLeod, 2010), remains at the center of the TPB model. The only two prior studies of academic cheating to have measured explicitly non-cognitive factors (McTernan et al., 2014; Murdock et al., 2008) did so in terms individual proneness to emotions such as shame and guilt, posed as personality constructs, within a decision-making framework.

The overarching PTLC hypothesis that student perceptions of class quality drive disintegrity behaviors as a function of moral judgment was very clearly supported by the significant indirect effects of *Teacher quality* that were conveyed by *Justifiability of cheating* to both types of disintegrity. Perceiving that teachers were 'worse' affected moral judgment negatively (in favor of cheating), and appeared, thereby, to indirectly cause disintegrity to increase. The prominent mediating role played by *Justifiability of cheating* with respect to contextual effects supported the PTLC hypothesis in that it represented the degree to which students felt obliged to be honest according to whether they thought their teachers met the obligations entailed by teaching 'well'.

9.3.3 Peer norms

The hypothesis that *Peer norms* would affect cheating behavior directly was not supported. Instead the mediated structural sequence involving '*Peer norms* → *Justifiability of cheating* → *Self-reported cheating*' was a central and significant theme in all PTLC models tested. The strength and consistency of this structural sequence is consonant with prior research indicating that cheating-related peer norms influence whether individuals believe that cheating is an appropriate in-group behavior (Eisenberg, 2004; Galloway, 2012; Gino et al., 2009; Hartshorne & May, 1928; McCabe, Treviño, & Butterfield, 2001; Steiner, 1930; Schab, 1980; Walker et al., 1966). The fact that *Peer norms* tended to mediate the effects of individuals' perceptions of teacher quality on the justifiability of cheating in PTLC models is consistent, moreover, with the assertion of social comparison theory (Broeckelman-Post, 2008; Festinger, 1954; Koul et al., 2009; Nora & Zhang, 2010) that individuals develop judgments in conformity with the judgments they perceive among their peers. The validity of this causal sequence, in which *Peer norms* conveys variance from *Teacher quality* to *Justifiability of cheating*, was supported by strong direct and indirect beta coefficients in all models tested, except the female sample model at Time 2.

The decision to position *Peer norms* as a mediator of the effects of class quality on the justifiability of cheating reflected the hypothesis that rules against cheating are viewed as morally imperative only insomuch as the context within which those rules exist is viewed as morally legitimate. By communicating the social acceptability of cheating to individual students, peer norms were expected to convey group-level evaluations of class quality that would shape individual evaluations of the same. The variable *Peer norms* was thus expected to mediate the variance explained by *Teacher quality* in *Justifiability of cheating*.

It is important to note that the structural relations observed among *Teacher quality*, *Peer norms*, and *Justifiability of cheating* do not appear to be the result of multicollinearity. The complete mediation at Time 2 by *Peer norms* of the effect of *Teacher quality* on *Justifiability of cheating* was accompanied, for instance, by moderate correlations between *Peer norms* and both *Justifiability of cheating* (r = .471) and *Teacher quality* (r = -.473), in addition to a slightly lower correlation between *Teacher quality* and *Justifiability of cheating* (r = -.388). These correlations were similar at Time 1, with the exception of that between *Peer norms* and

Justifiability of cheating (r = .600). While a correlation of .600 is undoubtedly strong, it still falls well below levels commonly used to demarcate multicollinearity, such as r > .800 (Field, 2009). At Time 1, moreover, *Peer norms* served as only a partial mediator at Time 1, such that its stronger association with *Justifiability of cheating* did not appear to affect the degree to which it mediated the effect of *Teacher quality* on *Justifiability of cheating*. *Peer norms* was also re-positioned in 'equivalent models' (Kline, 2011) as a correlate (see Appendices Q and AB) and then as a predictor (see Appendices R and AC) of class quality variables. None of these equivalent models were significantly better in overall model fit than the hypothesized PTLC model in which *Peer norms* was a mediated of class context on downstream variables (see Figures 6.4 and 7.3, respectively). These observations, together with the significant indirect effects of *Teacher quality* conveyed by *Peer norms* to *Justifiability of cheating* in the PTLC model at Time 1 (β = -.150, p < .001), Time 2 (β = .121, p < .01), and in the longitudinal PTLC model (see Appendix AF), suggest that *Peer norms* is a lynchpin in the causal sequence 'Class quality → *Peer norms* → *Justifiability of cheating* → *Self-reported cheating*'.

Longitudinal peer effects. A substantial autoregressive path coefficient in the longitudinal model suggested that peer effects also carried over from year to year, regardless of class context, and may, as such, have prejudiced students' perceptions of teacher quality and of whether cheating is justifiable, before they ever entered the classroom. Such carryover may have emanated from students' personal tendencies to perceive higher levels of cheating, or from continuity in their peer group associations between years. Higher levels of perceived peer support for cheating may also have emanated from variables outside of class contexts, such as school culture or national culture (Crittenden et al., 2009a; Magnus et al., 2008; Teixeira et al., 2010), that were not included in Model 4.

9.3.4 Academic context quality

The measure of *Good teaching* from the *Course Experience Questionnaire* (CEQ), which has been used to study associations between student experience and learning behavior at universities around the world for more than three decades (Barnhardt & Ginns, 2014), was used in the present study to gather data on student evaluations of class quality. This adds to a very small literature of student learning theory (SLT) research conducted in secondary schools (Ginns, Martin, & Papworth, 2014; Selmes, 1986), and an even smaller literature using SLT course experience measures at the secondary level (Ramsden et al., 1988, 1989). The present work also joins a small number of studies that have used CEQ measures to research academic cheating (Jurdi et al., 2011a, 2011b; Norton et al., 2001), and is the first study found to have done so at the secondary-level.

Academic context quality was represented in the PTLC model principally in terms of the second-order factor *Teacher quality*, which included two first-order factors: *Good teaching* and *Assessment quality*. *Teacher quality* is conceptually similar to several teacher-related measures found to negatively predict self-reported cheating and pro-cheating attitudes in prior studies, including *Teacher commitment* (Murdock et al., 2001), *Teacher caring* and *competence* (Murdock et al., 2004), and *Teacher credibility* (Anderman et al., 2010). The hypothesis that student evaluations of class quality affect self-reported cheating behavior directly was not supported by PTLC model results. The effect of *Teacher quality* on *Self-reported cheating* was mediated, instead, by *Peer norms* and *Justifiability of cheating*.

An important function served by *Teacher quality* in Model 4 was to completely mediate the effects of *Honesty-trustworthiness self-concept*. This relationship, theorized to represent cynicism (discussed further in section 9.3.6), transmitted a number of significant effects downstream in the model, especially to *Peer norms* and *Justifiability of cheating*, both in the Time 2 cross-sectional model and the longitudinal model.

While *Usefulness of curriculum*, a proxy measure for students' overall interest in a class, remained psychometrically distinct from the second-order factor *Teacher quality*, none of the often substantial bivariate correlations it shared with other factors manifested as a significant beta coefficient in any structural model tested. Stepwise regression analysis conducted with respect to co-ed models at Times 1 and 2 indicated that the effects of *Usefulness of Curriculum* were muted by the presence of *Teacher quality* (see Appendices N and Y). The large correlation between these two variables appeared to account for the variance that *Usefulness of curriculum* shared with downstream constructs. These observations led to the estimation of equivalent models at Times 1 and 2 that generally supported the possibility that *Teacher quality* could be modeled as a mediator of *Usefulness of curriculum*. It was not clear that the relationship should be unidirectional, however, and no clear theoretical or empirical support for modeling *Teacher quality* as a mediator of *Usefulness of curriculum* (as opposed to the other way around) was discovered in the literature.

9.3.5 Performance goal structure

Inclusion of *Performance goal structure* in the PTLC model extends a large and growing body of work on cheating in the field of achievement goal theory (AGT) (e.g. Bong, 2008; Brown-Wright et al., 2012; Murdock & Anderman, 2006; Yang, Huang & Chen, 2013). While the results of these studies have generally suggested that an orientation to personal performance goals is associated with higher levels of cheating, findings related to performance goal structures have been mixed (Anderman et al., 1998, 2010; Anderman & Midgley, 2001, 2004; Bong, 2008; Tas & Tekkaya, 2010; Anderman et al., 1998; Stephens & Gehlbach, 2007). Results of the present study are consistent with the findings of Anderman et al. (1998) that the perception of a performance goal structure in Science class is associated with higher levels of cheating. This finding is also consistent with experimental studies in which cheating has been caused by key aspects of a classroom performance structure, such as extrinsic motivation (Lobel & Levanon, 1988; Mills, 1958) and knowledge of peer performance (Shelton & Hill, 1969; Taylor & Lewit, 1966). The present work also provides evidence that performance goal structures predict surface learning strategies among secondary school students.

Performance goal structure was a significant predictor of *Justifiability of cheating* in every model tested, and was its strongest predictor in the Time 2 iteration of the longitudinal model. Its predictive effects on *Justifiability of cheating* were, in all of these models, conveyed to *Surface learning strategies* and *Self-reported cheating* in roughly equal portions. In the longitudinal model, a significant amount of variance in the tendency to perceive performance goals at Time 2 was carried over from Time 1. This might reflect variance from either the broader school environment or from students' personal achievement goal orientations. Achievement goal emphases are seen, for instance, to operate at three levels in schools: (1) students' personal goal orientations, (2) classroom goal structures, and (3) school goal structures (Anderman & Midgley, 1997). What goals students perceive to be emphasized in a given class may, therefore, reflect variance at the individual or school levels. The fact that such perceptions were more prevalent among male respondents at both times might, additionally, reflect variance from personological factors.

While respondents closely associated a mastery goal structure with teacher quality that was 'good', as evidenced by high multicollinearity between measures of *Mastery goal structure* and *Good teaching* in the Pilot Study, respondents did not associate a performance goal structure with teaching that was 'bad'. While perceiving a performance goal structure predicted the justifiability of cheating in every model, it does not appear to have done so as an aspect of class quality. Performance goal structures may, instead, have oriented student efforts to 'earning' grades at the expense of learning by communicating "messages about the purposes of instruction" that "emphasize the demonstration of ability and competing favorably with others as the main reasons for engaging in academic work" (Anderman & Midgley, 2004). The pursuit of grades at the expense of learning is a key conceptual similarity between surface learning and cheating that could explain why *Performance goal structure* indirectly predicted both disintegrity variables, as a function of *Justifiability of cheating*. Instead of affecting the degree of obligation students felt to meet a teacher's expectations, performance goal structures may have shifted what the expectations were, from learning priorities to grade priorities.

9.3.6 Self-concept

Self-concept measures were included in the 'Person' component of the PTLC model, based on a broad body of evidence that personological factors explain substantial amounts of variance in ethical behavior (e.g. Burton, 1963; Gino et al., 2011; Miller et al., 2007; Whitley, 1998). Measures of self-concept account for how individuals evaluate themselves in terms of various aspects of self-image, based on their perceptions of past performance and social comparison in those areas (Bong & Clark, 1999; Bong & Skaalvik, 2003). The two self-concept measures chosen for this study, *Honesty-trustworthiness self-concept* and *Subject self-concept* proved to be highly consistent over time, having the largest autoregressive path coefficients in the longitudinal model. Each Time 1 measure explained more than 50% of the variance in its Time 2 counterpart, yet the two types of self-concept shared only small bivariate correlations with one another (*r*'s = .206 - .212; see Table 8.2). These findings add to a small body of prior research in which self-concept has been measured in relation to academic integrity (Arvidson, 2004; Rost & Wild, 1994).

A notable difference in how self-concept factors performed at Time 1 *versus* Time 2 was the disappearance of their direct effects on outcome variables at the latter time. Direct effects of both self-concept constructs on integrity-related factors at Time 2 were overwhelmingly mediated by *Teacher quality*. This consolidation of personological effects through mediator variables at Time 2 may have been a maturation effect (Kline, 2009), stemming from the higher cognitive functioning of students who were one year older at Time 2, an effect of having transitioned to higher grade levels at Time 2, or a Hawthorne effect (Kline, 2009), in which students who participated in this project at Time 1 were stimulated to reflect more carefully on themes of honesty, justice, and cheating during the year leading up to Time 2 data collection.

Honesty-trustworthiness self-concept. Two effect patterns were exhibited by *Honesty-trustworthiness self-concept*. At Time 1, it exerted significant inverse effects, both direct and indirect, on all three downstream variables related to cheating (*Peer norms*, *Justifiability of cheating*, and *Self-reported cheating*). In the Time 2 co-ed model, by contrast, all effects of this factor were mediated by *Teacher quality*. This difference between cross-sectional models was also observed between iterations of the longitudinal model.

While these results are weaker than anticipated based on prior experimental studies in which moral self-concept has been observed to substantially influence dishonest behavior (Ariely, 2012; Gino et al., 2011; Mazar et al., 2008; Shalvi et al., 2011), the strongest downstream effects exerted by *Honesty-trustworthiness self-concept* were, in the co-ed model at Time 1 and in both gender-specific models at Time 2, associated with *Justifiability of cheating*. This pattern suggests that concern for maintaining a positive moral self-concept often weighed on students' judgments of whether cheating was justifiable. These effects were not large, however. The 'moral motivation' that has been associated with felt moral obligation (Schroeder et al., 2010) appeared, instead, to relate principally to students' perceptions of context.

Mediation of *Honesty-trustworthiness self-concept* by *Teacher quality* at Time 2 was interpreted to suggest cynicism (Mills & Keil, 2005), in that students who rated themselves as more dishonest were more likely to view their teachers in the same light, and *vice versa*. This mediation effect was first observed in the female sample model at Time 2, and found, subsequently, to carry-over to the Time 2 co-ed model, likely due to the larger proportion of females in the Time 2 sample (61%). This finding is consistent with statistical associations observed in prior research between *Honesty-trustworthiness self-concept* and perceived teacher quality (Hay, 2000; Martin et al., 2006), and with evidence that cynicism may be more prevalent among females, and may increase with age during adolescence (Galbraith & Merrill, 2012; Simon et al., 2004).

Subject self-concept. The measure of *Subject self-concept* used in the present study was related specifically to the subject of Science, in which more cheating has often been observed than in other subject areas at the secondary level (Miller et al., 2007; Murdock et al., 2001; Schab, 1991). The positive predictive effects of *Subject self-concept* on both types of disintegrity, observed in many of the models tested, were consistent with strong evidence that ability-related self-beliefs, such as self-efficacy, are inverse predictors of cheating among secondary school students (Bong, 2008; Finn & Frone, 2004; Lee et al., 2014; Murdock et al., 2004; Nora & Zhang, 2010; Tas & Tekkaya, 2010). These associations were, however, stronger with *Surface learning strategies* than with *Self-reported cheating*.

Subject self-concept was also a strong predictor of students' perceptions of class context factors, implying that how students perceive their ability in a particular subject area informs how positively they evaluate classes in that subject area. While the strength of these effects was similar for males and females at both time points, females tended to report lower self-concept related to Science and correspondingly worse evaluations of their Science teachers.

9.3.7 Differences between gender-specific models

Several differences between gender-specific structural models were observed at both Times 1 and 2, of which all were related to whether, and to what degree, males and females judged cheating to be justifiable. Among males, for instance, *Performance goal structure* was a stronger predictor of *Justifiability of cheating*, which was, in turn, a stronger predictor of *Surface learning strategies*. Among females, by contrast, *Peer norms* was a stronger predictor of *Justifiability of cheating* than it was among males, yet the function of *Peer norms* as a complete mediator of *Teacher quality* on *Justifiability of cheating* was observed exclusively among males. *Teacher quality* exerted a direct effect on *Justifiability of cheating* in both female sample models. The effect of perceived peer norms on the justifiability of cheating was more prominent among females, in other words, but played a greater role shaping the effect of teacher quality on the justifiability of cheating among males.

A notable feature of the female sample model at Time 2 was the failure of all hypothesized downstream effects of *Honesty-trustworthiness self-concept* to achieve statistical significance. *Honesty-trustworthiness self-concept* was found, instead, to exert indirect effects on downstream constructs as a function of *Teacher quality*. Such effects in female sample models generally carried over to co-ed models to a greater extent than did those of male sample models, due to larger proportions of female participants at Time 1 (59%), Time 2 (61%), and in the longitudinal sample (68%).

9.4 The PTLC hypothesis: Reducing the scope of the 'belief-behavior incongruity'

The results of this study suggest that mainstream contemporary scholarship tends to over-subscribe to the BBI and neutralization perspectives on academic cheating. The fact that the BBI is inconsistent with what rational-cognitive theories of moral psychology predict has been held to require the neutralization framework as a means of bridging the incongruity, or what Blasi (1983) described as the "fracture within the very core of the self" (p. 201). The neutralization perspective entails, however, as argued in section 2.6, the problematic assertion that students who cheat overcome the internal discomforts associated with the BBI by intentionally deceiving themselves.

A central assertion of the dual-process paradigm of moral psychology, that moral judgment is only partially rational (Cushman et al., 2010; Haidt, 2001, 2007), eliminates the need to assume that adolescents' judge academic cheating according to abstract ethical beliefs. Cheating is, on its face, a much broader concept than academic cheating. All forms of social contract violation can, for instance, be described as cheating (Cosmides & Tooby, 2013). By this broader view, students who believe they are being mistreated or neglected, i.e. 'cheated', in specific academic contexts may cease to hold the view that rules mandating honesty are morally imperative within those contexts, because the social contract that legitimates those rules is no longer in force (Rettinger, 2007). While this view may be entirely irrational on a student's part, identifying such rules with the conventional domain, instead of with the moral one (Eisenberg, 2004; Turiel, 1983, 2002, 2006), may still involve sincerely felt judgment that reflects genuine perceptions and beliefs, as opposed to techniques of neutralization.

From a social contract perspective, cheating can be viewed as wrong in a broad ethical sense, even while violating rules that mandate honesty may be viewed as morally permissible under certain circumstances. Such reasoning is exemplified by the concept of criminal defense in Anglo-American legal systems. A criminal defense may concede that a law has been broken, but maintain that the infraction was not criminal due to extenuating circumstances (Morawetz, 1986). Such defenses can exonerate individuals who have broken laws, by highlighting for a judge or jury the special circumstances that led to the violation. Criminal defenses are, in other words, taken seriously in courtrooms, whereas similar defenses offered by students in studies of cheating are uniformly dismissed as deceptions. While schools should, indeed, not let students 'off the hook' who break rules that mandate honesty, the possibility that students may sincerely believe that their justifications for cheating are valid is of value to scholarship on the psychology of cheating.

Results of the present study indicate that a substantial proportion of cheating is done under the belief that external circumstances have, in fact, rendered it a conventional transgression. The belief appears to be sincere, in other words, as opposed to an intentional self-deception, in that it reflects lower moral obligation. This can be stated with a healthy measure of confidence because the linkages in the PTLC model between (1) whether cheating was viewed as justifiable, and (2) the circumstances hypothesized to lead students to view it that way, i.e. poor teaching and widespread peer cheating, were inferred statistically. They could not, in other words, have been distortions, lies, or neutralization techniques offered up by participants in the study. These results are also consistent with experimental studies in which negative perceptions of class context have been found to cause students to judge that cheating is more justifiable (Murdock et al., 2004, 2007). Cheating that students judge to be justifiable cannot also violate their moral beliefs. Such cheating, described in the present work as 'cheating as a conventional violation', would not, therefore, cause BBI.

Results of the present study strongly endorse the hypothesis that cheating is often viewed by students as a conventional violation. This does not suggest, however, that cheating never induces BBI. There is strong evidence to suggest that many acts of cheating are felt by cheaters to be moral violations (e.g. Jensen et al., 2002; Stephens & Nicholson, 2008). The PTLC models tested in the present study explained, on average, roughly two-thirds of the variance in self-reported cheating, of which the majority was explained by the justifiability of cheating. The remaining third of variance in self-reported cheating that was not explained by the justifiability of cheating may reflect cheating as a moral violation. This implies that at least one-third of the cheating reported in this study may have been incongruent with respondents' moral beliefs.

9.5 Implications for educators: Addressing the problem of disintegrity

Many educators care dearly about student honesty, although the degree of this caring is often shown most clearly in the punishments meted out to cheaters, whether involving zeroes on major assignments or expulsions from school. Merit-based educational systems rely on integrity, lest they become mere fronts for otherwise meaningless activity. Cheating is a systemic threat, and the system tends to react vigorously. The focus on punishment and threats of punishment tends, however, to push solely in the direction of behavioral conformity, while ignoring moral maturity. The present study indicates that schools and educators should augment the emphasis they place on honor codes and punitive regimes with an emphasis on reducing the degree to which students judge cheating to be justifiable. Reducing the justifiability of cheating should, according to any of the PTLC models tested herein, be associated with a direct reduction in cheating behavior. While attempts to reduce cheating by improving students' moral judgment have been made many times, they have uniformly been made from a rational-cognitive point of view and have generally had small to negligible effects on cheating behavior. Examples include 'ethics education' (Bebeau & Thoma, 1999; Evans & Craig, 1990a), 'ethical philosophy programming' (Seider et al., 2013, p. 7), 'cognitive dissonance interventions' (Vinski & Tryon, 2009), and 'cognitive inoculation treatments' (Compton & Pfau, 2008, p. 104). The small to negligible effects of such programs on actual cheating behavior (e.g. Houston, 1983b, Sieder et al., 2012; Spear & Miller, 2012; Tittle & Rowe, 1973; Vinski & Tryon, 2009) are, in fact, consistent with the tenuous relationship between moral cognition and moral behavior (Blasi, 1980).

While the issue of how to manage non-cognitive moral processes has not been dealt with previously in literature on academic integrity, three general suggestions for how individuals can take control of their own non-cognitive impulses are offered by Haidt (2007):

> "We can [1] use conscious verbal reasoning, such as considering the costs and benefits of each course of action… [2] reframe a situation and see a new angle or consequence, thereby triggering a second flash of intuition that may compete with the first. And [3] we can talk with people who raise new arguments which then trigger in us new flashes of intuition followed by various kinds of reasoning" (p.999).

These three suggestions are framed in terms of curbing impulses that have already occurred within an individual. Stepping back, however, they also suggest preventative approaches that can be implemented at a classroom or school level. Cognition-intensive programs and moral appeals intended to improve adolescent ethics reflect the first and third of these suggestions by, for instance, focusing on moral reasoning skills in order to strengthen their abilities to cope with non-cognitive impulses, and by presenting arguments designed to counter likely justifications for feeling that it is okay to cheat. Such programs focus, in essence, on improving cognitive control. The present research indicates that educators should also attend to Haidt's (2007) second suggestion, i.e. reframing situations, by managing student perceptions of at least three factors: (1) the quality of educator skill and caring, (2) the intended goals of learning, and (3) the students' own morality.

9.5.1 Enforcing rules

Ultimatum game studies show that people are often inclined to punish 'cheaters', even when doing so is damaging to themselves, thus foregoing their rational self-interest for the sake of moral ideals (e.g. Boles et al., 2000; Gneezy, 2005; Haidt, 2001; Pillutla & Murnighan, 1996). This '*Homo moralis* view' of human nature (Haidt, 2007), which is at the crux of the PTLC hypothesis, appears to cut in two directions with respect to cheating in schools. On one hand, deeming rules that forbid academic cheating to be merely conventional may be a non-rational response by students to the perception that teachers are 'cheating' on their professional responsibilities; on the other hand, the emphasis that school honor policies tend to place on punitive consequences may also be imbued with a tendency for non-rational responses by educators to cheating done by students. Research on cheating does suggest that careful monitoring and strict consequence systems are needed to prevent widespread cheating in connection to 'contagion effects' (Gino et al., 2009; Walker et al., 1966). But if students are cheating in response to perceptions, as this and many other studies suggest is often the case, then perception management is another, and perhaps better, means of addressing the problem.

Rules, monitoring, and consequence systems are important to reducing the incidence of cheating, because they increase the risk of cheating, and thereby decrease its prevalence under all circumstances. Research shows that the incidence of cheating tends to be lower when honor policies are consistently enforced and clearly understood by students (Burrus et al., 2007; McCabe & Katz, 2009). But the punitive regimes that typically accompany such policies have two major weaknesses: (1) they address only the symptoms of the problem, and (2) most of the cheating that goes on in schools is never detected.

9.5.2 Managing perceptions

Perceptions of poor teacher quality, widespread peer cheating, and performance goal structures have been shown in experimental and non-experimental studies to lead to higher levels of cheating. The amount of concern individuals have for their moral self-concept, which involves perceiving themselves as moral people, has also been found experimentally and non-experimentally to cause systematic variation in cheating behavior. These studies were reviewed extensively in preceding chapters (e.g. see sections 2.3.3 and 5.1), and will not be reviewed again here.

Perceptions of teachers. Nowhere in this study is it proposed that students' perceptions of 'poor' teacher quality are objectively valid, or that two wrongs ever make a right. PTLC model results demonstrate that, to the contrary, students' perceptions of teachers are strongly associated with their subject self-concept, honesty-trustworthiness self-concept, and level of interest in class curricula. Substantial variance was also observed to carry over from measures of *Teacher quality* at Time 1 to its counterpart measure at Time 2, which suggests that students were exhibiting developed tendencies for liking Science teachers more or less. Teachers in this study faced factors, therefore, that were initially beyond their control, and that shaped how students perceived them from the beginning of the school year.

The present research is consistent with evidence from prior studies that suggests teachers can reduce cheating in their classes by actively shaping students' perceptions of themselves as competent and respectful (Murdock et al., 2008), caring (Murdock et al., 2004), and credible (Anderman et al., 2010). This does not mean that teachers should necessarily change how they teach and assess students, but that they should strive to shape how students *perceive* them. Teachers may, for instance, find appropriate ways to tout their own honesty and integrity, exploit opportunities to convey their caring for students, demonstrate their effort, communicate their strengths in terms of background training, knowledge, and experience, defend their particular approaches to pedagogy and assessment, and uphold (and enforce) the principle of mutual respect in their classes (Anderman et al., 2010; McCrosky & Young, 1981; Progue & Ah-Yun, 2006; Selman, 1980). Teachers should additionally look to their school administrators for public support of their efforts to make positive impressions on students, such as by touting their personal and professional accomplishments, hard work, caring, and integrity in front of students. The present research suggests that improving students' perceptions of teachers should reduce cheating by increasing the moral obligation they feel to respect the teachers' rules.

Learning goals. Disintegrity among students in the present study was associated with learning goals in two senses: (1) whether teachers emphasized performance goals, which appear more likely to orient students to grades over learning, and (2) whether the learning objectives emphasized on assessments were authentic and transparent. These findings, together with large amounts of prior research reviewed in Chapter Two, indicate that teachers and institutions should heavily emphasize mastery goals, and strive to design assessment tasks around purposes that are authentic and clearly understandable.

A practical reality faced by many Western educators is that performance goals are inherent to the educational systems within which they work. It is not enough, therefore, to recommend that teachers emphasize mastery goals. Grades are important, and popular curricular programs such as Advanced Placement (AP) and the International Baccalaureate (IB) entail assessment programs that explicitly rank individual performance. To the extent that performance emphases are inevitable, teachers should actively shun performance-avoidance goals in favor of performance-approach goals. While promoting performance goals at all is generally discouraged (Brophy, 2005), performance-approach goals have had weak-to-negligible associations with cheating behavior in published research (Bong, 2008; Niiya et al., 2008; Tas & Tekkaya, 2010), and appear to be of greatest concern in academically advanced classes (Stephens & Gehlback, 2007; Taylor et al., 2002). Inasmuch as promoting performance-approach goals in class is necessary, the messages they send that emphasize grades and peer comparisons over learning should be counterbalanced with strong, clear emphasis on mastery goals. The two types of goal are not mutually-exclusive. We can, indeed, strive to perform well by mastering material, even on a normative or competitive basis.

Design-it-out. Teachers can also counteract the messages that performance goals send with assessment design. A recent and growing body of scholarship on how cheating can be 'designed-out' of academic contexts has highlighted the need for assessment tasks that are uniquely personalized and meaningful, such that (1) the intrinsic value of what is to be assessed can compete with the extrinsic value that students ascribe to grades and peer comparisons (Heckler et al., 2013; Howard & Davies, 2009; Sisti, 2007), and (2) the opportunity to cheat is limited or eliminated by the design, itself (Carroll & Appleton, 2001; Gannon-Leary et al., 2009). If surface learning strategies are, as the present research suggests, tied to justifications of cheating, meaningless work that calls for surface-level learning may, in fact, lead students to believe cheating is justified.

Many principles for how to design-out cheating are also consistent with the design principles of 'constructive alignment', which Biggs (1996, 1999) developed in conjunction with the 3-P Model (see Figures 3.1 and 3.6). Students are held by the constructive alignment perspective to adopt surface or deep approaches to learning in response to "the messages [they] receive; and again, assessment becomes the chief source of these messages" (Biggs, 1991, p. 26). The adoption of surface learning goals often reflects a failure, Biggs (1996) argues, to constructively align the components of the 3-P Model, "so that [1] objectives express the kinds of understanding that we want from students, [2] the teaching context encourages students to undertake the learning activities likely to achieve those understandings, and [3] the assessment tasks tell students what activities are required of them" (p. 57). Designing courses according to the principles of constructive alignment is intended to 'entrap' students in the dynamics of healthy constructivist learning by emphasizing the process of learning over its end-products (i.e. grades), and by structurally limiting opportunities to engage in surface learning.

Grouping surface learning behavior together with cheating, under the rubric 'disintegrity', highlights the common goal of both the 'design-it-out' and 'constructive alignment' conceptions of good assessment and course design: to channel students' efforts into processes of learning with integrity. Both conceptions aim to inspire interest in the substance and meaning of what is being learned, to facilitate the construction of personal understandings and skills, and to structurally improve the odds that students will exhibit integrity in their academic work.

Self-perceptions. The third type of student perception that educators should strive to manage is moral self-concept. High moral self-concept has been shown to act as a buffer against the temptation to cheat in experimental studies (Gino et al., 2011; Mazar et al., 2008), and was, in keeping with such findings, observed in the present study to inversely predict whether students judged cheating to be justifiable. Recent research by Bryan et al. (2013) suggests that educators can manage students' moral self-concepts by, in effect, casting a spotlight on them.

Bryan et al. (2013) found that asking participants in an experimental study on cheating to 'not be cheaters' reduced cheating behavior far more than asking them 'not to cheat'. The difference was subtle but powerful. Asking participants not to cheat focused attention on the action of cheating, whereas asking them to not *be* cheaters focused attention on their moral 'beings', i.e. their moral self-concepts. It is not known how well this technique works with adolescents, or whether it loses effect if repeated frequently over long periods of time. When educators talk to students about academic integrity, however, a sparing deployment of the request to "not *be* cheaters", and reminders that 'we *are* not cheaters', may help reduce cheating by bringing focus to bear on students' moral identities.

9.5.3 Focusing on the Grade Eight – Grade Nine transition

A final recommendation for addressing cheating at the secondary school level is to pay special attention to the Grade Eight-to-Nine transition. This is especially important at schools in which Grade Eight is incorporated into a middle school division that is segregated from the high school division. Results in the present study are consistent with the only other longitudinal study of cheating over this transition (Anderman & Midgley, 2004). Students who made the transition from Grade Eight to Grade Nine during the research reported herein reported a higher incidence of cheating at the latter time (57%) than at the former (48%). The longitudinal nature of this study and the study conducted by Anderman and Midgley (2004) strengthens the basis for believing that the observed increase was caused by factors related to the grade-level transition. Policy initiatives aimed at reducing cheating would be well advised, therefore, to prioritize rising Ninth Graders.

9.6 A bad idea: Literal teacher-learner contracts

'Contract' is used in the PTLC framework as a metaphorical moral heuristic. It is metaphorical in the sense that the psychological dynamic it represents is *not* a literal contract, and a moral heuristic in the sense that it organizes the considerations involved in moral judgment (Sinnott-Armstrong et al., 2010). Literalizing contractarian relationships as actual contracts introduces a high level of what MacNeil (1980) refers to as 'procedural regularity', which involves externalizing the principals that govern proper behavior. Procedural regularity often becomes desirable, MacNeil (1980) argues, "when good faith and trust decline below certain levels" (p. 68). Between actors, in this case teachers and students, a high level of procedural regularity also sends the message that good faith and trust *are* low, and may encourage individuals to plan and act as if that were true.

> "The rise of procedural regularity respecting student-university relations is often hailed as a great step toward equality and justice. It can just as well be viewed as the result of a vast decline in trust and perceptions about an absence of good faith. So viewed it is a huge step backward." (MacNeil, 1980, p. 68)

While the urge to create literal contracts between students and teachers in order to clarify their relationships is understandable, and has been condoned in print (Deech, 2009;

Faucher & Caves, 2009; Frost, Hamlin & Barczyk, 2007; Gaffney-Rhys & Jones, 2010), using literal contracts to manage moral judgment is *non-sequitur*. As illustrated in the epigraph to this chapter, morality does not exist except through personal judgment and free will (Blasi, 1980). Literalizing psychological teaching-learning contracts would effectively bureaucratize the substance of what students and teachers should share of their own accord, in the natural course of their relationships. It would 'gut', in other words, the very substance of what gives teacher-learner relationships the potential to be uplifting, inspiring, and moral.

9.7 Limitations

Foremost among the limitations faced by the present study were relatively small sample size, data collection at only two time points and exclusively within the context of Science class, scalar non-invariance, sundry measurement dysfunction, and the use of self-reported data. While the sample was diverse, both geographically and culturally, its size restricted the range of modeling procedures that could be undertaken, as well as the statistical power associated with the more complex, multivariate models examined. The relationship of sample size to model complexity was expressed throughout the study as the $N{:}q$ ratio, where N is sample size and q is the number of free model parameters. Low $N{:}q$ ratios were more of an issue with respect to gender-specific models, which involved much smaller samples than co-ed models. Estimation of multivariate models with small samples at Time 2 and for the longitudinal analysis was accomplished by representing the variance of individual factors with weighted composite scores. This improved the $N{:}q$ ratios of multivariate models, but never above $N{:}q = 4.3$, implying persistently low statistical power. Analyses with low statistical power tend to under-represent effect size magnitudes and significance levels. A number of nearly-significant beta paths would likely have achieved significance with higher statistical power, such as that from *Teacher quality* to *Peer norms* in the Time 2 female sample model where the $N{:}q$ ratio was 3 ($\beta = -.270$, $p = .063$), and that from *Teacher quality* to

Justifiability of cheating in the Time 2 co-ed model where the *N:q* ratio was 1.5 (-.229, p = .077). The suppression and mediation effects noted in the male sample model at Time 2 may also have been caused by parameter estimates that were near zero due to low statistical power in what was the smallest sample used in the present study (N = 115). An *N:q* ratio of 10 - 20, as could be achieved with a sample of 2000 – 4000 for a cross-sectional test of Model 4 (q = 199), is recommended to optimize statistical power (Kline, 2011).

A further sample-related limitation of the study was the inability to purge school-level variance from the overall dataset using multilevel analysis using the COMPLEX command in M*plus*, due to having too few schools. Multilevel modeling in M*plus* requires that a minimum of twenty clusters be specified (L. K. Muthén personal communication, 15 November 2012), whereas the Main Study sample was drawn from only eleven schools. The COMPLEX command purges variance due to clustering, as by school, in the computation of standard errors and chi-squared statistics for a model (Muthén & Muthén, 1998-2010). Purging school-level variance would have improved the independence of observations, an important assumption of structural equation modeling (Kline, 2011).

The present study was also limited by the collection of data at only two time points and exclusively within the context of Science class. While collection from successive grade-levels allowed a comparison of changes within-groups and between-groups over two years, a two-wave design is inherently restricted to modeling only two types of within-groups change, namely increase or decrease. A multiwave design involving three or more instances of data collection would have enabled growth curve modeling for the various factors under investigation. Growth curves modeling accommodates more complex patterns of change over time than cross-sectional or two-wave designs (Bollen & Curran, 2006; McArdle & Nesselroade, 2003). The focus within Science class contexts was also limiting in that the relationships observed between variables in the present study could be different in other class contexts. Cheating has been found to be more common, for instance, in majors such as business at the university level (McCabe, 1997), and in 'hard' subjects such as Science at the secondary school level (Miller et al., 2007; Murdock et al., 2001; Schab, 1991). It is not clear, therefore, to what extent this work might generalize to 'soft' subjects such as English and History.

Scalar non-invariance was another limitation in the present study. A small but significant amount of scalar non-invariance was observed between gender groups at Times 1 and 2, as well as between the Time 1 and 2 response sets used for the longitudinal model. Scalar non-invariance, which involves group differences in the *y*-intercepts of the regression equations of observed variables, did not impede the use of composite variables to represent factor variances in multivariate models. It did, however, introduce ambiguity to the interpretation of differences observed between groups. Differential item functioning (DIF) analysis revealed a dispersion of low-level mean differences by gender in items that may, in sum, have conferred different operational definitions to several measures. The male conception of good teaching, for instance, might place greater emphasis on the timeliness and quality of feedback, whereas females might place greater emphasis on whether a teacher seems to care about what students have to say.

Psychometric dysfunction in several measures may have limited the generalizability of results of both the female sample model and co-ed model at Time 2. Congeneric models for *Performance goal structure* (Midgley et al., 2000) and *Surface learning strategies* (Simon et al., 2004)demonstrated significant measurement weakness in the female sample at Time 2 that seemed to carry across to the Time 2 co-ed analysis. Weakness demonstrated by these measures in Time 2 cross-sectional analyses cast doubt on the validity of their effects in the co-ed and female sample models. It is worth noting that these two measures did, however, perform acceptably well in longitudinal analyses.

While pains were taken to eliminate or modify weak measures in a pilot study conducted in the early stages of the overall program of this research, two *post-hoc* modifications were made at Time 2 that could not be cross-validated on separate data sets. Firstly, the congeneric model for *Surface learning strategies* was modified by freeing a covariance parameter between two of its indicator items in order to address what appeared to be a method effect. Secondly, a beta path was freed between *Honesty-trustworthiness self-concept* and *Teacher quality*, in order to address substantial model misfit to the female sample at Time 2. These two changes, which resulted in Model 4, cast some degree of doubt on the validity of their effects in the model.

Additionally, three constructs were dropped at Time One of the Main Study due to psychometric dysfunction. These included *Deep learning strategies* (Anderman et al., 1998), *Experience of classroom rules* (Gregory et al., 2010), and *Appropriate workload* (Wilson et al., 1997). The poor performance of these and other factors, noted above, might be related to the very high level of social-linguistic diversity in the current sample, and, in the case of *Appropriate workload*, also because it was originally designed for research at the university level.

The loss of *Appropriate workload* further limited the representation of moral obligation the in the PTLC model to the single measure *Justifiability of cheating*. The study is therefore limited in what it says about moral obligation, broadly, which would also likely encompass judgments of whether one should follow other types of rules and social expectations, such as working diligently and respecting others.

Limitations must also be addressed in relation to causal language. Causal language was used in discussions of the present research within the context of a body of prior experimental research that directly informed the PTLC model. No direct observations of causal processes could be made in the present study due to its non-experimental, passive-observational design. An individual respondent's belief while taking the questionnaire that cheating was justifiable in Science class, for instance, cannot be inferred, solely on the basis of this research, to have caused his or her cheating behavior in that context during the preceding year.

Additionally, not all hypotheses in the PTLC model were directly supported by prior experimental evidence. While all hypotheses were predicated upon strong causal theories, no experimental support was cited for the hypothesized downstream effects of *Subject self-concept*, for instance, nor for any of the hypothesized effects exerted upon *Peer norms* or *Surface learning strategies*, whereas the hypothesized effects of *Justifiability of cheating* on *Self-reported cheating* had only indirect experimental support, principally from ultimatum game research (e.g. Boles et al., 2000; Gneezy, 2005; Haidt, 2001; Pillutla & Murnighan, 1996).

The correlational, self-report design of the present research additionally limited the study in several ways. Firstly, self-report measures do not represent features of objective reality, but only features of what respondents perceive and remember at the time of the questionnaire. Students in the present study were not, for instance, able to report on what their peers actually thought about cheating when they responded to items of the *Peer norms* measure, but only on what they *believed* their peers thought. Nor would it be warranted to judge, solely based on perceived teacher quality data in the present study, whether, or in what way, the teachers in question should improve. The self-report design thus confined consideration of class context to what students remembered and perceived, as highlighted in the discussion of managing perceptions (section 9.5.2).

Similarly, the use of a self-report measure of cheating behavior introduced an unknown amount of variance from intra-psychic factors such as the accuracy of respondents' memories and their willingness to report cheating honestly. Self-reported data is prone to a variety of group- and individual-level response biases, or tendencies to answer questionnaire items in certain ways despite what is factually accurate (Austin, Deary, Gibson, et al., 1998). A noted type of response bias to which *Self-reported cheating* may be especially prone is 'socially desirable responding', or the preference for giving answers that are more socially acceptable over those that are true (Harding et al., 2007; Miller et al., 2008; Paulhus, 1991; Walker, 2010). The potential for intra-psychic variance in the measurement of *Self-reported cheating* limits the interpretability of PTLC model results as it may have both biased how much cheating actually reported, and inflated correlations with other intra-psychic variables, such as *Justifiability of cheating*. This limitation is notably important in interpreting the consistently high statistical associations between *Justifiability of cheating* and *Self-reported cheating*, which were used to represent moral judgment and cheating behavior, respectively. While such biases have been found in prior studies of academic cheating to be small (Harding et al., 2007; Mayhew et al., 2009), a high degree of shared intra-psychic variance would misrepresent the relationship between moral judgment and actual behavior, and could be seen to argue for modeling the two variables as a single higher-order factor.

9.8 Future research

The present work indicates that understanding why secondary school students cheat is aided by allowing for the role of non-rational processes in moral judgment. Such processes have been observed in numerous experimental studies, of which only a portion is reviewed above. Future research on cheating should benefit from continuing to push past the limits of rational-cognition, to the fuller range of mechanisms incorporated by a dual-process view of moral psychology. Qualitative research that seeks to discover why adolescents cheat would, in particular, benefit from designing questioning formats to avoid, inasmuch as possible, the potential for moral dumbfounding.

Qualitative research would be useful, for instance, on the distinction between cheating as a conventional violation *vs.* cheating as a moral violation. Such research might avoid moral dumbfounding by focusing firstly on whether students' judgments of the seriousness or triviality of particular types of cheating (e.g. test cheating *vs.* plagiarism *vs.* homework copying) fluctuate under various hypothetical conditions. The better we understand what conditions shape students' moral judgments, the better we should be able to craft counter-narratives and perception management strategies to manage those perceptions. Extending such research to exploring *why* students arrive at those judgments should anticipate difficulties associated with the adolescent ability to cognize and articulate such information, i.e. moral dumbfounding.

The social contract perspective on academic cheating featured in this study also needs to be tested experimentally. While social contract hypotheses have strong experimental support in the field of evolutionary psychology, none of these experiments involved academic cheating. Experimental evidence of how breaking social contracts affects academic cheating would cross-validate PTLC model results, and might provide valuable insight into, conversely, how social contracts can be managed and preserved.

The PTLC model entails a four-category framework that can be populated in future research with variables not investigated here, such as personality constructs, additional moral obligation constructs, other integrity-related behaviors, and other types of learner perception. Several recent publications suggest that student perceptions of socio-cultural factors, such as the level of national corruption, might impart significant variance to *Peer norms* and *Justifiability of cheating* (Crittenden et al., 2009a; Magnus et al., 2008; Teixeira et al., 2010). This research has not yet been carried out at the secondary level, and could speak to how cheating in secondary schools, and efforts to stop it, shape, and are shaped by, broader socio-cultural processes. Another category of variable to consider for inclusion in future PTLC models is that of interaction terms, such as between *Usefulness of curriculum* and *Teacher quality* (Murdock et al., 2004).

Recent research related to expectancy-value theory (Eccles & Wigfield, 1994) suggests, for instance, that task-value may be an important moderator of the effects of students' self-perceived ability (measured herein as *Subject self-concept*) and cheating, as a function of fear-of-failure (Lee et al., 2014). Including an 'efficacyXtask-value' interaction variable as a predictor of cheating behavior could augment the PTLC model by explaining variance in cheating that is likely to bear minimal relation to moral judgment. It may well be that variables related to anxiety and fear-of-failure would substantially increase the amount of variance in cheating already explained by the PTLC model. Models developed within the theory of planned behavior (TPB) have also explained as much as 71% of the variance in the intention to cheat (Mayhew et al., 2009). This compares favorably to the amount of variance explained in the justifiability of cheating in the present study (see Table 9.1), but, by focusing on intentions, it remains bound by the limits of the rational-cognitive paradigm. An integration of these models, in which the justifiability of cheating replaces the intention to cheat, could substantially increase variance explained in justifiability of cheating, which this research suggests should be a focal point of efforts to improve student ethics.

A particularly valuable extension of the present work would involve growth-curve modeling with a multiwave design (Bollen & Curran, 2006; McArdle & Nesselroade, 2003). The tendency for self-reported cheating behavior to increase over the Grade Eight to Grade Nine transition, and yet decrease over the Grade Nine to Grade Ten transition, has important practical implications for schools in which Grade Eight and Grade Nine are segregated into the Middle and High School divisions (see also Anderman & Midgley, 2004). A consistent spike in cheating behavior following the Grade Eight to Grade Nine transition would suggest that schools' efforts to improve student integrity should be brought to bear especially at the beginning of Grade Nine. Collecting data in such schools at the end of Grade Eight, and then at both the middle and end of Grade Nine, as well as, ideally, one or two points during Grade Ten, would help confirm this pattern.

Future research should also focus on gender differences among secondary school students with respect to the structural elements of the PTLC model and, more generally, with respect to non-rational processes underlying moral judgment. Invariance testing and gender-specific model results in the present research suggest that males and females differ in how they arrive at judgments of whether to engage in disintegrity. The effect of *Teacher quality* on *Justifiability of cheating* was, for instance, fully mediated among males by *Peer norms*, whereas, among females, it was direct. The effects of self-concept variables and *Performance goal structure* on downstream constructs also appeared to be more prominent among males. Future research that can utilize larger gender-specific samples would be in a position to cross-validate such differences with higher statistical power.

Integrating non-contractarian perspectives. The PTLC model could also be integrated into broader models that explain more variance in cheating behavior by including variable relations that are non-contractarian in nature. An under-investigated source of variance that is of potential interest to future research on academic integrity, in general, is socio-historical identity. As pointed out in Section 2.2, for instance, cheating behavior in Russia has been associated with a wide-spread tendency to 'hate' students who inform on others, which Magnus et al. (2002) identify as a lingering effect of anti-government sentiment during the Soviet era. Additionally, surface approaches to learning have been found to be more likely in cultures where education is viewed as a revered form of economic and social mobility (e.g. Phan, 2009b). Studies incorporating data from Transparency International's *Corruption Perceptions Index* (Crittenden et al., 2009a; Magnus et al., 2002) have found significant correlations, moreover, between academic cheating and perceived national corruption that may well reflect specific socio-historical factors. Future research in this area might use multilevel modeling to analyze data collected from members of multiple well-defined cultural groups. Significant findings in this area could help elucidate national- and cultural-level differences in the psychology of disintegrity behavior, and may have political implications that further elevate the relevance of moral and educational psychology (Anderman, 2011).

9.9 Chapter summary

The results of demographic, measurement, and structural analyses were discussed in this chapter in relation to (1) the hypotheses tested in the PTLC model, (2) the literatures and sub-disciplines within educational and moral psychology to which this work pertains, and (3) implications for educational practice. The evolution of the PTLC model was traced from Model 1, which included seventeen first-order factors, to Model 4, which included ten first-order factors, one second-order factor, and two covariates (gender and grade-level). The two outcome variables retained in Model 4, *Self-reported cheating* and *Surface learning strategies*, were shown, additionally, to be advisedly grouped together as 'disintegrity'.

Prominent in all of the models discussed was strong support for the hypothesis that the justifiability of cheating mediates the effects of students' perceptions of context on their disintegrity behaviors. This hypothesis was supported, in fact, to the near-complete exclusion of all hypothesized direct effects of context on cheating and surface learning strategies. The fact that the justifiability of cheating behavior fluctuated inversely with evaluations of class quality is consistent with the metaphor of a contract between teachers and learners, by which the felt obligations of students may be diminished or annulled to the extent that they believe a teacher has failed at his or her obligations.

It was argued that the correspondence of higher cheating behavior to higher justifiability of cheating across all PTLC models indicated that students often did not view cheating as a moral violation, but as a conventional one. Inasmuch as an act of cheating is viewed as a conventional violation, it should not conflict with an individual's abstract moral beliefs. Belief-behavior incongruities and neutralization techniques appear, therefore, to apply to a narrower scope of cheating behavior than recent literature indicates.

The large amount of variance explained in self-reported cheating behavior by whether cheating was judged to be justifiable suggested, moreover, that the measure *Justifiability of cheating* was capturing variance from both cognitive and non-cognitive processes. This is consistent with the dual-process paradigm of moral psychology, wherein non-cognitive processes are held to be managed, restrained, and articulated by cognitive ones. By asking for students' fully-formed, or 'terminal', judgments of whether cheating was justifiable, without asking them to explain *why* they hold such judgments, or whether they *intend* to cheat, the measure *Justifiability of cheating* avoided moral dumbfounding, or the difficulty that individuals often face in cognizing judgments that are at least partly of non-cognitive origin.

Implications of the present research were discussed next, highlighting that student-participants in the present study demonstrated perceptual tendencies, instead of behavioral tendencies, related to cheating. Cheating behavior at Time 2 was not rooted in cheating behavior at Time 1, but in perceptions and judgments that carried across time. It was argued, in view of these findings, that educators should complement honor codes and punitive consequences with efforts to manage student perceptions of teacher quality, learning goals, and moral self-concept, so as to mitigate the tendencies, observed in the longitudinal model, to perceive academic contexts in ways that negatively affect felt moral obligation. Together, these recommendations are consistent with what has been referred to as 'design-it-out' and 'constructive alignment', which are both design-based methods for structurally eliminating opportunities to engage in disintegrity behaviors, and that additionally counterbalance the message, associated with performance goals, that normative performance is more important than learning.

References

Aiken L., & West, S. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage Publishing Inc.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes 50*, 179-211. doi: 10.1016/0749-5978(91)90020-T

Ajzen, I. (2002). *Attitudes, personality, and behavior.* Buckingham, England: Open University Press.

Ajzen, I., & Sexton, J. (1999). Depth of processing, belief congruence, and attitude-behavior correspondence. In S. Chaiken & Y. Trope (Eds.), *Dual-processing theories in social psychology* (pp. 117-138). New York: The Guilford Press.

Alkharusi, H., Aldhafri, S., Alnabhani, H., & Alkalbani, M. (2013). The impact of students' perceptions of assessment tasks on self-efficacy and perception of task value: A path analysis. *Social Behavior and Personality*, *10*, 1681-1692. doi: 10.2224/sbp.2013.41.10.1681

Alschuler, A.S., & Blimling, G.S. (1995). Curbing epidemic cheating through systemic change. *College Teaching, 43*, 123-125. doi: 10.1080/87567555.1995.9925531

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84, 261-276. doi: 10.1037/0022-0663.84.3.261

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology, 80*, 260-267. doi: 10.1037/0022-0663.80.3.260

Anderman, E.M., Cupp, P.K., Lane, D. (2010). Impulsivity and academic cheating. *Journal of Experimental Education, 78,* 135-150. doi: 10.1080/00220970903224636

Anderman, E.M., Griesinger, T., & Westerfield, G. (1998). Motivation and cheating during early adolescence. *Journal of Educational Psychology, 90*, 84-93. doi: 10.1037/0022-0663.90.1.84

Anderman, E.M., & Midgley C. (1997). Changes in achievement goal orientations, perceived academic competence, and grades across the transition to middle school levels. *Contemporary Educational Psychology, 22*, 269-298. doi: 10.1006/ceps.1996.0926

Anderman, E.M., & Midgley C. (2004). Changes in self-reported academic cheating across the transition from middle school to high school. *Contemporary Educational Psychology, 29*, 499-517. doi: 10.1016/j.cedpsych.2004.02.002

Anderman, E.M. (2011). Educational psychology in the Twenty-First Century: Challenges for our community. *Educational Psychologist, 46*, 185-196. doi: 10.1080/00461520.2011.587724

Anderman, L.H., Freeman, T.M., & Mueller, C.E. (2007). The "social" side of social context: Interpersonal and affiliative dimensions of students' experiences and academic dishonesty. In E.M. Anderman & T.B. Murdock (Eds.), *Psychology of academic cheating* (pp. 203-228). Burlington, MA: Elsevier.

Anderson, J.C., & Gerbing, D.W. (1988). Structural equation modeling in practice: A review and recommended two step approach. *Psychological Bulletin, 103*, 411-423. doi: 10.1037/0033-2909.103.3.411

Antion, D.L., & Michael, W.B. (1983). Short-term predictive validity of demographic affective, personal, and cognitive variables in relation to two criterion measures of cheating behaviors. *Educational and Psychological Measurement, 43*, 467-482. doi: 10.1177/00131644830430021

Ariely, D. (2012). *The (honest) truth about dishonesty*. London: HarperCollins*Publishers*.

Arnold, S.D. (2012). Assessing student learning online: Overcoming reliability issues. *IADIS International Conference on Cognition and Exploratory Learning in Digital Age*, 189-196. Retrieved from Education Resource Information Center.

Aronson, E. (1968). Dissonance theory: Progress and problems. In R.P. Abelson, E. Aronson, WJ. McGuire, T.M. Newcomb, M.J. Rosenberg & P.H. Tannenbaum (Eds.). *Theories of cognitive consistency: A sourcebook* (pp. 5-27). Chicago: Rand McNally.

Aronson, E., & Mettee, D.R. (1968). Dishonest behavior as a function of differential levels of induced self-esteem. *Journal of Personality and Social Psychology, 9*, 121-127. doi: 10.1037/h0025853

Arvidson, C.J. (2004). *The anatomy of academic dishonesty: Cognitive development, self-concept, neutralization techniques, and attitudes toward cheating* (Doctoral dissertation, University of North Texas.) Retrieved from http://digital.library.unt.edu/ark:/67531/metadc4641/m2/1/high_res_d/dissertation.pdf

Asch, S.E. (1956). Studies of independence and conformity. I. A minority of one against a unanimous majority. *Psychological Monographs*, 70, 1–70. doi: 10.1037/h0093718

Ashworth, P., Bannister, P., & Thorne, P. (1997). Guilty in whose eyes? University students' perceptions of cheating and plagiarism in academic work and assessment. *Studies in Higher Education, 22*, 187-203. doi: 10.1080/03075079712331381034

Asparouhov, T., & Muthén, B.O. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 16*. doi: 10.1080/10705510903008204.

Austin, E.J., Deary, I.J., Gibson, G.J., McGregor, M.J., & Dent, J.B. (1998). Individual response spread in self-report scales: Personality correlations and consequences. *Personality and Individual differences, 24*, 421-438. doi: 10.1016/S0191-8869(97)00175-X

Bacha, N.N., Bahous, R., & Nabhani, M. (2012). High schoolers' views on academic integrity. *Research Papers in Education, 27*, 365-381. doi: 10.1080/02671522.2010.550010

Baird, J.S. (1980). Current trends in college cheating. *Psychology in the Schools, 17*, 515-522. doi: 10.1002/1520-6807(198010)17:4<515::AID-PITS2310170417>3.0.CO;2-3

Bamberger, Y.M. (2014). Encouraging girls into science and technology with feminine role model: Does this work? *Journal of Science and Technology, 23*, 549-561. doi: 10.1007/s10956-014-9487-7

Bandalos, D.L., & Finney, S.J. (2010). Factor analysis: Exploratory and confirmatory. In G.R. Hancock and R.O. Mueller (Eds.), *The reviewers guide to quantitative methods in the social sciences* (pp. 93-114). New York: Routledge.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215. doi: 10.1016/0146-6402(78)90002-4

Bandura, A. (1986). *Social foundations of thought and action: A social-cognitive view*. Englewood Cliffs, NJ: Prentice-Hall

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review, 3*, 193-209. doi: 10.1207/s15327957pspr0303_3

Barnes, E. (1904). Student honor: A study in cheating. *International Journal of Ethics, 14*, 481-488. doi: 10.1086/206091

Barnes, C.M., Schaubroeck, J., Huth, M., & Ghumman, S. (2011). Lack of sleep and unethical conduct. *Organizational Behavior and Human Decision Processes, 115*, 169-180. doi: 10.1016/j.obhdp.2011.01.009

Barnhardt, B., & Ginns, P. (2014). An alienation-based framework for student experience in higher education: New interpretations of past observations in student learning theory. *Higher Education*. doi: 10.1007/s10734-014-9744-y

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personal and Individual Differences, 42*, 815 – 824. doi: 10.1016/j.paid.2006.09.018

Batane, T. (2010). Turning to Turnitin to fight plagiarism among university students. *Journal of Educational Technology & Society*, *13*(2), 1-12.

Baumrind, D. (1987). A developmental perspective on adolescent risk taking in contemporary America. *Adolescent Social Behavior and Health, 37*, 93-125. doi: 10.1002/cd.23219873706

Beasley, E.M. (2014). Students reported for cheating explain what they think would have stopped them. *Ethics & Behavior, 24*, 229-252. doi: 10.1080/10508422.2013.845533

Bebeau, M.J, & Thoma, S.J. (1999). "Intermediate" conceptions and the connection to moral education. *Educational Psychology Review, 11*, 343-360. doi: 10.1023/A:1022057316180

Beck, L., & Ajzen, I. (1991). Predicting dishonest actions using the theory of planned behavior. *Journal of Research in Personality, 25*, 285-301. doi: 10.1016/0092-6566(91)90021-H

Becker, G.S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy, 76*, 169-217. doi: 10.1086/259394

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246. doi: 10.1037/0033-2909.107.2.238

Bentler, P.M., & Chou, C. (1987). Practical issues in structural equation modeling. *Sociological Methods and Research, 16*, 78-117.

Bergman, R. (2002). Why be moral? A conceptual model from developmental psychology. *Human Development*, *45*, 104-124. doi: 10.1159/000048157

Berliner, D.C. (2002). Educational research: The hardest science of all. *Educational Researcher, 31*, 18-20. doi: 10.3102/0013189X031008018

Bernardi, R.A., Baca, A.V., Landers, K.S., & Witek, M.B. (2008). Methods of cheating and deterrents to classroom cheating: An international study. *Ethics & Behavior, 18*, 373-391. doi: 10.1080/10508420701713030

Biggs, J.B. (1987). *Student approaches to learning and studying*. Hawthorn: Australian Council for Educational Research.

Biggs, J.B. (1991). Student learning in the context of school. In J. B. Biggs (ed.), *Teaching for Learning* (pp. 7-29). Hawthorn: Australian Council for Educational Research, Ltd.

Biggs, J.B. (1993). From theory to practice: A cognitive systems approach. *Higher Education Research and Development, 12*, 73-86. doi: 10.1080/0729436930120107

Biggs, J.B. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*, 1-18. doi: 10.1007/BF00138871

Biggs, J.B. (1999). What the student does: Teaching for enhanced learning. *Higher Education Research & Development, 18*, 57-75. doi: 10.1080/0729436990180105

Biggs, J.B., Kember, D., & Leung, D.Y.P. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, *71*, 133-149. doi: 10.1007/BF03173004

Biggs, J.B., & Tang, C. (2011). *Teaching for quality learning at university* (4th ed.). Maidenhead: Open University Press.

Bing, M.N., Davison, H.K., Vitell, S.J., Ammeter A.P., Garner B.L., & Novicevic, M.M. (2012). An experimental investigation of an interactive model of academic cheating among business school students. *Academy of Management Learning & Education, 11*, 28-48. doi: 10.5465/amle.2010.0057

Bisping, T.O., Patron, H., & Roskelley, K. (2008). Modeling academic dishonesty: The role of student perceptions and misconduct type. *The Journal of Economic Education, 39*, 4-21. doi: 10.3200/JECE.39.1.4-21

Bjorklund, F., Haidt, J., & Murphy, S. (2000). Moral dumbfounding: When intuition finds no reason. *Lund Psychological Reports,* 1(2), 1-23.

Black, P.J., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7-74. doi: 10.1080/0969595980050102

Black, P.J., & Wiliam, D. (2006a). Assessment for Learning in the Classroom, in J. Gardner (Ed.), *Assessment and Learning* (pp. 9-25), London: Sage Publications, Ltd.

Black, P.J., & Wiliam, D. (2006b). Developing a Theory of Formative Assessment, in J. Gardner (Ed.), *Assessment and Learning* (pp. 81-100). London: Sage Publications, Ltd.

Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin, 88*, 1-45. doi: 10.1037/0033-2909.88.1.1

Blasi, A. (1983). Moral cognition and moral action: A theoretical perspective. *Developmental Review, 3*, 178-210. doi: 10.1016/0273-2297(83)90029-1

Boles, T.L., Croson, R.T.A., & Murnighan, J.K. (2000). Deception and retribution in repeated ultimatum bargaining. *Organizational Behavior and Human Decision Processes, 83*, 235-259. doi: 10.1006/obhd.2000.2908

Bollen, K.A. (1989). Sample size and Bentler and Bonett's nonnormed index. *Psychometrika*, *51*, 375-377. doi: 10.1007/BF02294061

Bollen, K.A., & Curran, P.J. (2006). *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.

Bong, M. (2008). Effects of parent-child relationships and classroom goal structures on motivation, help-seeking avoidance, and cheating. The Journal of Experimental Education, 76, 191-217. doi: 10.3200/JEXE.76.2.191-217

Bong, M., Cho, C., Ahn, H.S., & Kim, H.J. (2012). Comparison of self-beliefs for predicting student motivation and achievement. *The Journal of Educational Research, 105*, 336-352. doi: 10.1080/00220671.2011.627401

Bong, M., & Clark, R.E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychology, 34*, 139-154. doi: 10.1207/s15326985ep3403_1

Bong, M., Skaalvik, E.M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*, 1-40. doi: 10.1023/A:1021302408382

Boomsma, A., & Herzog, W. (2013). R function swain: Correcting structural equation model t-statistics and indexes under small-sample and/or large-model conditions. http://www.ppsw.rug.nl/~boomsma/swain.pdf.

Bouville, M. (2007). *Cheating and neutralization*. Cambridge, UK: University of Cambridge. http://www.mathieu.bouville.name/education-ethics/Bouville-cheating-neutralization.pdf

Bowers, W.J. (1964). *Student dishonesty and its control in college*. New York: Bureau of Applied Research, Columbia University.

Bracey, G.W. (2005). Cheating update. *The Phi Delta Kappan, 86*, 637-638. doi: 10.1177/003172170508600815

Brandes, B. (1986). *Academic honesty: A special study of California students*. Sacramento, CA: California State Department of Education, Bureau of Publications.

Brent, E., & Atkisson, C. (2011). Accounting for cheating: An evolving theory and emergent themes. *Research in Higher Education, 52*, 640-658. doi: 10.1007/s11162-010-9212-1

Breslow, N.E., & Day, N.E. (1980). *Statistical methods in cancer research. Volume I – The Analysis of Case-Control Studies.* Lyon: International Agency for Research on Cancer (IARC Scientific Publications No. 32).

Briggs, K., Workman, J.P., & York, A.S. (2013). Collaborating to cheat: A game theoretic exploration of academic dishonesty in teams. *Academy of Management Learning & Education, 12*, 4-17. doi: 10.546S/Qmle.2011.0M0

Broeckelman-Post, M.A. (2008). Faculty and student classroom influences on academic dishonesty. *IEEE Transactions on Education*, *51*, 206-211. doi: 10.1109/TE.2007.910428

Brophy, J. (2005). Goal theorists should move on from performance goals. *Educational Psychologist, 40*, 167-176. doi: 10.1207/s15326985ep4003_3

Brosnan, S.F., & de Waal, F.B.M. (2003). Monkeys reject unequal pay. *Nature, 425*, 297-299. doi: 10.1038/nature01963

Brosnan, S.F., & de Waal, F.B.M. (2002). A proximate perspective on reciprocal altruism. *Human Nature, 13,* 129-152. doi: 10.1007/s12110-002-1017-2

Brousseau, G. 1984. The Crucial Role of the Didactical Contract in the Analysis and Construction of Situations in Teaching and Learning mathematics. In H. Steiner (Ed.), *Theory of Mathematics Education Paper 54* (pp. 110-119). Institut fur Didaktik der Mathematik der Universitat Bielefeld.

Brown, T.A. (2006). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.

Brown-Wright, L., Tyler, K.M., Stevens-Watkins, D., Thomas, D., Mulder, S., Hughes, T., Stevens-Morgan, R., Roan-Belle, C., Gadson, N., & Smith, L.T. (2012). Investigating the link between home-school dissonance and academic cheating among high school students. *Urban Education, 48*, 314-334. doi: 10.1177/0042085912469203

Browne, M.W., & Crudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Part, CA: Sage Publications.

Brüne, M., Juckel, G., & Enzi1, B. (2013). "An eye for an eye"? Neural correlates of retribution and forgiveness. *PLoS ONE, 8*, 1-11. doi: 10.1371/journal.pone.0073519

Bryan, C.J., Adams, G.S., & Monin, B. (2013). When cheating would make you a cheater: Implicating the self prevents unethical behavior. *Journal of Experimental Psychology, 142*, 1001-1005. doi: 10.1037/a0030655

Buckley, M.R., Novicevic, M.M., Halbesleben, J.R.B., & Harvey, M. (2004). Course management and students' expectations: theory-based considerations. *International Journal of Educational Management, 18*, 138-144. doi: 10.1108/09513540410522261

Burrus, R.T., McGoldrick, K., & Schuhmann, P.W. (2007). Self-reports of student cheating: Does a definition of cheating matter? *The Journal of Economic Education, 38*, 3-16. doi: 10.3200/JECE.38.1.3-17

Burton, R.V. (1963). Generality of honesty reconsidered. *Psychological Review, 70*, 481-499. doi: 10.1037/h0047594

Bushway, A., & Nash, W.R. (1977). School cheating behavior. *Review of Educational Research, 47(4)*, 623-632. doi: 10.3102/00346543047004623

Byrne, B. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York: Routledge.

Byrne, B.M., Shavelson, R.J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105,* 456–466. doi: 10.1037/0033-2909.105.3.456

Byrne, B.M., & van de Vijver, F.J.R. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing, 10*, 107-132. doi: 10.1080/15305051003637306

Calabrese, R.L., and Cochran, J.T. (1990). The relationship of alienation to cheating among a sample of American adolescents. *Journal of Research and Development in Education, 23(2)*, 65-72.

Callahan, D. (2004). *The cheating culture*. New York: Houghton Mifflin Harcourt.

Campbell, J.D. (1990). Self-esteem and clarity on the self-concept. *Journal of Personality and Social Psychology, 59*, 538-549. doi: 10.1037/0022-3514.59.3.538

Canning, R.R. (1956). Does an honor system reduce classroom cheating? An experimental answer. *The Journal of Experimental Education, 24*(4), 291-296.

Carrell, S.E., Malmstrom, F.V., & West, J.E. (2008). Peer effects in academic cheating. *Journal of Human Resources, 43*, 173-207. doi: 10.1353/jhr.2008.0013

Carroll, J., & Appleton, J. (2001). *Plagiarism: A good practice guide.* Oxford: Oxford Brookes University.

Carson, D.C. (2013). Perceptions of prosocial and delinquent peer behavior and the effect on delinquent attitudes: A longitudinal study. *Journal of Criminal Justice, 41*, 151-161. doi: 10.1016/j.jcrimjus.2013.01.005

Chaiken, S., & Trope, Y. (Eds.) (1999). *Dual-processing theories in social psychology*. New York: The Guilford Press.

Chao, K.-J., Hung, I.-C., & Chen, N.-S. (2012). On the design of online synchronous assessments in a synchronous cyber classroom. *Journal of Computer Assisted Learning, 28*, 379-395. doi: 10.1111/j.1365-2729.2011.00463.x

Chen, F. Curran, P.J., Bollen, K.A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research, 36*, 462-494. doi: 10.1177/0049124108314720

Cheung, D. (2000). Evidence of a single second-order factor in student ratings of teaching effectiveness. *Structural Equation Modeling, 7*, 442-460, doi: 10.1207/S15328007SEM0703_5

Cheung, D., & Ng, P.H. (2000). Science teachers' beliefs about curriculum design. *Research in Science Education, 30*, 357-375. doi: 10.1007/BF02461556

Cheung, G.W., & Rensvold R.B. (2002). Evaluation of goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 233-255. doi: 10.1207/S15328007SEM0902_5

Chou, C.P., Bentler, P.M., & Satorra, A. (1991). Scaled test statistics and robust standard errors for non-normal covariance structure analysis: A Monte Carlo study. *British Journal of Mathematical and Statistical Psychology*, 44, 347-357. doi: 10.1111/j.2044-8317.1991.tb00966.x

Cizek, G.J. (2003). *Cheating on tests: How to do it, detect it, and prevent it*. Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1970). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment* (Vol. 1). Cambridge, UK: Cambridge University Press.

Colby, A., Kohlberg, L., Gibbs, J., Lieberman, M., Fischer, K., & Saltzstein, H.D. (1983). A longitudinal study of moral judgment. *Monographs of the Society for Research in Child Development, 48*, 1-124. doi: 10.2307/1165935

Colnerud, G., & Rosander, M. (2009). Academic dishonesty, ethical norms and learning. *Assessment & Evaluation in Higher Education, 34* 505-517. doi: 10.1080/02602930802155263

Compton, J., & Pfau, M. (2008). Inoculating against pro-plagiarism justifications: Rational and affective strategies. *Journal of Applied Communication Research, 36*, 98-119. doi: 10.1080/00909880701799329

Conger, A.J. (1974). A revised definition for suppressor variables: A guide to their identification and interpretation. *Educational and Psychological Measurement 34*, 35-46. doi: 10.1177/001316447403400105

Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Boston: Houghton Mifflin.

Coote, L. (2011). *Fundamentals of structural equation modelling*. Lecture notes for ACSPRI 2011 Winter Program. Australian National University, Canberra.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31*, 187-276. doi: 10.1016/0010-0277(89)90023-1

Cosmides, L., & Tooby, J. (2013). Evolutionary psychology: New perspectives on cognition and motivation. *The Annual Review of Psychology, 64*, 201-229. doi: 10.1146/annurev.psych.121208.131628

Costello, A.B., & Osborne, J. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation, 10(7)*, 1-9.

Covey, M.K., Saladin, S., & Killen, P.J. (1989). Self-monitoring, surveillance, and incentive effects on cheating. *The Journal of Social Psychology, 129*, 673-679. doi: 10.1080/00224545.1989.9713784

Crittenden, V.L., Hanna, R.C., & Peterson, R.A. (2009a). Business students' attitudes toward unethical behavior: A multi-country comparison. *Marketing Letters*, *20*, 1-14. doi: 10.1007/s11002-008-9051-4

Crittenden, V.L., Hanna, R.C., & Peterson, R.A. (2009b). The cheating culture: A global societal phenomenon. *Business Horizons*, 52, 337-346. doi:10.1016/j.bushor.2009.02.004

Cronbach L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika 16*, 297–334. doi: 10.1007/bf02310555

Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*, 438-481. doi: 10.2307/1170281

Crown, D.F., & Spiller, M.S. (1998). Learning from the literature on collegiate cheating: A review of empirical research. *Journal of Business Ethics, 17*, 683-700. doi: 10.1023/A:1017903001888

Curran, P. J., West, S. G., & Finch, J. F. (1997). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*, 16-29. doi: 10.1037/1082-989X.1.1.16

Curry, L. (1984). Student commitment and school organization in relation to on-task behavior and achievement. *Contemporary Educational Psychology, 9*, 171-184. doi: 10.1016/0361-476X(84)90018-3

Cushman, F., Young, L., & Greene, J.D. (2010). Multi-system moral psychology. In J.M. Doris (Ed.), *The moral psychology handbook* (pp. 47-68). Oxford: Oxford University Press.

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. London: Vintage Books.

David, P. (1973). Correlates of cheating behavior in a ghetto elementary school. *Graduate Research in Urban Education & Related Disciplines, 7(1)*, 35-63..

Davis, J.A. (1985). The logic of causal order, Quantitative Applications in the Social Sciences Series, (No. 55). Sage Publications, Richmond, Virginia.

Davis, S.F., Grover, C.A., Becker, A.H., & McGregor, L.N. (1992). Academic dishonesty: Prevalence, determinants, techniques, and punishments. *Teaching of Psychology, 19*, 16-20. doi: 10.1207/s15328023top1901_3

Davis, S.F., & Ludvigson, H.W. (1995). Additional data on academic dishonesty and a proposal for remediation. *Teaching of Psychology, 22*, 119-121. doi: 10.1207/s15328023top2202_6

Davy, J.A., Kincaid, J.F., Smith, K.J., & Trawick, M.A. (2007). An examination of the role of attitudinal characteristics and motivation on the cheating behavior or business students. *Ethics & Behavior, 17*, 281-302. doi: 10.1080/10508420701519304

Day, N.E., Hudson, D., Dobles, P.R., & Warls, R. (2011). Student or situation? Personality and classroom context as predictors of attitudes about business school cheating. *Social Psychology of Education, 14*, 261-282. doi: 10.1007/s11218-010-9145-8

De Bruin, G.P., & Rudnick, H. (2007). Examining the cheats: The role of conscientiousness and excitement seeking in academic dishonesty. *South African Journal of Psychology, 37*, 153-164. doi: 10.1177/008124630703700111

Deci, E.L., & Ryan, R.M. (2011). Levels of analysis, regnant causes of behavior and well-being: The role of psychological needs. Psychological Inquiry, 22, 17-22. doi: 10.1080/1047840X.2011.545978

Deech, R. (2009). The student contract. *The Law Teacher, 43*, 3-13 doi: 10.1080/03069400802703102

DePalma, M.T., Madey, S.F., & Bornschein, S. (1995). Individual differences and cheating behavior: Guilt and cheating in competitive situations. *Personality and Individual Differences, 18*, 761-769. doi: 10.1016/0191-8869(94)00208-A

de Waal, F.B.M. (1991). The Chimpanzee's sense of social regularity and its relation to the human sense of justice. *American Behavioral Scientist 34*, 335-349. doi: 10.1177/0002764291034003005

de Waal, F.B.M (2014). Natural normativity: The 'is' and 'ought' of animal behavior. *Behaviour, 151*, 185-204. doi: 10.1163/1568539X-00003146

Desalegn, A.A., & Berhan, A. (2014). Cheating on examinations and its predictors among undergraduate students at Hawassa University College of Medicine and Health Science, Hawassa, Ethiopia. *BMC Medical Education, 14*, 89-100. doi: 10.1186/1472-6920-14-89

Desruisseaux, P. (1999, April 30). Cheating is reaching epidemic proportions worldwide, researchers say. *The Chronical of Higher Education*, *45*(34), A45. Retrieved from http://chronicle.com/article/Cheating-Is-Reaching-Epidemic/5854/

DeVries, D.L., & Ajzen, I. (1971). The relationship of attitudes and normative beliefs to cheating in college. *The Journal of Social Psychology, 83*, 199-207. doi: 10.1080/00224545.1971.9922463

Diekhoff, G.M., LaBeff, E.E., Clark, R.E., Williams, L.E., Francis, B., & Haines, V.J. (1996). College cheating: Ten years later. *Research in Higher Education, 37*, 487-502. doi: 10.1007/BF01730111

Ding, L., Velicer, W.F., & Harlow, L.L. (1995). Effects of estimation method, number of indicators per factor, and improper solutions on structural equation modeling fit indices. *Structural Equation Modeling: A Multidisciplinary Journal, 2*, 119-143. doi: 10.1080/10705519509540000

Diseth, Å. (2007). Students' evaluation of teaching, approaches to learning, and academic achievement. *Scandinavian Journal of Educational Research, 51*, 185−204. doi: 10.1080/00313830701191654

Diseth, Å., Palleson, S., Brunborg, G.S., & Larsen, S. (2010). Academic achievement among first semester undergraduate psychology students: the role of course experience, effort, motives and learning strategies. *Higher Education, 59*, 339-352.

Dix, E.L., Emery, L.F., & Le, B. (2014). Committed to the honor code: An investment model analysis of academic integrity. *Social Psychology of Education, 17*, 179-196. doi: 10.1007/s11218-013-9246-2

Doherty, C., & Shield, P. (2012). Teachers' work in curricular markets: Conditions of design and relations between the International Baccalaureate diploma and the local curriculum. *Curriculum Inquiry, 42*, 414-441. doi: 10.1111/j.1467-873X.2012.00596.x

Dorman, J.P., & Knightley, W. M. (2006). Development and validation of an instrument to assess secondary school students' perceptions of assessment tasks. *Education Studies, 32*, 47-58. doi: 10.1080/03055690500415951

Eisenberg, J. (2004). To cheat or not to cheat: Effects of moral perspective and situational variables on students' attitudes. *Journal of Moral Education, 33*, 163-178. doi: 10.1080/0305724042000215276

Elias, R.Z. (2009). The impact of anti-intellectualism attitudes and academic self-efficacy on business students' perceptions of cheating. *Journal of Business Ethics, 86*, 199-209. doi: 10.1007/s10551-008-9843-8

Elliot, A.J. (2005). A conceptual history of the achievement goal construct. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 52–72). New York, NY: Guilford Press.

Elliot, A.J., & McGregor, H.A. (2001). A 2 X 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*, 501-519. doi: 10.1037//0022-3514.80.3.501

Elliot, A.J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology, 100*, 613-628. doi: 10.1037/0022-0663.100.3.613

Elliot, A.J., Murayama, K., & Pekrun, R. (2011). A 3 x 2 achievement goal model. *Journal of Educational Psychology, 103*, 632-648. doi: 10.1037/a0023952

Emmer, E.T., & Stough, L.M. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational Psychologist, 36*, 103-112. doi: 10.1207/S15326985EP3602_5

Engler, J.N., & Landau, J.D. (2011). Source is important when developing a social norms campaign to combat academic dishonesty. *Teaching of Psychology, 38*, 46-48. doi: 10.1177/0098628310390848

Entwistle, N.J., & Ramsden, P. (1983). *Understanding student learning*. London: Croom Helm.

Erickson, G. and Erickson, L. (1984). Females and science achievement: evidence, explanations and implications. *Science Education, 68*, 63–89. doi: 10.1002/sce.3730680202

Evans, E.D., & Craig, D. (1990a). Adolescent cognitions for cheating as a function of grade level and achievement status. *Journal of Adolescent Research, 5*, 325-345. doi: 10.1177/074355489053005

Evans, E.D., & Craig, D. (1990b). Teacher and student perceptions of academic cheating in middle and senior high school. *The Journal of Educational Research, 84*(1), 44-52.

Evans, E.D., Craig, D., & Mietzel, G. (1993). Adolescents' cognitions and attributions for academic cheating: A cross-national study. *The Journal of Psychology 127*, 585-602. doi: 10.1080/00223980.1993.9914897

Evans, R. (2006). Evaluating an electronic plagiarism detection service: The importance of trust and the difficulty of proving students don't cheat. *Active Learning in Higher Learning, 7*, 87-99. doi: 10.1177/1469787406061150

Fang & Casadevall (2013 May/June). Why we cheat: Scientists are unraveling the causes of fraud and devising new strategies for rooting them out. *Scientific American Mind*, *24*(2) 31-37.

Farrell, A.D. (1994). Structural equation modeling with longitudinal data: Strategies for examining group differences and reciprocal relationships. *Journal of Consulting and Clinical Psychology, 62*, 477-487. doi: 10.1037//0022-006X.62.3.477

Faucher, D., & Caves, S. (2009). Academic dishonesty: Innovative cheating techniques and the detection and prevention of them. *Teaching and Learning in Nursing, 4*, 37-41. doi: 10.1016/j.teln.2008.09.003

Fehr, E., & Gächter, S. (1998). Reciprocity and economics: The economic implications of *Homo reciprocans*. *European Economic Review, 42*, 845-859. doi: 10.1016/S0014-2921(97)00131-1

Feinberg, J.M. (2009). Perception of cheaters: The role of past and present academic achievement. *Ethics & Behavior, 19*, 310-322. doi: 10.1080/10508420903035299

Feldman, S.E., & Feldman, M.T. (1967). Transition of sex differences in cheating. *Psychological Reports, 20*, 957-958. doi: 10.2466/pr0.1967.20.3.957

Fellonar, P., Román, S., & Cuestas, P.J. (2007). University students' academic performance: An integrative conceptual framework and empirical analysis. *British Journal of Educational Psychology, 77*, 873-891. doi: 10.1348/000709907X189118

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117-140. doi: 10.1177/001872675400700202

Feynman, R. (1974). Cargo Cult Science [Video file]. Retrieved from https://www.youtube.com/ watch?v=yvfAtIJbatg

Fiddick, L., Spampinato, M.V., & Grafmand, J. (2005). Social contracts and precautions activate different neurological systems: An fMRI investigation of deontic reasoning. *NeuroImage, 28*, 778-786. doi: 10.1016/j.neuroimage.2005.05.033

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage Publications.

Finn, K.V., & Frone, M.R. (2004). Academic performance and cheating: Moderating role of school identification and self-efficacy. *The Journal of Educational Research, 97*, 115-121. doi: 10.3200/JOER.97.3.115-121

Fletcher, J.F. (1966). *Situation ethics: The new morality*. Philadelphia: Westminster Press.

Franklyn-Stokes, A., & Newstead, S.E. (1995). Undergraduate cheating: Who does what and why? *Studies in Higher Education, 20*, 159-172. doi: 10.1080/03075079512331381673

Fransson, A. (1977). On qualitative differences in learning. IV – Effects of motivation and test anxiety on process and outcome. *British Journal of Educational Psychology, 47*, 244-257. doi: 10.1111/j.2044-8279.1977.tb02353.x

Frost, J., Hamlin, A., & Barczyk, C. (2007). A survey of professors' reactions to scenarios of academic dishonesty in American universities. *Journal of Business Inquiry, 7*, 11-18.

Gaffney-Rhys, R. and Jones, J. (2010). Issues surrounding the introduction of formal student contracts. *Assessment and Evaluation in Higher Education, 35*, 711-735. doi: 10.1080/ 02602930902977749

Galbraith, C.S., & Merrill, G.B. (2012). Academic and work-related burnout: A longitudinal study of working undergraduate university business students. *Journal of College Development, 53*, 452-463. doi: 10.1353/csd.2012.0044.

Gallant, T.B. (2008). *Academic integrity in the Twenty-First Century*. Hoboken, NJ: Wiley Periodicals.

Gallant, T.B., Anderson, M.G., & Killoran, C. (2013). Academic integrity in a mandatory physics lab: The influence of post-graduate aspirations and grade point averages. *Science and Engineering Ethics, 19*, 219-235. doi: 10.1007/s11948-011-9325-8

Galloway, M.K. (2012). Cheating in advantaged high schools: Prevalence, justifications, and possibilities for change. *Ethics & Behavior, 22*, 378-399. doi: 10.1080/10508422.2012.679143

Gannon-Leary, P., Trayhurn, D., & Home, M. (2009). Good images, effective messages? Working with students and educators on academic practice understanding. *Journal of Further and Higher Education, 33*, 435-448. doi: 10.1080/03098770903272511

Garavalia, L., Olson, E., Russell, E., & Christensen, L. (2007). How do students cheat? In E.M. Anderman & T.B. Murdock (Eds.), *Psychology of academic cheating* (pp. 33-55). Burlington, MA: Elsevier.

Genereux, R.L., & McLeod, B.A. (1995). Circumstances surrounding cheating: A questionnaire study of college students. *Research in Higher Education, 36*, 687-704. doi: 10.1007/BF02208251

Gijbels, D., Segers, M., & Struyf, E. (2008). Constructivist learning environments and the (im)possibility to change students'' perceptions of assessment demands and approaches to learning. *Instructional Science, 36*, 431-443. doi: 10.1007/s11251-008-9064-7

Ginns, P., Martin, A.J., & Papworth, B. (2014). Student learning theory goes (back) to (high) school. *Instructional Science, 42*, 485-504. doi: 10.1007/s11251-013-9291-4

Ginns, P., Prosser, M., and Barrie, S. (2007). Students' perceptions of teaching quality in higher education: the perspective of currently enrolled students. *Studies in Higher Education 32*, 603-615. doi: 10.1080/03075070701573773

Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and differentiation in unethical behavior: The effect of one bad apple in the barrel. *Psychological Science, 20*, 393-398. doi: 10.1111/j.1467-9280.2009.02306.x

Gino, F., & Pierce, L. (2009). The abundance effect: Unethical behavior in the presence of wealth. *Organizational Behavior and Human Decision Processes, 109*, 142-155. doi: 10.1016/j.obhdp.2009.03.003

Gino, F., Schweitzer, M.E., Mead, N., & Ariely, D. (2011). Unable to resist temptation: How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes, 115*, 191-203. doi: 10.1016/j.obhdp.2011.03.001

Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review, 95*, 384-394. doi: 10.1257/0002828053828662

Gray, B.T. (1997). *Higher-order factor analysis*. Paper presented at the Annual Meeting of the Southwest Educational Research Association, Austin, TX. Paper retrieved from http://files.eric.ed.gov.ezproxy1.library.usyd.edu.au/fulltext/ED407418.pdf

Grayson, D.A., Mackinnon, A., Jorn, A.F., Creasey, H., & Broe, G.A. (2000). Item bias in the Center in the Epidemiologic Studies Depression Scale: Effects of physical disorders and disability in an elderly community sample. *The Journal of Gerontology Series B: Psychological Sciences and Social Sciences, 55*, 273-282. doi: *10.1093/geronb/55.5.P273*

Greenberg, J. (1993). Stealing in the name of justice: Informational and interpersonal moderators of theft reactions to underpayment inequity. *Organizational Behavior and Human Decision Processes 54*, 81-103. doi: 10.1006/obhd.1993.1004

Greene, J.D., Morelli, S.A., Lowenberg, K., Nystrom, L.E., & Cohen, J.D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition, 107*, 1144-1154. doi:10.1016/j.cognition.2007.11.004

Greene, J.D., Nystrom, L.E., Engell, A.D., & Darley, J.M. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron, 44*, 389-400. doi: 10.1016/j.neuron.2004.09.027

Greene, J.D., & Paxton, J.M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences, 106*, 12506–12511. doi: 10.1073_pnas.0900152106

Greene, J.D., & Haidt, J. (2002). How (and where) does moral judgment work? *TRENDS in Cognitive Sciences, 6*, 517-523. doi: 10.1016/S1364-6613(02)02011-9

Gregory, A., Cornell, D., Fan, X., Sheras, P., Shih, T., & Huang, F. (2010). Authoritative school discipline: High school practices associated with lower bullying and victimization. *Journal of Educational Psychology, 2*, 483-496. doi: 10.1037/a0018562

Griggs, R.A., & Cox, J.R. (1983). The effects of problem content and negation on Wason's selection task. *Quarterly Journal of Experimental Psychology, 35A*, 533. doi: 10.1080/14640748308402486

Grotius, H. (1625/1962). *The law of war and peace: De jure belli ac pacis libri tres* (F.W. Kelsey, Trans.). Indianapolis: Bobbs-Merrill.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814-834. doi: 10.1037/0033-295X.108.4.814

Haidt, J. (2003). The emotional dog does learn new tricks: A reply to Pizarro and Bloom (2003). *Psychological Review, 110*, 197-198. doi: 10.1037/0033-295X.110.1.197

Haidt, J. (2007). The new synthesis in moral psychology. *Science, 316*, 998-1002. doi: 10.1126/science.1137651

Haines, V.J., Diekhoff, G.M., LaBeff, E.E., & Clark, R.E. (1986). College cheating: Immaturity, lack of commitment and the neutralizing attitude. *Research in Higher Education, 25*, 342-354. doi: 10.1007/BF00992130

Hall, N. (2007). Structural equations and causation. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 132*, 109-136. doi: 10.1007/s11098-006-9057-9

Hamlen, K.R. (2012). Academic dishonesty and video game play: Is new media use changing conceptions of cheating? *Computers & Education, 59*, 1145-1152. doi: 10.1016/j.compedu.2012.06.001

Harding, T.S., Carpenter, D.D., & Finelli, C.J. (2012). An exploratory investigation of the ethical behavior of engineering undergraduates. *Journal of Engineering Education, 101*, 346-374. doi: 10.1002/j.2168-9830.2012.tb00053.x

Harding, T.S., Carpenter, D.D., Finelli, C.J., & Passow, H.J. (2004). Does academic dishonesty relate to unethical behavior in professional practice? An exploratory study. *Science and Engineering Ethics, 10*, 1-14. doi: 10.1007/s11948-004-0027-3

Harding, T.S., Mayhew, M.J., Finelli, C.J., & Carpenter, D.D. (2007). The theory of planned behavior as a model of academic dishonesty in engineering and humanities undergraduates. *Ethics & Behavior, 17*, 255-279. doi: 10.1080/10508420701519239

Hargreaves, D.H. (1995). School culture, school effectiveness and school improvement. *School effectiveness and school improvement, 6*, 23-46. doi: 10.1080/0924345950060102

Harlen, W. (2006). The Role of Assessment in Developing Motivation for Learning. In J. Gardner (Ed.), *Assessment and Learning* (pp. 61-80). London: Sage Publications, Ltd.

Harmon, O.R., & Lambrinos, J. (2008). Are online exams an invitation to cheat? *The Journal of Economic Education, 39*, 116-125. doi: 10.3200/JECE.39.2

Harris P.L., Núñez, M., & Brett, C. (2001). Let's swap: Early understanding of social exchange by British and Nepali children. *Memory & Cognition, 29*, 757–764. doi: 10.3758/BF03200478

Harter, S. (1999). *The construction of the self: A developmental perspective*. New York: Guilford Press.

Hartshorne, H., & May, M.A. (1928). *Studies in deceit*. Norwood, MA: Norwood Press.

Hattie, J., & Timperley, H. (2009). The power of feedback. *Review of Educational Research, 77*, 81-112. doi: 10.3102/00346543029848

Heckler, N.C., Forde, D.R., & Bryan, C.H. (2013). Using writing assignment designs to mitigate plagiarism. *Teaching Sociology, 41*, 94-105. doi: 10.1177/0092055X12461471

Hershberger, S.L. (2003). The growth of structural equation modeling: 1994 – 2001. *Structural Equation Modeling: A Multidisciplinary Journal, 10*, 35-46. doi: 10.1207/S15328007SEM1001_2

Herzog, W., & Boomsma, A. (2009). Small-sample robust estimators of noncentrality-based and incremental model fit. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*, 1-27. doi: 10.1080/10705510802561279

Herzog, W., Boomsma, A., & Reinecke, S. (2007). The Model-Size Effect on Traditional and Modified Tests of Covariance Structures. *Structural Equation Modeling: A Multidisciplinary Journal, 14*, 361-390. doi: 10.1080/10705510701301602

Hidi, S., & Renninger, K.A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*, 111-127. doi: 10.1207/s15326985ep4102_4

Higbee, J.L., Schultz, J.L., & Sanford, T. (2011). Student perspectives on behaviors that constitute cheating. *Contemporary Issues in Education Research, 4*(10), 1-8.

Hollinger, RC., & Lanza-Kaduce, L. (1996). Academic dishonesty and the perceived effectiveness of countermeasures: An empirical survey of cheating at a major public university. *NASPA Journal, 33*(4), 292-306.

Holmes-Smith, P. (2012). *Advanced structural equation modeling using AMOS*. Lecture notes for ACSPRI 2012 Winter Program. University of Technology, Sydney.

Holmes-Smith, P., & Rowe, K.J. (1994). *The development and use of congeneric measurement models in school effectiveness research: Improving the reliability and validity of composite and latent variables for fitting multilevel and structural equation models*. Paper presented at the International Congress for School Effectiveness and Improvement, The World Congress Centre, Melbourne, Australia, January 3rd – 6th, 1994.

Horn, J.L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185. doi: 10.1007/BF02289447

Horst, P. (1941). The role of the predictor variables which are independent of the criterion. *Social Science Research Council, 48*, 431--436.

Houston, J.P. (1977). Cheating behavior, anticipated success-failure, confidence, and test importance. *Journal of Educational Psychology, 69*, 55-60. doi: 10.1037/0022-0663.69.1.55

Houston, J.P. (1983a). College classroom cheating, threat, sex and prior performance. *College Student Journal*, *17*(3), 229-235.

Houston, J.P. (1983b). Kohlberg-type moral instruction and cheating behavior. *College Student Journal, 17*(2), 196-204.

Houston, J.P. (1986). Classroom answer copying: Roles of acquaintanceship and free versus assigned seating. *Journal of Educational Psychology, 78*, 230-232. doi: 10.1037/0022-0663.78.3.230

Howard, R.M., & Davies, L.J. (2009). Plagiarism in the internet age. *Literacy 2.0, 66*(6), 64-67.

Hsiao, C., & Yang, C. (2011). The impact of professional unethical beliefs on cheating intention. *Ethics & Behavior, 21*, 301-316. doi: 10.1080/10508422.2011.585597

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1-55. doi: 10.1080/10705519909540118

Hu, L., Bentler, P.M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin, 112*, 351-362.

Hudd, S.S., Apgar C., Bronson, E.F. Lee, R.G. (2009). Creating a campus culture of integrity: Comparing the perspectives of full- and part-time faculty. *The Journal of Higher Education, 80*, 146-177. doi: 10.1353/jhe.0.0039

Jackson, D.L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N:q hypothesis. *Structural Equation Modeling: A Multidisciplinary Journal, 10*, 128-141. doi: 10.1207/S15328007SEM1001_6

Jacobson, L.I., Stephen, E.B., & Millham, J. (1970). Individual differences in cheating during a temptation period when confronting failure. *Journal of Personality and Social Psychology, 15*, 48-56. doi: 10.1037/h0029204

Jang, H., Reeve, J., Ryan, R. M., & Kim, A. (2009). Can self-determination theory explain what underlies the productive, satisfying learning experiences of collectivistically-oriented South Korean adolescents? *Journal of Educational Psychology, 101*, 644-661. doi: 10.1037/a0014241

Jensen, L.A., Arnett, J.J., Feldman, S.S., & Caufman, E. (2002). It's wrong, but everybody does it: Academic dishonesty among high school and college students. *Contemporary Educational Psychology, 27*, 209-228. doi: 10.1006/ceps.2001.1088

Johnson, C.D., & Gormly, J. (1972). Academic cheating: The contribution of sex, personality, and situational variables. *Developmental Psychology, 6*, 320-325. doi: 10.1037/h0032123

Johnson, P.B., & Richter, L. (2004). Research note: What if we're wrong? Some implications of the systemic distortions in adolescents' self-reports of sensitive behaviors. *Journal of Drug Issues, 34*, 951-970. doi: 10.1177/002204260403400412

Jordan, A.E. (2001). College student cheating: The role of motivation, perceived norms, attitudes, and knowledge of institutional policy. *Ethics & Behavior, 11*, 233-247. doi: 10.1207/S15327019EB1103_3

Jöreskog, K.G. (1963). *Statistical estimation in factor analysis: A new technique and its foundation*. Stockholm: Almqvist & Wiksell.

Jöreskog, K.G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika, 34*(2), 183-202.

Jöreskog, K.G. (1973). A general method for estimating a linear structural equation system. In A.S. Goldberger & O.D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 85-112). New York: Seminar.

Jöreskog, K.G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Chicago: Scientific Software International

Jöreskog, K.G., Sörbom, D., Du Toit, S., & Du Toit, M. (2001). *LISREL 8: New statistical features*. Lincolnwood: Scientific Software International.

Josephson Institute (2000-2012). Report card on the ethics of American youth. Retrieved from Character Counts website: http://charactercounts.org/programs/reportcard/

Jurdi, R. Hage, H.S., & Chow, H.P.H. (2011a). Academic dishonesty in the Canadian classroom: Behaviours of a sample of university students. *Canadian Journal of Higher Education, 41*, 1-35. doi: 10.1111/hea.12000_38

Jurdi, R., Hage, H.S., & Chow, H.P.H. (2011b). What behaviours do students consider academically dishonest? Findings from a survey of Canadian undergraduate students. *Social Psychology of Education, 15*, 1-23. doi: 10.1007/s11218-011-9166-y

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kant, I. (1797/1996). *The metaphysics of morals* (W. S. Pluhar, Trans.). Indianapolis: Hackett Publishing Company.

Keesing, R.J. (1974). Theories of culture. *Annual Review of Anthropology, 3*, 73-97. doi: 10.1146/annurev.an.03.100174.000445

Kelley, E.L. (1927). *Interpretation of educational measurements*. Yonkers, NY: World.

Kelley, K., & Maxwell, S.E. (2010). Multiple regression. In G.R. Hancock and R.O. Mueller (Eds.) *The reviewer's guide to quantitative methods in the social sciences* (pp.281-297). New York: Routledge.

Kelly, J.A., & Worell, L. (1978). Personality characteristics, parenting behaviors, and sex of subject in relation to cheating. *Journal of Research in Personality, 12*, 179-188. doi: 10.1016/0092-6566(78)90094-6

Keltikangas-Järvinen, L., & Lindeman, M. (1997) Evaluation of theft, lying, and fighting in adolescence. *Journal of Youth and Adolescence, 26*, 467-483. doi: 10.1023/A:1024585406173

Kember, D. (2004). Interpreting student workload and the factors which shape students' perceptions of their workload. *Studies in Higher Education, 29*. doi: 10.1080/0307507042000190778

Kember, D., & Gow, L. (1989). A model of student approaches to learning encompassing ways to influence and change approaches. *Instructional Science*, 22, 263-288. doi: 10.1007/BF00118014

Kenny, D.A. (2014a). *Estimation with instrumental variables*. Retrieved from: http://davidakenny.net/cm/iv.htm.

Kenny, D.A. (2014b). *Measuring model fit*. Retrieved from http://davidakenny.net/cm/fit.htm.

Kenny, D.A., Kaniskan, B., & McCoach, D.B. (2011). *The performance of RMSEA in models with small degrees of freedom*. Unpublished paper, University of Connecticut, retrieved from http://davidakenny.net/cm/fit.htm

Kidwell, L.A. (2001). Student honor codes as a tool for teaching professional ethics. *Journal of Business Ethics, 29*, 45-49. doi: 10.1023/A:1006442925586

King, D.L., & Case, C.J. (2014). E-cheating: Incidence and trends among college students. *Issues in Information Systems, 15*(1), 20-27.

Kisamore, J.L., Stone, T.H., & Jawahar, I.M. (2007). Academic integrity: The relationship between individual and situational factors on misconduct contemplations. *Journal of Business Ethics, 75*, 381-394. doi: 10.1007/s10551-006-9260-9

Kline, R.B. (2009). *Becoming a behavioral science researcher.* New York: The Guilford Press.

Kline R.B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: The Guilford Press.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*, *314*, 829-832. doi: 10.1126/science.1129156

Kohlberg L. (1958). *The Development of Modes of Thinking and Choices in Years 10 to 16* (doctoral dissertation). Chicago, IL: University of Chicago.

Kohlberg, L. (1968). The child as a moral philosopher. *Psychology Today, 2*, 25-30. doi: 10.1037/e400132009-002

Kohlberg, L. (1971). *From* Is *to* Ought*: How to Commit the Naturalistic Fallacy and Get Away with It in the Study of Moral Development*. New York: Academic Press.

Kohn, A. (2007a). Who's cheating whom? *The Education Digest, 73(5)*, 4-11.

Kohn, A. (2007b). Foreword. In E.M. Anderson & T.B. Murdock (Eds.), *Psychology of academic cheating* (pp. XI-XIX). Burlington, MA: Elsevier.

Kornilova, T.V., Kornilov, S.A., & Chumakova, M.A. (1990). Subjective evaluations of intelligence and academic self-concept predict academic achievement: Evidence from a selective student population. *Learning and Individual Differences, 19*, 596-608. doi: 10.1016/j.lindif.2009.08.001

Koul, R., Clariana, R.B., Jitgarun, K., & Songsriwittaya, A. (2009). The influence of achievement goal orientation on plagiarism. *Learning and Individual Differences 19*, 506-512. doi: 10.1016/j.lindif.2009.05.005

Labeff, E.E., Clark, R.E., Haines, V.J., & Diekhoff, G.M. (1990). Situational ethics and college student cheating. *Sociological Inquiry, 60*, 190-198. doi: 10.1111/j.1475-682X.1990.tb00138.x

Latova, N.V., & Latov, I.V. (2008). Cheating in the educational process. *Russian Education and Society, 50*, 17-43. doi: 10.2753/RES1060-9393500702

Lattin, J., Carroll, D.J., & Green, P.E. (2003). *Analyzing multivariate data*. Pacific Grove, CA: Duxbury Press.

Lau, G.K.K., Yuen, A.H.K., & Park, J. (2013). Toward an analytical model of ethical decision making in plagiarism. *Ethics & Behavior, 23*, 360-377. doi: 10.1080/10508422.2013.787360

Lavelle, L. (2008). GMAT cheating controversy grows. *Businessweek*, p. 13.

Lawless, C.J., & Richardson, J.T.E. (2002). Approaches to studying and perceptions of academic quality in distance education. *Higher Education, 44*, 257-282. doi: 10.1023/A:1016315114558

Lee, J., Bong, M., & Kim, S. (2014). Interaction between task values and self-efficacy on maladaptive achievement strategy use. *Educational Psychology: An International Journal of Experimental Educational Psychology*. doi: 10.1080/01443410.2014.895296

Lee, K., Ashton, M.C., Shin, K. (2005). Personality correlates of workplace anti-social behavior. *Applied psychology: An international review, 54*, 81-98. doi: 10.1111/j.1464-0597.2005.00197.x

Leming, J.S. (1980). Cheating behavior, subject variables, and components of the internal-external scale under high and low risk conditions. *The Journal of Educational Research, 74*(2), 83-87.

Leonard, V., & LeBrasseur, R. (2008). Individual assignments and academic dishonesty – Exploring the learning conundrum. *Australian Educational Researcher, 35*, 37-56. doi: 10.1007/BF03216874

Leung, K.C., Marsh, H.W., Yeung, A.S., & Abduljabbar, A.S. (2015). Validity of social, moral and emotional facets of Self-Description Questionnaire II. *The Journal of Experimental Education, 83*, 1-23. doi: 10.1080/00220973.2013.876229

Levitt, S.D., & Dunbar, S.J. (2005). *Freakonomics: A rogue economist explores the hidden side of everything*. New York: Morrow.

Lim, V.K.G., & See, S.K.B. (2001). Attitudes toward, and intentions to report, academic cheating among students in Singapore. *Ethics & Behavior, 11*, 261-274. doi: 10.1207/S15327019EB1103_5

Lobel, T.E., & Levanon, I. (1988). Self-esteem, need for approval, and cheating behavior in children. *Journal of Educational Psychology, 80*, 122-123. doi: 10.1037/0022-0663.80.1.122

Locke, J. (1689/1996). *An essay concerning human understanding*. In K.P. Winkler (Ed.). Indianapolis: Hackett Publishing Company.

Lorenzo-Seva, U., & Ferrando, P.J. (2006). FACTOR: A computer program to fit the exploratory analysis model. *Behavior Research Methods, 38*, 88-91. doi: 10.3758/BF03192753

Lorenzo-Seva, U., & Ferrando, P.J. (2007). *Factor, version 7.0*. Tarragona, Spain: Rovira I Virgili University.

Ludeman, W.W. (1938). A study of cheating in public schools. *American School Board Journal, 96*, 45.

Ma, H., Lu, E.R., Turner, S., & Wan, G. (2007). An empirical investigation of digital cheating and plagiarism among middle school students. *American Secondary Education, 35*(2), 69-82.

MacCallum, R.C., & Austin, J.T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201-226. doi: 10.1146/annurev.psych.51.1.201

MacCallum, R.C., Browne, M.W., & Cai, L. (2006). Testing differences between nested covariance structure models: Power analysis and null hypotheses. *Psychological Methods, 11*, 19-35. doi: 10.1037/1082-989X.11.1.19

MacCallum, R.C., Browne, M.W., & Sugawara, H.M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149. doi: 10.1037/1082-989X.1.2.130

Machery, E., & Mallon, R. (2010). Evolution of morality. In J.M. Doris (Ed.), *The moral psychology handbook* (pp. 3-46). Oxford: Oxford University Press.

MacKinnon, D.P., Krull, J.L., & Lockwood, C.M. (2000). Equivalence of the mediation, confounding, and suppression effect. *Prevention Science, 1*, 173-181. doi: 10.1023/A:1026595011371

MacKinnon, D.P., Fairchild, A.J., & Fritz, M.S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593-614. doi: 10.1146/annurev.psych.58.110405.085542

MacNeil, I.R. (1974). The many futures of contracts. *Southern California Law Review, 47*, 691-816.

MacNeil, I.R. (1980). *The new social contract: An inquiry into modern contractual relations*. New Haven: Yale University Press.

Maehr, M.L. (1983). On doing well in science: Why Johnny no longer excels; why Sarah never did. In S.G. Paris, G.M. Olson, & H.W. Stevenson (Eds.), *Learning and Motivation in the Classroom* (pp. 179-210). Hillsdale, NJ: Erlbaum.

Magnus, J.R., Poltervich, V.M., Danilov, D.L., & Savvateev, A.V. (2002). Tolerance of cheating: An analysis across countries. *The Journal of Economic Education, 33*, 125-135. doi: 10.1080/00220480209596462

Markland, D. (2007). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modeling. *Personality and Individual Differences, 42*, 851-858. doi: 10.1016/j.paid.2006.09.023

Mallon, R., & Nichols, S. (2010). Rules. In Doris, J.M. and the Moral Research Group (Eds.), *The moral psychology handbook* (pp. 297-320). Oxford, UK: Oxford University Press.

Marsh, H.W. (1989). *The Self Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent self-concept: An interim test manual and a research monograph.* Preliminary Draft: University of Western Sydney.

Marsh, H.W. (1992). *SDQ-II: Manual*. Sydney, Australia: SELF Research Center, University of Western Sydney.

Marsh, H.W. (1994). Sport motivation orientations: Beware of the jingle-jangle fallacies. *Journal of Sport and Exercise Psychology, 16*(4), 365-380.

Marsh, H.W., Byrne, B.M., & Yeung, S.Y. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and revised recommendations. *Educational Psychologist*, 34, 155–167. doi: 10.1207/s15326985ep3403_2

Marsh, H.W., Craven, R.G., Hinkley, J.W., & Debus, R.L. (2003). Evaluation of the big-two factor theory of academic motivation orientations: An evaluation of jingle-jangle fallacies. *Multivariate Behavioral Research, 38*, 189-224. doi: 10.1207/S15327906MBR3802_3

Marsh, H.W., Ellis, L.A., Parada, R. H., Richards, G., Heubeck, B. G. (2005). A short version of the self description questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychological Assessment, 17*, 81-102. doi: 10.1037/1040-3590.17.1.81

Marsh, H.W., Ginns, P., Morin, A.J.S., Nagengast, B., & Martin, A.J. (2011). Use of student ratings to benchmark universities: Multilevel modeling of responses to the Australian Course Experience Questionnaire (CEQ). *Journal of Educational Psychology, 103*, 733-748. doi: 10.1037/a0024221.

Marsh, H.W., Hau, K., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexs and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*, 320-341. doi: 10.1207/s15328007sem1103_2

Marsh, H.W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin, 97,* 562–582. doi: 10.1037/0033-2909.97.3.562

Marsh, H.W., Nagengast, B., & Morin, A.J. (2013). Measurement invariance of Big-Five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and La Dolce Vita effects. *Developmental Psychology*, *49*, 1194-1218. doi: 10.1037/a0026913

Martin, A.J. (2011). Prescriptive Statements and Educational Practice: What Can Structural Equation Modeling (SEM) Offer? *Educational Psychology Review, 23*, 235-244. doi: 10.1007/s10648-011-9160-0

Martin, D.E., Rao, A., & Sloan, L.R. (2009). Plagiarism, integrity, and workplace deviance: A criterion study. *Ethics and Behavior, 19*, 36-51. doi: 10.1080/10508420802623666

Marton, F. (1976). What does it take to learn? Some implications of an alternative view of learning. In N.J. Entwistle (Ed.), *Strategies for Research and Development in Higher Education*. Amsterdam: Swetz & Zeitlinger.

Marton, F., & Säljö, R. (1976). On qualitative differences in learning, outcome and process I. *British Journal of Educational Psychology, 46*, 4-11. doi: 10.1111/j.2044-8279.1976.tb02980.x

Marton, F., & Säljö, R. (1997). Approaches to Learning. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The Experience of Learning* (pp. 39-58). Edinburgh: Scottish Academic Press.

Mather, H. (1999). *Contract law and morality*. Westport, Ct: Greenwood Press.

Mayhew, M.J., Hubbard, S.M., Finelli, C.J., Harding, T.S., & Carpenter, D.D. (2009). Using structural equation modeling to validate the theory of planned behavior as a model for predicting student cheating. *The Review of Higher Education, 32*, 441-468. doi: 10.1353/rhe.0.0080

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*, 633-644. doi: 10.1509/jmkr.45.6.633

McArdle, J.J., & Nesselroade, J.R. (2003). Growth curve analysis in contemporary psychology. In J.A. Schinka & W.F. Velicer (Eds.), *Handbook of psychology: Research methods in psychology* (pp. 447-480). New York: Wiley.

McCabe, D.L. (1992). The influence of situational ethics on cheating among college students. *Social Inquiry, 62*, 365-374. doi: 10.1111/j.1475-682X.1992.tb00287.x

McCabe, D.L. (1997). Classroom cheating among natural science and engineering majors. *Science and Engineering Ethics, 3*, 443-445. doi: 10.1007/s11948-997-0046-y

McCabe, D.L. (1999). Academic dishonesty among high school students. *Adolescence, 34*(136), 681-687.

McCabe, D.L. (2005). It takes a village: academic dishonesty and educational opportunity. *Liberal Education, 91*(3), 26-31.

McCabe, D.L., Butterfield, K.D., & Treviño, L.K. (2003). Faculty and academic integrity: The influence of current honor codes and past honor code experiences. *Research in Higher Education, 44*, 367-385. doi: 10.1023/A:1023033916853

McCabe, D.L., Feghali, T., & Abdallah, H. (2008). Academic dishonesty in the Middle East: Individual and contextual factors. *Research in Higher Education, 49*, 451-467. doi: 10.1007/s11162-008-9092-9

McCabe, D.L., & Katz, D. (2009). Curbing cheating. *Education Digest, 75*(1), 16-19.

McCabe, D.L., & Treviño, L.K. (1993). Academic dishonesty: Honor codes and other contextual influences. *The Journal of Higher Education*, *64*, 522-538. doi: 10.1023/A:1014893102151

McCabe, D.L., & Treviño, L.K. (1997). Individual and contextual influences on academic dishonesty: A multicampus investigation. *Research in Higher Education, 38*, 379-396. doi: 10.1023/A:1024954224675

McCabe, D.L., Treviño, L.K., & Butterfield, K.D. (2002). Honor codes and other contextual influences on academic integrity: A replication and extension to modified honor code settings. *Research in Higher Education, 43*, 357-378. doi: 10.1023/A:1014893102151

McCrosky, J., & Young, T. (1981). Ethos and credibility: The construct and its measurements after two decades. *Central States Speech Journal*, *32*, 24–34. doi: 10.1080/10510978109368075

McNabb, L. (2010). An update on student authentication: Implementation in context. *Continuing Higher Education Review, 74*(1), 43-52

McTernan, M., Love, P., & Rettinger, D. (2014). The influence of personality on the decision to cheat. *Ethics & Behavior*, *24*, 53-72. doi: 10.1080/10508422.2013.819783

Meade, A.W., Johnson, E.C., & Braddy, P.W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*, 568-592. doi: 10.1037/0021-9010.93.3.568

Meade, J. (1992). Cheating: Is academic dishonest par for the course? *Prism, 1*(7), 30-32.

Medina, V. (1990). *Social contract theories: Political obligation or anarchy?* Savage, Md: Rowman & Littlefield Publishers.

Meece, J.L., Anderman, E.M., & Anderman, L.H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology, 57*, 487-503. doi: 10.1146/annurev.psych.56.091103.070258

Midgley, C., Maehr, M.L., Hruda, L., Anderman, E.M., Anderman, L., Freeman, K.E., Gheen, M., Kaplan, A., Kumar, R., Middleton, M.J., Nelson, J., Roeser, R., & Urdan, T. (2000). *Manual for the patterns of adaptive learning scales (PALS)*. Ann Arbor, MI: University of Michigan.

Miller, A.D., Murdock, T.B., Anderman, E.M., & Poindexter, A.L. (2007). Who are all these cheaters? Characteristics of academically dishonest students. In E. M. Anderman & T. B. Murdock (Eds.), *Psychology of Academic Cheating* (pp. 9-32). San Diego, CA: Elsevier.

Miller, A., Shoptaugh, C., & Parkerson, A. (2008). Under reporting of cheating in research using volunteer college students. *College Student Journal, 42*(2), 326-339.

Miller, A., Shoptaugh, C., & Wooldridge J. (2011). Reasons not to cheat, academic-integrity responsibility, and frequency of cheating. *The Journal of Experimental Education, 79*, 169-184. doi: 10.1080/00220970903567830

Mills, J. (1958). Changes in moral attitudes following temptation. *Journal of Personality, 26*, 517-531. doi: 10.1111/j.1467-6494.1958.tb02349.x

Mills, C.M., & Neil, F.C. (2005). The development of cynicism. *Psychological Science, 16*, 385-390. doi: 10.1111/j.0956-7976.2005.01545.x

Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology, 55*, 1-22. doi: 10.1146/annurev.psych.55.042902.130709

Morawetz, T. (1986). Reconstructing the criminal defenses: The significance of justification. *The Journal of Criminal Law and Criminology, 77*, 277-307. doi: 10.2307/1143335

Moton, J., Fitterer, A., Brazier, E., Leonard, J., & Brown, A. (2013). Examining online college cyber cheating methods and prevention measures. *The Electronic Journal of e-Learning, 11*(2), 139-146.

Mueller, G.R., & Hancock, R.O. (2010). Structural equation modeling. In G.R. Hancock and R.O. Mueller (Eds.) *The reviewer's guide to quantitative methods in the social sciences* (pp. 371-383). New York: Routledge.

Mulaik, S.A. (2009). *Linear causal modeling with structural equations*. New York: CRC Press.

Murphy, P.K., & Alexander, P.A. (2000). A motivated exploration of motivation terminology. *Contemporary Educational Psychology, 25*, 3-53. doi: 10.1006/ceps.1999.1019

Murdock, T.B., & Anderman, E.M. (2006). Motivational perspectives on student cheating: Toward an integrated model of academic dishonesty. *Educational Psychologist, 41*, 129-145. doi: 10.1207/s15326985ep4103_1

Murdock, T.B., Beauchamp, A.S., & Hinton, A.M. (2008). Predictors of cheating and cheating attributions: Does classroom context influence cheating and blame for cheating? *European Journal of Psychology of Education, 23*, 477-492. doi: 10.1007/BF03172754

Murdock, T.B., Hale, N.M., & Weber, M.J. (2001). Predictors of cheating among early adolescents: Academic and social motivations. *Contemporary Educational Psychology, 26*, 96-115. doi: 10.1006/ceps.2000.1046

Murdock, T.B., Miller, A.D., & Goetzinger, A. (2007). Effects of classroom context on university students' judgments about cheating: Mediating and moderating processes. *Social Psychology of Education, 10*, 141-169. doi: 10.1007/s11218-007-9015-1

Murdock, T.B., Miller, A.D., & Kohlhardt, J. (2004). Effects of classroom context variables on high school students' judgments of the acceptability and likelihood of cheating. *Journal of Educational Psychology, 96*, 765-777. doi: 10.1037/0022-0663.96.4.765

Murdock, T.B., & Stephens, J.M. (2007). Is cheating wrong? Students' reasoning about academic dishonesty. In E.M. Anderson & T.B. Murdock (Eds.), *Psychology of academic cheating* (pp. 229-251). Burlington, MA: Elsevier.

Muthén, B.O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika, 29*(1), 81-117.

Muthén, L.K., & Muthén, B.O. (1998-2010). *Mplus User's Guide. Sixth Edition*. Los Angeles, CA: Muthén & Muthén.

Muthén, L.K., & Muthén, B.O. (2014). *Can a grouping variable change?* http://www.statmodel.com/discussion/messages/11/9132.html?1412357058

Muthén, L.K., & Muthén, B.O. (2014). *Non-positive definite matrix*. http://www.statmodel.com/ discussion/messages/11/1602.html?1398918861.

Nagin, D.S., & Pogarsky, G. (2003). An experimental investigation of deterrence: Cheating, self-serving bias, and impulsivity. *Criminology, 41*, 167-194. doi: 10.1111/j.1745-9125.2003.tb00985.x

Narvaez, D. (2010). Moral complexity: The fatal attraction of truthiness and the importance of mature moral functioning. *Perspectives on Psychological Sciences*, *5*(2), 163-181. doi: 0.1177/1745691610362351

Nathanson, C., Paulhus, D.L., & Williams, K.M. (2006). Predictors of a behavioral measure of scholastic cheating: Personality and competence but not demographics. *Contemporary Educational Psychology, 31*, 97-122. doi: 10.1016/j.cedpsych.2005.03.001

Newstead, S.E., Franklyn-Stokes, A., & Armstead, P. (1996). Individual differences in student cheating. *Journal of Educational Psychology, 88*, 229-241. doi: 10.1037/0022-0663.88.2.229

Nichols, S.L., & Berliner, D.C. (2007). The pressure to cheat in a high stakes testing environment. In E.M. Anderson & T.B. Murdock (Eds.), *Psychology of academic cheating* (pp. 289-311). Burlington, MA: Elsevier.

Nie, Y., Lau, S., & Liau, A.K. (2011). Role of academic self-efficacy in moderating the relation between task importance and test anxiety. *Learning and Individual Differences, 21*, 736-741. doi: 10.1016/j.lindif.2011.09.005

Niiya, Y., Ballantyne, R., North, M.S., & Crocker, J. (2008). Gender, contingencies of self-worth, and achievement goals as predictors of academic cheating in a controlled laboratory setting. *Basic and Applied Social Psychology, 30*, 76-83. doi: 10.1080/01973530701866656

Ning, H.K., & Downing, K. (2010) Connections between learning experience, study behaviour and academic performance: A longitudinal study. *Educational Research, 52*, 457-468. doi: 10.1080/00131881.2010.524754

Nora, W.L.Y., & Zhang, K.C. (2010). Motives of cheating among secondary students: the role of self-efficacy and peer influence. *Asia Pacific Educational Review, 11*, 573-584. doi: 10.1007/s12564-010-9104-2

Norton, L.S., Tilley, A.J., Newstead, S.E., & Franklin-Stokes, A. (2001). The pressures of assessment in undergraduate courses and their effect on student behaviours. *Assessment & Evaluation in Higher Education*, *26*, 269-284. doi: 10.1080/02602930120052422

Nucci, L.P. (2001). *Education in the moral domain.* Cambridge, UK: Cambridge University Press.

Núñez, M., & Harris, P.L. (1998). Psychological and deontic concepts: separate domains or intimate connection? *Mind & Language, 13*, 153–70. doi: 10.1111/1468-0017.00071

Nye, C.D., & Drasgow, F. (2011). Assessing goodness of fit: Simple rules of thumb simply do not work. *Organizational Research Methods, 14*, 548-570. doi: 10.1177/1094428110368562.

Ogilvie, J., & Stewart, A. (2010). The integration of rational choice and self-efficacy theories: A situational analysis of student misconduct. *The Australian and New Zealand Journal of Criminology, 43*, 130-155. doi: 10.1375/acri.43.1.130

Olafson, L., Schraw, G., Nadelson, L., Nadelson, S., & Kehrwald, N. (2013). Exploring the judgment-action gap" College students and academic dishonesty. *Ethics & Behavior, 23*, 148-162. doi: 10.1080/10508422.2012.714247

Olson, K.R., & Spelke, E.S. (2008). Foundations of cooperation in young children. *Cognition, 108*, 222-231. doi: 10.1016/j.cognition.2007.12.003

O'Rourke, J., Barnes, J., Deaton, A., Fulks, K., Ryan K., & Rettinger, D.A. (2010). Imitation is the sincerest form of cheating: The influence of direct knowledge and attitudes on academic dishonesty. *Ethics & Behavior, 20* 47-64. doi: 10.1080/10508420903482616

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25*, 1049-1079. doi: 10.1080/0950069032000032199

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, *66*, 543–548. doi: 10.3102/00346543066004543

Pandey, S., & Elliott, W. (2010). Suppressor variables in social work research: Ways to identify in multiple regression models. *Journal of the Society for Social Work and Research, 1*, 28-40. doi: 10.5243/jsswr.2010.2

Parr, F.W. (1936). The problem of student dishonesty. *Journal of Higher Education, 6*, 318-326. doi: 10.2307/1974907

Passow, H.J., Mayhew, M.J., Finelli, C.J., Harding, T.S., & Carpenter, D.D. (2006). Factors influencing engineering students' decision to cheat by type of assessment. *Research in Higher Education, 47*, 643-684. doi: 10.1007/s11162-006-9010-y

Paulhus, D.L. (1991). Measurement and control of response bias. In J.P. Robinson, P.R. Shaver, & L.S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes*. San Diego, CA: Academic, pp. 17–59.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge, UK: Cambridge University Press.

Peters, E., Västfjäll, D., Gärling, T., & Slovic, P. (2006). Affect and decision-making: A "hot" topic. *Journal of Behavioral Decision-Making*, *19*, 79–85. doi:10.1002/bdm.528

Phan, H.P. (2008). Unifying different theories of learning: theoretical framework and empirical evidence. *Educational Psychology, 28*, 325-340. doi: 10.1080/01443410701591392

Phan, H.P. (2009a). Exploring students' reflective thinking practice, deep processing strategies, effort, and achievement goal orientations. *Educational Psychology, 29*, 297-313. doi: 10.1080/01443410902877988

Phan, H.P. (2009b). Amalgamation of future time orientation, epistemological beliefs, achievement goals and study strategies: Empirical evidence established. *British Journal of Educational Psychology, 79*, 155-173. doi: 10.1348/000709908X306864

Pillutla, M.M., & Murnighan, J.K. (1996). Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational behavior and human decision processes, 68*, 208-224. doi: 10.1006/obhd.1996.0100

Pintrich, P.R. (2000). An achievement goal theory perspective on issues in motivational terminology, theory, and research. *Contemporary Educational Psychology, 25*, 92-104. doi: 10.1006/ceps.1999.1017

Pogue, L.L., & Ah-Yun, K. (2006). The effect of teacher nonverbal immediacy and credibility on student motivation and affective learning. *Communication Education*, *55*, 331–344. doi: 10.1080/03634520600748623

Pratt, C.B., & McLaughlin, G.W. (1989). An analysis of predictors of college students' ethical inclinations. *Research in Higher Education, 30*, 195-219. doi: 10.1007/BF00992717

Pulvers, K., & Diekhoff, G.M. (1999). The relationship between academic dishonesty and college classroom environment. *Research in Higher Education, 40*, 487-498. doi: 10.1023/A:1018792210076

R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rakovski, C.C., & Levy, E.S. (2007). Academic dishonesty: Perceptions of business students. *College Student Journal, 41(2)*, 466-481.

Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. *Studies in Higher Education, 16*, 129-150. doi: 10.1080/03075079112331382944

Ramsden, P. (1992). *Learning to teach in higher education*. London: Routledge.

Ramsden, P., Bowden, J.A., & Martin, E. (1988). *Students' learning and perceptions of teaching: School effectiveness reconsidered*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Ramsden, P., & Entwistle, N.J. (1981) Effects of academic departments on students' approaches to studying. *British Journal of Educational Psychology, 51,* 368-383. 10.1111/j.2044-8279.1981.tb02493.x

Ramsden, P., Martin, E., & Bowden, J. (1989). School environment and sixth form pupils' approaches to learning. *British Journal of Educational Psychology, 59,* 129–142. doi: 10.1111/j.2044-8279.1989.tb03086.x.

Rawls, J. (1971). *A theory of justice*. Cambridge, Ma: Belknap Press.

Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence for fixed congeneric components. *Multivariate Behavioral Research, 32,* 329–354. doi: 10.1207/s15327906mbr3204_2

Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy, 35*, 299-331. doi: 10.1016/S0005-7894(04)80041-8

Raykov, T. (2009). Evaluation of scale reliability for unidimensional measures using latent variable modeling. *Measurement and Evaluation in Counseling and Development, 42*. doi: 10.1177/0748175609344096.

Raykov, T. (2011). Scale validity evaluation with congeneric measures in hierarchical designs. *British Journal of Mathematical Statistical Psychology, 64*, 427-438. doi: 10.1348/000711009X484286

Raykov, T., & Maroulides, G. A. (2006). *A first course in structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.

Reeve, J., & Tseng, C.-M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology, 36*, 257-267. doi: 10.1016/j.cedpsych.2011.05.002

Renard, L. (1999, December/January). Cut and Paste 101: Plagiarism and the Net. *Educational Leadership,* 57(4), p. 41. Retrieved from http://www.ascd.org/publications/

educational-leadership/dec99/vol57/num04/Cut-and-Paste-101@-Plagiarism-and-the-Net.aspx

Renninger, K.A., Hidi, S., & Krapp, A. (1992). *The role of interest in learning and development*. New York: Lawrence Erlbaum Associates, Inc.

Rest, J.R. (1986). *Moral development: Advances in research and theory*. New York: Praeger Publishers.

Rettinger, D.A. (2007). Applying decision theory to academic integrity decisions. In E.M. Anderman & T.B. Murdock (Eds.), *Psychology of academic cheating* (pp. 141-167). Burlington, MA: Elsevier.

Rettinger, D.A., & Kramer, Y. (2009). Situational and personal causes of student cheating. *Research in Higher Education, 50*, 293-313. doi: 10.1007/s11162-008-9116-5

Richardson C.B., Mulvey, K.L., & Killen, M. (2012). Extending social domain theory with a process-based account of moral judgments. *Human Development, 55*, 4-25. doi: 1 0.1159/000335362

Richardson, J.T.E. (1990). Reliability and replicability of the Approaches to Studying Questionnaire. *Studies in Higher Education, 15*, 155-168. doi: 10.1080/03075079012331377481

Richardson, J.T.E. (1994). A British Evaluation of the Course Experience Questionnaire. *Studies in Higher Education, 19*, 59-68. doi: 10.1080/03075079412331382143

Richardson, J.T.E. (2005). Students' perceptions of academic quality and approaches to studying in distance education. *Journal of Educational Psychology, 93*, 387-415. doi: 10.1080/0141192052000310001

Rosenberg, M. (1965). *Society and Adolescent Self-image*. Middletown, CT: Wesleyan University Press.

Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review*, *60*, 141–156. doi: 10.2307/2096350

Rost, D.H., and Wild, K.P. (1994). Cheating and achievement-avoidance at school: Components and assessment. *British Journal of Educational Psychology 64*, 119-132. doi: 10.1111/j.2044-8279.1994.tb01089.x

Roth, M.T. (1995). Mesopotamian legal traditions and the laws of Hammurabi. *Chicago-Kent Law Review, 71*(1), 13-39.

Rotter, J.B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs 80(1)*, 1-28. doi: 10.1037/h0092976

Rousseau J.J. (1762/1968). *The social contract*. Harmondsworth: Penguin.

Rowe K.K.J. (2002). The measurement of latent and composite variables from multiple items or indicators: Applications in performance indicator systems. Paper presented at the RMIT Statistics Seminar Series (Background paper for Keynote Address), Camberwell, VIC: Australian Council for Educational Research. http://www.acer.edu.au/research/ programs/learningprocess.html.

Rowe K.K.J. (2004). Analysing and reporting performance indicator data: 'Caress' the data and user beware! Background paper to invited address presented at the 2004 Public Sector Performance & Reporting Conference (under the auspices of the International Institute for Research—IIR), Sydney, 19–22 April 2004. http://www.acer.edu.au/learning_processes

Rowe K.K.J., & Hill, P.W. (1998). Modeling educational effectiveness in classrooms: The use of multi-level structural equations to model students' progress. *Education Research and Evaluation: An International Journal on Theory and Practice, 4*. doi: 10.1076/edre.4.4.307.6953

Sampson, R.J. (2013). The place of context: A theory and strategy for criminology's hard problems. *Criminology, 51*, 1-31. doi: 10.1111/1745-9125.12002

Schab, F. (1969). Cheating in high school: Differences between the sexes. *The Journal for the National Association for Woman Deans, 33*, 39-42.

Schab, F. (1980). Cheating in high school: Differences between the sexes (revisited). *Adolescence, 15*(60), 959-965.

Schab F. (1991). Schooling without learning: Thirty years of cheating in high school. *Adolescence, 26*(104), 839-847.

Schlenker, B.R. (1997). Personal responsibility: Applications of the Triangle Model. In L. L. Cummings, & B. Staw (Eds.), *Research in organizational behavior*, (Vol. 19, pp. 241-301). Greenwich, Connecticut: JAI Press.

Schmelkin, L.P., Gilbert, K., Spencer, K.J., Pincus, H.S., & Silva, R. (2008). A multidimensional scaling of college students' perceptions of academic dishonesty. *The Journal of Higher Education, 79*, 587-607. doi: 10.1353/jhe.0.0021

Schmelkin, L.P., Gilbert, K.A., & Silva, R. (2010). Multidimensional scaling of high school students' perceptions of academic dishonesty. *The High School Journal, 93*, 156-165. doi: 10.1353/hsj.2010.0001

Schnall, S., Haidt, J., Clore, G.L., & Jordan, A.H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology, 34*, 1096-1109. doi: 10.1177/0146167208317771

Schraw, G., Olafson, L., Kuch, F., Lehman, T., Lehman, S., & McCrudden, M. (2007). Interest and academic cheating. In E.M. Anderson & T.B. Murdock (Eds.), *Psychology of academic cheating* (pp. 59-85). Burlington, MA: Elsevier.

Schroeder, T., Roskies, A.L., & Nichols, S. (2010). Moral motivation. In J.M. Doris (Ed.), *The moral psychology handbook* (pp. 72-110). Oxford: Oxford University Press.

Schumacker, R.E., & Lomax, R.G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). New York: Routledge.

Schwartz, S.J., Klimstra, T.A., Luyckx, K., Hale, W.W., & Meeus, W.H.J. (2012). Characterizing the self-system over time in adolescence: Internal structure and associations with internalizing symptoms. *Journal of Youth and Adolescence, 41*, 1208-1225. doi: 10.1007/s10964-012-9751-1

Schwieren, C., & Weichselbaumer, D. (2010). Does competition enhance performance or cheating? A laboratory experiment. *Journal of Economic Psychology, 31*, 241-253. doi: 10.1016/j.joep.2009.02.005

Scott, M.B., & Lyman, S.M. (1968). Accounts. *American Sociological Review, 33*, 46-62. doi: 10.2307/2092239

Segers, M., & Dochy, F. (2006). Enhancing student learning through assessment: Alignment between levels of assessment and different effects on learning. *Studies in Educational Evaluation, 32*, 171-179. doi: 10.1016/j.stueduc.2006.08.003

Seider, S., Novick, S., & Gomez, J. (2013). The effects of privileging moral or performance character development in urban adolescents. *The Journal of Early Adolescence, 33*, 786-820. doi: 10.1177/0272431612468318

Selman, R. (1980). *The growth of interpersonal understanding: Developmental and clinical analyses*. New York: Academic Press.

Selmes, I. (1986). Approaches to normal learning tasks adopted by senior secondary school pupils. *British Educational Research Journal, 12*, 15-27. doi: 10.1080/0141192860120102

Shalvi, S., Eldar, O., & Bereby-Meyer, Y. (2012). Honesty requires time (and lack of justifications). *Psychological Science, 23*, 1264-1270. doi: 10.1177/0956797612443835

Shalvi, S., Handgraaf, M.J.J., & De Dreu, C.K.W. (2011). Ethical manoeuvring: Why people avoid both major and minor lies. *British Journal of Management, 22*, S16-S27. doi: 10.1111/j.1467-8551.2010.00709.x

Sheard, J., Dick, M., Markham, S., Macdonald, I., & Walsh, M. (2003). Cheating and plagiarism: Perceptions and practices of first year IT students. Proceedings of the 7th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education, Aarhus, Denmark. doi: 10.1145/637610.544468

Shelton, J., & Hill, J.P. (1969). Effects of cheating of achievement anxiety ad knowledge of peer performance. *Developmental Psychology, 1*, 449-455. doi: 10.1037/h0028010

Shipley, L.J. (2009). Academic and professional dishonesty: Student views of cheating in the classroom and on the job. *Journalism & Mass Communication Educator*, *64*, 39-53. doi: 10.1177/107769580906400104

Shirk, E., & Hoffman, R.W. (1961). The academic setting of the dishonest student. *Improving College and University Teaching, 9*, 130-134. doi: 10.1080/00193089.1961.10534169

Simkin, M.G., & McLeod, A. (2010). Why do college students cheat? *Journal of Business Ethics, 94*, 441-453. doi: 10.1007/s10551-009-0275-x

Simon, A.S., Carr, J.R., Mccullough, S.M., Morgan, S.J., Olsen, T., & Ressel, M. (2004). Gender, students' perceptions, institutional commitments and academic dishonesty: Who reports in academic dishonesty cases? *Assessment and Evaluation in Higher Education, 29*, 75-90. doi: 10.1080/0260293032000158171

Simon, H.A. (1982). *Models of bounded rationality*. Cambridge, MA: The MIT Press.

Simon, J., Dewitte, S., & Lens, W. (2004). The role of different types of instrumentality in motivation, study strategies, and performance: Know why you learn, so you'll know what you learn! *British Journal of Educational Psychology*, *74*, 343-360. doi: 10.1348/0007099041552314

Sinnott-Armstrong, W., Young, L., & Cushman, F. (2010). Moral intuitions. In J.M. Doris (Ed.), *The moral psychology handbook* (pp. 246-272). Oxford: Oxford University Press.

Sisti, D.A. (2007). How do high school students justify internet plagiarism? *Ethics & Behavior, 17*, 215-231. doi: 10.1080/10508420701519163

Sivo, S.A., Fan, X., Witta, L.E., & Willse, J.T. (2006). The search for "optimal" cutoff properties: Fit index criteria in structural equation modeling. *Journal of Experimental Education, 74*, 267-288. doi: 10.3200/JEXE.74.3.267-288

Skarlicki, D.P., & Folger, R. (1997). Retaliation in the workplace: The roles of distributive, procedural, and interactional justice. *Journal of Applied Psychology, 82*, 434–443. doi: 10.1037//0021-9010.82.3.434

Smith, K.J., Davy, J.A., & Easterling, D. (2011). An examination of cheating and its antecedents among marketing and management majors. *Journal of Business Ethics, 50*, 63-80. doi: 10.1023/B:BUSI.0000020876.72462.3f

Smith, C.P., Ryan, E.R., & Diggins, D.R. (1972). Moral decision making: Cheating on examinations. *Journal of Personality, 40(4),* 640-660.

Sneddon, A. (2007). A social model of moral dumbfounding: Implications for studying moral reasoning and moral judgment. *Philosophical Psychology, 20*, 731-748. doi: 10.1080/09515080701694110

Sørensen, J. (2002). The strength of corporate culture and the reliability of firm performance. *Administrative Science Quarterly, 47*, 70-91. doi: 10.2307/3094891

Spear, J.A., & Miller, A.N. (2012). The effects of instructor fear appeals and moral appeals on cheating-related attitudes and behavior of university students. *Ethics & Behavior, 22*, 196-207. doi: 10.1080/10508422.2012.659603

Spencer, S.H., Barrett, P.M., & Turner, C.M. (2003). Psychometric properties of the Spence Children's Anxiety Scale with young adolescents. *Journal of Anxiety Disorders, 17*, 605-625. doi: 10.1016/S0887-6185(02)00236-0

Srull, T.K., Karabenick, S.A. (1975). Effects of personality-situation locus of control congruence. *Journal of Personality and Social Psychology, 32*, 617-628. doi: 10.1037//0022-3514.32.4.617

Stearns, S.A. (2001). The student-instructor relationship's effect on academic integrity. *Ethics & Behavior, 11*, 275-285. doi: 10.1207/S15327019EB1103_6

Steiner, M.A. (1930). Does the school strengthen or weaken the tendency to cheat? *The Journal of Educational Research, 22(5)*, 388-395.

Steiner, M.A. (1932). Cheating in school. *The School Review, 40(7)*, 535-546.

Steininger, M., Johnson, R.E., & Kirts, D.K. (1964). Cheating on college examinations as a function of situationally aroused anxiety and hostility. *Journal of Educational Psychology, 55*, 317-324. doi: 10.1037/h0042396

Stephens, J.M., & Gehlbach, H. (2007). Under pressure and underengaged: Motivational profiles and academic cheating in high school. In E. M. Anderman & T. B. Murdock (Eds.), *Psychology of academic cheating* (pp. 107-134). San Diego, CA: Elsevier.

Stephens, J.M., & Nicholson, H. (2008). Cases of incongruity: Exploring the divide between adolescents' beliefs and behavior related to academic dishonesty. *Educational Studies, 34*, 361-376. doi: 10.1080/03055690802257127

Stone, T.L., Kisamore, J.L., Jawahar, I.M., & Bolin, J.H. (2014). Making our measures match perceptions: Do severity and type matter when assessing academic misconduct offenses? *Journal of Academic Ethics, 12*, 251-270. doi: 10.1007/s10805-014-9216-0

Strom, P.S., & Strom, R.D. (2007). Cheating in middle school and high school. *The Educational Forum, 71*, 104-116. doi: 10.1080/00131720708984924

Struyven, K., Dochy, F., & Janssens, S. (2002). Student perceptions about assessment in higher education: A review. http://www.leeds.ac.uk/educol/documents/00002255.htm

Sugai G., & Horner, R. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child & Family Behavior Therapy, 24*, 23-50. doi: 10.1300/J019v24n01_03

Sykes, G.M., & Matza, D. (1957). Techniques of neutralization: A theory of delinquency. *American Sociological Review, 22*, 664-670. doi: 10.2307/2089195

Tait, H., & Entwistle, N.J. (1996). Identifying students at risk through ineffective study strategies. *Higher Education, 31*, 99-118. doi: 10.1007/BF00129109

Taras, M. (2010). Assessment for learning: Assessing the theory and evidence. *Procedia Social and Behavioral Sciences, 2*, 3015 – 3022. doi: 10.1016/j.sbspro.2010.03.457

Tas, Y., & Tekkaya, C. (2010). Personal and contextual factors associated with students' cheating in science. *The Journal of Experimental Education 78*, 440-463. doi: 10.1080/00220970903548046

Taylor, L., Pogrebin, M., & Dodge, M. (2002). Advanced placement—advanced pressures: Academic dishonesty among elite high school students. *Educational Studies, 33*(4), 403-421.

Taylor, S.T., & Lewit, D.W. (1966). Social comparison and deception regarding ability. *Journal of Personality, 34*, 94-104. doi: 10.1111/j.1467-6494.1966.tb01701.x

Teixeira, A.A.C., & Rocha, M.F. (2010). Cheating by economics and business undergraduate students: An exploratory international assessment. *Higher Education*, *59*, 663-701. doi: 10.1007/s10734-009-9274-1

Teodorescu, D., & Andrei, R. (2009). Faculty and peer influences on academic integrity: College cheating in Romania. *Higher Education, 57*, 267-282. doi: 10.1007/sl0734-008-9143-3

Thompson, B. (2000). Ten commandments of structural equation modeling. In L.G. Grimm & P.R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 261-283). Washington, DC: American Psychological Association.

Thomson, R., & Holland, J. (2002). Young people, social change and the negotiation of moral authority. *Children & Society*, *16*, 103-115. doi: 10.1002/chi.703

Thorkildsen, T.A., Golant, C.J., & Richesin, D.L. (2007). Reaping what we sow: Cheating as a mechanism of moral engagement. In E.M. Anderson & T.B. Murdock (Eds.), *Psychology of academic cheating* (pp. 171-202). Burlington, MA: Elsevier.

Thorkildsen, T.A., Nolen, S.B., & Fournier, J (1994). What is fair? Children's critiques of practices that influence motivation. *Journal of Educational Psychology, 86*, 475-486. doi: 10.1037/0022-0663.86.4.475

Thornberg, R. (2008). School children's reasoning about school rules. *Research Papers in Education, 23*, 37-52. doi: 10.1080/02671520701651029

Thucydidies (c420CE/1993). *On justice, power and human nature: Selections from the History of the Peloponnesian War.* (P. Woodruff, Trans.). Indianapolis: Hackett Publishing Company.

Tittle, C.R., & Rowe, A.R. (1973). Moral appeal, sanction threat, and deviance: An experimental test. *Social Problems, 20*, 488-498. doi: 10.1525/sp.1973.20.4.03a00080

Tomarken, J.A., & Waller, G.N. (2005). Structural equation modeling: Strengths, limitations and misconceptions. *Annual Review of Clinical Psychology, 1*, 31 – 65. doi: 10.1146/ annurev.clinpsy.1.102803.144239

Treviño, L.K. (1986). Ethical decision making in organizations: A person-situation interactionist model. *The Academy of Management Review, 11*(3), 601-617.

Trigwell, K., & Prosser, M. (1991). Improving the quality of student learning: the influence of learning context and student approaches to learning on learning outcomes. *Higher Education, 22*, 251-266. doi: 10.1007/BF00132290

Trivers, R.L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*, 35-57. doi: 10.1086/406755

Trushell, J., & Byrne, K. (2013). Education undergraduates and ICT-enhanced academic dishonesty: A moral panic? British Journal of Educational Technology, 44, 6-19. doi: 10.1111/j.1467-8535.2012.01381.x

Trushell, J., Byrne, K., & Hassan, N. (2013). ICT facilitated access to information and undergraduates' cheating behaviours. Computers & Education, 63, 151-159. doi: 10.1016/j.compedu.2012.12.006

Tsai, C. (2012). Peer effects on academic cheating among high school students in Taiwan. *Asia Pacific Educational Review, 13*, 147-155. doi: 10.1007/s12564-011-9179-4

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, UK: Cambridge University Press.

Turiel, E. (2002). *The culture of morality: Social development, context, and conflict*. Cambridge, UK: Cambridge University Press.

Turiel, E. (2006). The development of morality. In W. Damon and R.M. Lerner (Eds.), *Handbook of child Psychology* (6th ed., pp. 789–857). New York: Wiley

Tylor, E. (1920) [1871]. *Primitive Culture*. New York: J. P. Putnam's Sons

Tzelgov, J., & Henik, A. (1991). Suppression situations in psychological research: Definitions, implications and applications. *Psychological Bulletin, 109*, 524-536. doi: 10.1037/0033-2909.109.3.524

Vandenberg, R., & Lance, C. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4 – 70. doi: 10.1177/109442810031002

Van Gundy, K., Morton, B.A., Liu, H.Q., & Kline, J. (2006). Effects of web-based instruction on math anxiety, the sense of mastery, and global self-esteem: A quasi-experimental study of undergraduate statistics students. *Teaching Sociology, 34*(4), 370-388.

Van Lier, J., Revlin, R., & De Neys, W. (2013). Detecting cheaters without thinking: Testing the automaticity of the cheater detection module. *PLoS ONE 8*, 1-8. doi: 10.1371/ journal.pone.0053827

Vinski, E.J., & Tryon, G.S. (2009). Study of a cognitive dissonance intervention to address high school students' cheating attitudes and behaviors. *Ethics & Behavior, 19*, 218-226. doi: 10.1080/10508420902886692

Vitro, F.T., & Schoer, L.A. (1972). The effects of probability of test success, test importance, and risk of detection on the incidence of cheating. *Journal of School Psychology, 10*, 269-277. doi: 10.1016/0022-4405(72)90062-3

Vowell, P.R., & Chen, J. (2004). Predicting academic misconduct: A comparative test of foud sociological explanations. *Sociological Inquiry, 74*, 226-249. doi: 10.1111/j.1475-682X.2004.00088.x

Waldrip, B.G., Fisher, D. L., & Dorman, J. (2009). Identifying exemplary science teachers through their students' perceptions of the assessment process. *Research in Science & Technological Education, 27*, 117-129. doi: 10.1080/02635140802658958

Walker, J. (2010). Measuring plagiarism: Researching what students do, not what they say they do. *Studies in Higher Education, 35*, 41-59. doi: 10.1080/03075070902912994

Walker, R.E., Wiemeler, G.E., Procyk, M.R., & Knake, W.P. (1966). The contagion of cheating. *Psychology in the Schools, 3*, 359-360. doi: 10.1002/1520-6807(196610)3:4<359::AID-PITS2310030412>3.0.CO;2-W

Wang, J., & Wang, X. (2012). *Structural equation modeling: Applications using Mplus*. West Sussex, UK: Higher Education Press.

Ward, D.A., & Beck, W.L. (1990). Gender and dishonesty. *The Journal of Social Psychology, 130*, 333-339. doi: 10.1080/00224545.1990.9924589

Watson, G.R., & Sottile, J. (2010). Cheating in the digital age: Do students cheat more in online courses? *Online Journal of Distance Learning Administration, 13*(1). Retrieved from http://www.westga.edu/~distance/ojdla/spring131/watson131.html

Wellborn, S.N. (1980, October 20). Cheating in college become epidemic. *U.S. News and World Report, 89*, 39-42.

Welsh. J.F. (1993). *Student dishonesty in higher education: Social context and institution response.* Unpublished report, Kansas Board of Regents, Topeka, Kansas.

Wentzel, K.R. (1998). Social relationships and motivation in middle school: The role of parents, teachers, and peers. *Journal of Educational Psychology, 90*, 202–209. doi: 10.1037/0022-0663.90.2.202

West, T., Ravenscroft, S.P., & Shrader, C.B. (2004). Cheating and moral judgment in the college classroom: A natural experiment. *Journal of Business Ethics, 54*, 173-183. doi: 10.1007/s10551-004-9463-x

Whitley, B.E. (1998). Factors associated with cheating among college students: A review. *Research in Higher Education, 39*, 235-274. doi: 10.1023/A:1018863909149

Whitley, B.E., Nelson, A.B., & Jones, C.J. (1999). Gender differences in cheating attitudes and classroom cheating behavior: A meta-analysis. *Sex Roles, 41*, 657-680. doi: 10.1023/A:1018863909149

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review, 6*, 49–78. doi: 10.1007/BF02209024

Wilkinson, L., & the Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594-604. doi: 10.1037/0003-066X.54.8.594

Williams, B.T. (2008). Trust, betrayal, and authorship: Plagiarism and how we perceive students. *Journal of Adolescent & Adult Literacy, 51*, 350-354. doi: 10.1598/JAAL.51.4.6

Wilson, J.Q. (1993). *The moral sense.* New York: Free Press.

Wilson, K.L., & Fowler, J. (2005). Assessing the impact of learning environments on students' approaches to learning. *Assessment and Evaluation in Higher Education, 30*, 85-99. doi: 10.1080/0260293042003251770

Wilson, K.L., Lizzio, A., & Ramsden, P (1997). The development, validation, and application of the course experience questionnaire. *Studies in Higher Education, 22*, 33-53. doi: 10.1080/03075079712331381121

Wong, N., Len, W., & Watkins, D. (1996). Cross-cultural validation of models of approaches to learning: An application of confirmatory factor analysis. *Educational Psychology, 16*, 317-327. doi: 10.1080/0144341960160308

Wowra, S.A. (2007b). Moral identities, social anxiety, and academic dishonesty among American college students. *Ethics & Behavior, 17*, 303-321. doi: 10.1080/10508420701519312

Yang, S.C., Huang, C., & Chen, A. (2013). An investigation of college students' perceptions of academic dishonesty, reasons for dishonesty, achievement goals, and willingness to report dishonest behavior. *Ethics & Behavior*, *23*, 501-522. doi: 10.1080/10508422.2013.802651

Yuan, K. (2005). Fit indices versus test statistics. *Multivariate Behavioral Research, 40*, 115-148. doi: 10.1207/s15327906mbr4001_5

Zito, N., & McQuillan, P. (2011). "It's not my fault": Using neutralization theory to understand cheating by middle school students. *Current Issues in Education, 13*, 2-25. doi: 10.1007/s12646-012-0161-y

Zwagerman, S. (2008). The scarlet P: Plagiarism, panopticism, and the rhetoric of academic integrity. *College Composition and Communication, 59*(4), 676-710.

APPENDICES

Appendix A:

Composite score loadings, error variances, and normalization

1. Estimate scale reliability using a weighted reliability estimation method such as *Rho* (Raykov, 2009)
2. The factor loading for the composite score is the square root of the weighted reliability estimate.
3. The error variance for the composite score is one minus the weighted reliability estimate (e.g. 1 – *Rho*).
4. Under the MODEL command in M*plus* fix the factor loading and error variance with syntax of the following form:

```
Factor BY Compscore@[loading value];
Compscore@[error variance];
```

(Holmes-Smith, 2012)

Normal equivalent deviates (NEDs) may be created, furthermore, for composite variables, using the following SPSS syntax:

```
RANK VARIABLES=[insert variable name here] (A)
  /NORMAL
  /RANK
  /PRINT=YES
  /TIES=MEAN
  /FRACTION=BLOM.
```

Appendix B: Time 1 Questionnaire items

This survey is anonymous. Please DO NOT write your name. Instead, please provide the following information as an identification code:

First 2 letters of your LAST NAME	Number of years have you gone to THIS school	MONTH Of birth (as a number)	Number of OLDER brothers and sisters that you have	First 2 letters of the STREET YOU LIVE ON

SECTION A: BACKGROUND INFORMATION

1. Your age: _____ 2. Your grade-level: ______

3. Please indicate your gender:	☐ Male	☐ Female

4. Which word best describes your English language ability?				
☐ Fluent/native	☐ High	☐ Intermediate	☐ Low	☐ Beginner

5. What language is spoken most by YOUR FAMILY at home? (please give one, only)						
	1. ☐	English	5. ☐	Hindi	9. ☐	Japanese
	2. ☐	Spanish	6. ☐	Bengali	10. ☐	German
	3. ☐	Arabic	7. ☐	Portuguese	11. ☐	Filipino/Tagalog
	4. ☐	Mandarin	8. ☐	Russian	12. ☐	French
	13. ☐	Other	If 'other', which language?			

7. What is your parent's/guardian's level of education?	**Female parent or guardian**	**Male parent or guardian**
Less than high school diploma or certificate	1. ☐	1. ☐
High school diploma or certificate	2. ☐	2. ☐
Trade or apprenticeship	3. ☐	3. ☐
University or college degree	4. ☐	4. ☐
Advanced degree (eg. Masters, PhD, medical, business, law, etc.)	5. ☐	5. ☐

SECTION B: PERSON

SUBJECT SELF-CONCEPT (Marsh et al., 2005)

1. Science is one of my best subjects.
2. I am hopeless in Science class.
3. Work in Science classes is easy for me.
4. I get good grades/marks in Science.
5. I learn things quickly in Science classes.

HONESTY-TRUSTWORTHINESS SELF-CONCEPT (Marsh, 1992)

6. I sometimes take things that belong to other people.
7. I am honest.
8. I sometimes tell lies to stay out of trouble.
9. I always tell the truth.
10. Cheating on a test is OK if I do not get caught.
11. Honesty is very important to me.
12. I sometimes cheat.
13. When I make a promise I keep it.
14. I often tell lies.
15. People can really count on me to do the right thing.

SECTION C: MOTIVATIONAL CONTEXT

MASTERY GOAL STRUCTURE (Anderman & Midgley, 2004)

16. My Science teacher thinks mistakes are okay as long as we are learning.
17. My Science teacher wants us to understand our work, not just memorize it.
18. My Science teacher really wants us to enjoy learning new things.
19. My Science teacher recognizes us for trying hard.
20. My Science teacher gives us time to really explore and understand new ideas.

PERFORMANCE GOAL STRUCTURE (Anderman & Midgley, 2004)

21. My Science teacher points out those students who get good grades as an example to all of us.
22. My Science teacher lets us know which students get the highest scores on a test.

23. My Science teacher makes it obvious when certain students are not doing well on their math work.
24. My Science teacher tells us how we compare to other students.
25. My Science teacher calls on smart students more than on other students.

SECTION D: ACADEMIC CONTEXT (MEASURES OF PERCEIVED QUALTIY)

GOOD TEACHING (Wilson et al., 1997)

26. My Science teacher motivates students to do their best work.
27. *My Science teacher gives a lot of time to commenting on students' work.
28. *My Science teacher makes a good effort to understand problems students may be having with their work.
29. My Science teacher normally gives helpful feedback about how you are doing.
30. My Science teacher is very good at explaining things to us.
31. *My Science teacher works hard to make science interesting.
32. *My Science teacher doesn't really care about what students have to say.
33. My Science class really tries to get the best out of all students.

APPROPRIATE ASSESSMENT (Wilson et al., 1997)

34. To do well in my Science class, all you really need is a good memory.
35. *My Science teacher seems to care more about what you've memorized than what you've understood.
36. My Science teacher asks us too many questions just about facts.
37. *It would be possible to succeed in my Science class just by studying for tests and quizzes the night before.

CLEAR GOALS AND STANDARDS (Wilson et al., 1997)

38. *It's always easy in my Science class to know what I need to do to get a good grade.
39. *In my Science class, you usually have a clear idea of what you're supposed to do.
40. *It's often hard to figure out what the teacher expects of you in my Science class.
41. *The goals and purposes of my Science class are NOT made very clear.
42. My Science teacher makes it clear right from the start what she/he expects from students.

APPROPRIATE WORKLOAD (Wilson et al., 1997)

43. *The amount of work in my Science class is too large.
44. *It seems to me that my Science teacher tries to cover too much material.
45. *In my Science class, we are usually given enough time to understand the things we have to learn.
46. There's a lot of pressure on you as a student in my Science class.
47. *The large amount of work you have to do in my Science class means you can't understand it all completely.

USEFULNESS OF CURRICULUM (Rowe & Hill, 1998)

48. In my Science class I learn things that will be useful to me when I leave school.
49. *What I learn in my Science class will be useful to me later on as a student.
50. What I learn in my Science class is useful to me.
51. What I learn in my Science class will be useful to me in the future.

TRANSPARENCY OF ASSESSMENT (Dorman & Knightley, 2006)

52. *I understand what is needed in all graded assignments in my Science class.
53. *I know what is needed to successfully accomplish graded assignments in my Science class.
54. *I know in advance HOW I will be graded in my Science class.
55. *I am told in advance WHY I am being ask to do graded assignments in my Science class.
56. *I am told in advance WHEN I will be graded in my Science class.
57. *I am told in advance WHAT science topics and information I will be graded on in my Science class.
58. *I understand the purpose of graded assignments in my Science class.

AUTHENTICITY OF ASSESSMENT (Dorman & Knightley, 2006)

59. I am asked to apply my learning to real-life situations in my Science class.
60. *In my Science class, graded assignments are meaningful.
61. *In my Science class, graded assignments are useful.
62. *I find that in my Science class, graded assignments relate to the real world.
63. *In my Science class, graded assignments check my understanding of topics.
64. *In my Science class, graded assignments test my ability to use what I've learned.

65. *In my Science class, graded assignments examine my ability to answer important questions.

EXPERIENCE OF SCHOOL RULES (Gregory, 2010)

66. *Everyone knows the rules for how students should behave in my Science class.
67. *The rules in my Science class are fair.
68. *The punishment for breaking rules in my Science class is the same no matter who you are.
69. *My Science teacher makes sure that everyone follows the rules in class.
70. *If a rule is broken in my Science class, students know what the teacher will do about it.
71. *If a student breaks the rules in my Science class, the teacher will do something about it.

PEER NORMS (Mayhew et al., 2009)

72. *If I cheated on a test in Science class this year, most of my classmates would think that's okay.
73. *Most of my classmates in Science class this year would be willing to cheat on a Science test to avoid failing.
74. *Most of my classmates would NOT think it's okay if I cheated in Science class this year.
75. *Most of my classmates think that I should NOT cheat in Science class.
76. *My classmates will look down on me if I cheat in Science class this year.
77. *Most of my classmates expect me to cheat in my Science class this year.
78. *None of my classmates think it is okay to cheat in my Science class this year

SECTION E: LEARNING STRATEGIES

DEEP LEARNING STRATEGIES (Anderman et al., 1998)

79. When working on a Science problem, I try to see how it connects with something in everyday life.
80. When I make mistakes in Science, I try to figure out why.
81. I try to connect new work in Science to what I've learned before.
82. I take my time to figure out my work in Science.
83. If I can't solve a Science problem one way, I try to use a different way.
84. I spend some time thinking about how to do my Science work before I start it.
85. I ask myself questions when I work on Science to make sure I understand.

SURFACE LEARNING STRATEIGES (Simon et al., 2004)

86. *I study for Science class by rehearsing (repeating over and over) important information.
87. *I study for Science class by memorizing things I do not understand.
88. *I study for Science class by rehearsing and repeating the material over and over until I can write it exactly, word-for-word.
89. I study for Science class by skipping over parts I think the teacher will not ask questions about.
90. I study for Science class by skipping parts I do not understand.
91. I study for Science class by skipping parts I do not find important.

SECTION F: ACADEMIC INTEGRITY

JUSTIFIABILITY OF CHEATING IN SCIENCE CLASS (Murdock et al., 2004)

92. Students would have a good reason to cheat on a test in my Science class.
93. Students would be justified to cheat on an exam in my Science class.
94. **It's reasonable to cheat in my Science class.
95. **I can understand why students would cheat in my Science class.

ACTUAL CHEATING IN SCIENCE CLASS (Midgley et al., 2000)

96. *I sometimes cheat on Science tests, this year.
97. **I sometimes cheat on my Science class work, this year. (Anderman et al., 1998)
98. * I have cheated on Science class work by copying answers from other students this year.
99. * I have cheated in Science class this year.

Appendix C: Original and modified item wording

Adjusted	Original
Self-concept SUBJECT SELF-CONCEPT (Marsh et al., 2005) 1. Science is one of my best subjects. 2. I am hopeless in Science class. 3. I often need help in Science. 4. I look forward to Science classes. 5. I have trouble understanding anything that involves Science. 6. Work in Science classes is easy for me. 7. I do badly at tests of Science. 8. I enjoy studying for Science. 9. I get good grades/marks in Science. 10. I never want to take another Science course. 11. I have always done well in Science. 12. I hate Science. 13. I learn things quickly in Science classes. HONESTY-TRUST. SELF-CONCEPT (Marsh, 1992) 14. I sometimes take things that belong to other people. 15. I am honest. 16. I sometimes tell lies to stay out of trouble. 17. I always tell the truth. 18. Cheating on a test is OK if I do not get caught. 19. Honesty is very important to me. 20. I sometimes cheat. 21. When I make a promise I keep it. 22. I often tell lies. 23. People can really count on me to do the right thing.	**Self-concept** SUBJECT SELF-CONCEPT (Marsh et al., 2005) 1. Science is one of my best subjects. 2. I am hopeless in Science class. 3. I often need help in Science. 4. I look forward to Science classes. 5. I have trouble understanding anything that involves Science. 6. Work in Science classes is easy for me. 7. I do badly at tests of Science. 8. I enjoy studying for Science. 9. I get good grades/marks in Science. 10. I never want to take another Science course. 11. I have always done well in Science. 12. I hate Science. 13. I learn things quickly in Science classes. HONESTY-TRUST. SELF-CONCEPT (Marsh, 1992) 14. I sometimes take things that belong to other people. 15. I am honest. 16. I sometimes tell lies to stay out of trouble. 17. I always tell the truth. 18. Cheating on a test is OK if I do not get caught. 19. Honesty is very important to me. 20. I sometimes cheat. 21. When I make a promise I keep it. 22. I often tell lies. 23. People can really count on me to do the right thing.

Motivational Context Scales MASTERY GOAL STRUCTURE (Midgley et al., 2000) 24. My Science teacher thinks mistakes are okay as long as we are learning. 25. My Science teacher wants us to understand our work, not just memorize it. 26. My Science teacher really wants us to enjoy learning new things. 27. My Science teacher recognizes us for trying hard. 28. My Science teacher gives us time to really explore and understand new ideas. PERFORMANCE GOAL STRUCTURE (Anderman & Midgley, 2004) 29. My Science teacher points out those students who get good grades as an example to all of us. 30. My Science teacher lets us know which students get the highest scores on a test. 31. My Science teacher makes it obvious when certain students are not doing well on their math work. 32. My Science teacher tells us how we compare to other students. 33. My Science teacher calls on smart students more than on other students. **Academic Context Scales** GOOD TEACHING (Wilson et al., 1997) 34. [My Science teacher] motivates students to do their best work. 35. *[My Science teacher] [gives a lot of time to] commenting on students' work. 36. *[My Science teacher] makes a [good] effort to understand [problems] students may be having with their work.	**Motivational Context Scales** MASTERY GOAL STRUCTURE (Midgley et al., 2000) 24. My Science teacher thinks mistakes are okay as long as we are learning. 25. My Science teacher wants us to understand our work, not just memorize it. 26. My Science teacher really wants us to enjoy learning new things. 27. My Science teacher recognizes us for trying hard. 28. My Science teacher gives us time to really explore and understand new ideas. PERFORMANCE GOAL STRUCTURE (Anderman & Midgley, 2004) 29. My Science teacher points out those students who get good grades as an example to all of us. 30. My Science teacher lets us know which students get the highest scores on a test. 31. My Science teacher makes it obvious when certain students are not doing well on their math work. 32. My Science teacher tells us how we compare to other students. 33. My Science teacher calls on smart students more than on other students. **Academic Context Scales** GOOD TEACHING (Wilson et al., 1997) 34. The teaching staff of this course motivate students to do their best work. 35. Staff here put a lot of time into commenting on students' work. 36. The staff here make a real effort to understand difficulties students may be having with their work.

37. [My Science teacher] normally gives helpful feedback [about] how you are doing.	37. Teaching staff here normally gives helpful feedback on how you are doing.
38. [My Science teacher] is [very] good at explaining things to us.	38. Our lecturers are extremely good at explaining things to us.
39. *[My Science teacher] works hard to make [science] interesting.	39. Teaching staff here work hard to make subjects interesting.
40. *[My Science teacher] [doesn't really care about] what students have to say.	40. Staff here show no real interest in what students have to say.
41. [My Science class] really tries to get the best out of all students.	41. This course really tries to get the best out of all its students.
APPROPRIATE ASSESSMENT (Wilson et al., 1997)	APPROPRIATE ASSESSMENT (Wilson et al., 1997)
42. To do well in [my Science class], all you really need is a good memory.	42. To do well in this course, all you really need is a good memory.
43. *[My Science teacher] [seems to care more about] what you've memorized than what you've understood.	43. Staff seem more interested in testing what you've memorized than what you've understood.
44. My Science teacher asks us too many questions just about facts.	44. Too many staff ask us just about facts.
45. It would be possible to [succeed in] [my Science class] just by [studying for tests and quizzes the night before].	45. It would be possible to get through this course just by working hard around exam times.
CLEAR GOALS AND STANDARDS (Wilson et al., 1997)	CLEAR GOALS AND STANDARDS (Wilson et al., 1997)
46. *It's always easy [in my Science class] to know [what I need to do to get a good grade].	46. It's always easy here to know the standard of work expected.
47. *[In my Science class], you usually have a clear idea of [what you're supposed to do].	47. You usually have a clear idea of where you're going and what's expected of you.
48. *It's often hard to [figure out] what [the teacher expects] of you in [my Science] class.	48. It's often hard to discover what's expected of you in this course.
49. *The [goals and purposes] of [my Science class] are NOT made very clear.	49. The aims and objectives of this course are NOT made very clear.

50. [My Science teacher] makes it clear right from the start what [she/he] expects from students.	50. The staff here make it clear right from the start what they expect from students.
APPROPRIATE WORKLOAD (Wilson et al., 1997)	APPROPRIATE WORKLOAD (Wilson et al., 1997)
51. *The [amount of work] in my Science class is too [large].	51. The workload is too heavy.
52. *It seems to me that [my Science teacher] tries to cover too [much material].	52. It seems to me that the syllabus tries to cover too many subjects.
53. *In [my Science class], we are [usually] given enough time to understand the things we have to learn.	53. We are generally given enough time to understand the things we have to learn.
54. There's a lot of pressure on you as a student [in my Science class].	54. There's a lot of pressure on you as a student here.
55. *The [large amount] of work [you have to do] in [my Science class] means you can't understand it all completely.	55. The sheer volume of work to be got through in this course means you can't comprehend it all thoroughly.
USEFULNESS OF CURRICULUM (Rowe & Hill, 1998)	USEFULNESS OF CURRICULUM (Rowe & Hill, 1998)
56. In my [Science] class I learn things that will be useful to me when I leave school.	56. In my class I learn things that will be useful to me when I leave school.
57. *What I learn in [my Science] class will be useful to me [later on as a student].	57. What I learn in class will be useful to me when I go to secondary school.
58. What I learn in [my Science] class is useful to me.	58. What I learn in class is useful to me.
59. What I learn in [my Science] class will be useful to me in the future.	59. What I learn in class will be useful to me in the future.
TRANSPARENCY OF ASSESS. (Dorman & Knightley, 2006)	TRANSPARENCY OF ASSESS. (Dorman & Knightley, 2006)
60. *I understand what is needed in all [graded assignments in my Science class].	60. I understand what is needed in all Science assessment tasks.
61. *I know what is needed to successfully accomplish [graded assignments in my Science class].	61. I know what is needed to successfully accomplish Science assessment tasks.
62. *I know in advance HOW I will be [graded in my Science class].	62. I know in advance HOW I will be assessed.

63. *I am told in advance WHY I am being [ask to do graded assignments in my Science class].	63. I am told in advance WHY I am being assessed.
64. *I am told in advance WHEN I [will be graded in my Science class].	64. I am told in advance WHEN I am being assessed.
65. *I am told in advance WHAT science topics and information I [will be graded] on [in my Science class].	65. I am told in advance WHAT science topics and information I am being assessed on.
66. *I understand the purpose of [graded assignments in my Science class].	66. I understand the purpose of Science assessment.
AUTHENTICITY OF ASSESS. (Dorman & Knightley, 2006)	AUTHENTICITY OF ASSESS. (Dorman & Knightley, 2006)
67. I am asked to apply my learning to real-life situations in my Science class.	67. I am asked to apply my learning to real-life situations in Science class.
68. *[In my Science class], [graded assignments] are meaningful.	68. My Science assessment tasks are meaningful.
69. *[In my Science class], [graded assignments] are useful.	69. My Science assessment tasks are useful.
70. *I find [that in my Science class] [graded assignments] [relate] to the real world [in important ways].	70. I find science assessment tasks relevant to the real world.
71. *[In my Science class], [graded assignments] check my understanding of topics.	71. Science assessment tasks check my understanding of topics.
72. * In my Science class, [graded assignments] test my ability to [use what I've learned].	72. Assessment in Science class tests my ability to apply learning.
73. In my Science class, [Graded assignments] examine my ability to answer important questions.	73. Assessment in Science class examines my ability to answer important questions.
EXPERIENCE OF SCHOOL RULES (Gregory et al., 1989)	EXPERIENCE OF SCHOOL RULES (Gregory et al., 1989)
74. *Everyone knows the rules for [how students should behave in my Science class].	74. Everyone knows the school rules for student conduct.
75. *[The rules in my Science class] are fair.	75. The school rules are fair.

76. *The punishment for breaking [rules in my Science class] is the same no matter who you are.	76. The punishment for breaking school rules is the same no matter who you are.
77. *[My Science teacher makes sure that everyone follows the rules in class.]	77. School rules are strictly enforced.
78. *If a [rule is broken in my Science class], students know [what the teacher will do about it].	78. If a school rule is broken, students know what kind of punishment will follow.
79. *If a student breaks the rules [in my Science class, the teacher will do something about it].	79. If a student breaks the rules in this school, he or she will be punished.
PEER NORMS (Mayhew et al., 2009)	PEER NORMS (Mayhew et al., 2009)
80. *If I cheated on a test [in Science class this year, most of my classmates] would [think that's okay].	80. If I cheated on an in-class test, most people who are important to me (e.g., my family, friends, etc.) would approve of my behavior.
81. *[Most of my classmates in Science class this year] would be willing to cheat on [a Science test to avoid failing].	81. The people in my life whose opinions I value (e.g., my family, friends, etc.) would be willing to cheat on an in- class test or exam if they were in my situation.
82. *[Most of my classmates] would NOT [think it's okay] if I cheated [in Science class this year].	82. The people in my life whose opinions I value (e.g., my family, friends, etc.) would NOT approve if I cheated on an in-class test.
83. *Most [of my classmates] think that I should NOT cheat [in Science class this year].	83. Most people who are important to me (e.g., my family, friends, colleagues, teachers, etc.) think I should NOT cheat on an in-class test or exam.
84. *[My classmates] will look down on me if I cheat [in Science class this year].	84. Most people who are important to me (e.g., my family, friends, colleagues, teachers, etc.) will look down on me if I cheat on an in-class test or exam.
85. *[Most of my classmates] expect me to cheat [in Science class this year.]	85. People whose opinions I value (e.g., my family, friends, etc.) expect me to cheat on an in-class test or exam.
86. *[None of my classmates think it is okay to cheat in Science class this year.]	86. NO ONE who is important to me (e.g., my family, friends, etc.) thinks it is OK to cheat on an in-class test or exam.

Learning strategies DEEP LEARNING STRATEGIES (Anderman et al., 1998) 87. When working on a Science problem, I try to see how it connects with something in everyday life. 88. When I make mistakes in Science, I try to figure out why. 89. I try to connect new work in Science to what I've learned before. 90. I take my time to figure out my work in Science. 91. If I can't solve a Science problem one way, I try to use a different way. 92. *I spend some time thinking about how to do my Science [work] before I start it. 93. I ask myself questions when I work on Science to make sure I understand. SURFACE LEARNING STRATEGIES (Simon et al., 2004) 94. *I study [for Science class] by rehearsing (repeating over and over) [important information]. 95. *I study [for Science class] by memorizing [things] I do not understand. 96. *I study [for Science class] by rehearsing [and repeating] the material over and over until I can [write it exactly, word-for-word]. 97. I study [for Science class] by skipping over parts I think the teacher will not ask questions about. 98. I study [for Science class] by skipping parts I do not understand. 99. I study [for Science class] by skipping parts I do not find important.	**Learning strategies** DEEP LEARNING STRATEGIES (Anderman et al., 1998) 87. When working on a Science problem, I try to see how it connects with something in everyday life. 88. When I make mistakes in Science, I try to figure out why. 89. I try to connect new work in Science to what I've learned before. 90. I take my time to figure out my work in Science. 91. If I can't solve a Science problem one way, I try to use a different way. 92. I spend some time thinking about how to do my Science before I start it. 93. I ask myself questions when I work on Science to make sure I understand. SURFACE LEARNING STRATEGIES (Simon et al., 2004) 94. I study, or will study, by rehearsing (repeating over and over) the material different times. 95. I study, or will study, by memorizing something I do not understand. 96. I study, or will study, by rehearsing material until I can reproduce it literally. 97. I study, or will study, course material by skipping over parts I think the teacher will not ask questions about. 98. I study, or sill study, by skipping parts I do not understand. 99. I study, or will study, by skipping parts I do not find important.

Academic integrity	**Academic integrity**
JUSTIFIABILITY OF CHEATING (Murdock et al., 2004)	JUSTIFIABILITY OF CHEATING (Various)
100. Students would have a good reason to cheat on a test in [my Science class].	100. Students would have a good reason to cheat on a test in Ms. Jones's class. (Murdock et al., 2004)
101. Students would be justified to cheat on an exam in [my Science class].	101. Students would be justified to cheat on an exam in Dr. James' class. (Murdock et al., 2004)
102. *[It's reasonable] to cheat in [my] Science class.	102. Is it okay to cheat in Science class? (Anderman et al., 1998)
103. *[I can understand why students would cheat in my Science class.]	103. (Developed based on Murdock et al., 2004; Murdock et al., 2007)
ACTUAL CHEATING (Midgley et al., 2000)	ACTUAL CHEATING (Midgley et al., 2000)
104. *I sometimes [cheat on Science] tests, [this year].	104. I sometimes copy answers from other students during tests.
105. I sometimes cheat on my [Science] class work, [this year].	105. I sometimes cheat on my class work.
106. * I have cheated on Science class work by copying answers from other students [this year].	106. I sometimes copy answers from other students when I do my class work.
107. *I have cheated in Science class this year.	107. I cheat on my Science work (Anderman et al., 1998)

Appendix D: Time 2 questionnaire

UNIVERSITY OF SYDNEY
Faculty of Education and Social Work, NSW 2006
Dr. Paul Ginns, Senior Lecturer (Rm 914, Bld A35)
Ph/Fax: +61 (02) 9351-2611/5027 Email: paul.ginns@sydney.edu.au

STUDENT QUESTIONNAIRE: *ACADEMIC INTEGRITY IN CONTEXT*

Dear Student,

This questionnaire looks at how students' experiences in Science class affect their study strategies, attitudes towards academic cheating, and actual academic cheating in Science class. Science is the class chosen for research in *all* schools that participate in this study. We are very interested in your experiences, and what they say about how school experiences, in general, may be improved. Your answers will be combined with the answers from many other students to get an overall picture of how study strategies and academic integrity are affected by class experiences.

Your answers are confidential. Please do not write your name anywhere on the questionnaire form. Your individual answers will never be reported to your parents, your school, or anyone else. Since the questionnaire is *anonymous*, however, it cannot be withdrawn after you hand it in (because we will no longer know who filled it out). All questionnaires will be stored in a secure location at the University of Sydney in Australia. Only aggregated (group) scores from the overall study will be submitted for publication; a PhD thesis will also be produced. In this way, your individual answers will be hidden (as an anonymous part of a large group).

This questionnaire should take about 30 minutes to complete. It should be given only to students who participated in this project, last year. If you have any questions after reading this information, please contact Bradford Barnhardt at the University of Sydney on +61 2 9351 6260 or bbar6232@uni.sydney.edu.

Thank you very much for your time,

Dr. Paul Ginns (Chief Investigator, University of Sydney)
Bradford Barnhardt (PhD Student, University of Sydney)

INSTEAD OF WRITING YOUR NAME, please create an IDENTIFICATION CODE, below

Your answers on this questionnaire are confidential. Your individual answers will not be reported to your parents, your school or anyone else, ever.

IDENTIFICATION CODE			
LAST 2 letters of your FAMILY NAME	**LAST** two letters of your FIRST NAME	MONTH of birth (as a number)	**LAST** 2 numbers of your MOBILE PHONE (If you do not have your own mobile phone, put 00)

SECTION A: BACKGROUND INFORMATION

1. Your age:	__________	2. Your grade level:	__________

3. Please indicate your gender:	☐ Male	☐ Female

4. How do you rate your English skills?				
☐ Very good	☐ Good	☐ Average	☐ Poor	☐ Very poor

5. What language is spoken most by YOUR FAMILY at home? (please choose one, only)						
	1. ☐	English	5. ☐	Hindi	9. ☐	Japanese
	2. ☐	Spanish	6. ☐	Bengali	10. ☐	German
	3. ☐	Arabic	7. ☐	Portuguese	11. ☐	Filipino/Tagalog
	4. ☐	Mandarin	8. ☐	Russian	12. ☐	French
	13. ☐	Other	If 'Other', which language?			

6. What is your parent's/guardian's level of education?	**Female parent or guardian**	**Male parent or guardian**
Less than high school diploma or certificate	1. ☐	1. ☐
High school diploma or certificate	2. ☐	2. ☐
Trade or apprenticeship	3. ☐	3. ☐
University or college degree	4. ☐	4. ☐
Advanced degree (eg. Masters, PhD, medical, business, law, etc.)	5. ☐	5. ☐

SECTION B: ABOUT ME

HOW DO YOU SEE YOURSELF?

In this section, we ask how you see yourself in terms of Science and Honesty – not just this year, but overall. Remember, your answers are anonymous and confidential. Thank you!

For each of the items below, please mark a number to indicate your level of agreement.

1—agree strongly
2—agree
3—neutral
4—disagree
5—disagree strongly

		Agree Strongly				**Disagree Strongly**
1	People can really count on me to do the right thing.	**1**	**2**	**3**	**4**	**5**
2	I learn things quickly in Science class.	**1**	**2**	**3**	**4**	**5**
3	I get good grades in Science.	**1**	**2**	**3**	**4**	**5**
4	Science is one of my best subjects.	**1**	**2**	**3**	**4**	**5**
5	I always tell the truth.	**1**	**2**	**3**	**4**	**5**
6	When I make a promise I keep it.	**1**	**2**	**3**	**4**	**5**

* Items that were added to the questionnaire at Time 2

		Agree Strongly				Disagree Strongly
7	I am honest.	1	2	3	4	5
8	Honesty is very important to me.	1	2	3	4	5
9	I often tell lies.	1	2	3	4	5
10	Work in Science class is easy for me.	1	2	3	4	5
11	I am hopeless in Science class.	1	2	3	4	5

SECTION C: ABOUT MY SCIENCE CLASS

HOW ABOUT YOUR SCIENCE CLASS, THIS YEAR??

The last section asked how you see yourself as a Science student, overall. In this section, we ask about your experience in Science class, THIS YEAR. What's good and bad about it??

We chose Science class as the context for our questions in this section because we had to focus on just one subject. All students who do this survey are asked the same questions.

Will this make your teacher look good or bad? No. The answers you and your classmates give in this section will never be reported to your school or to your science teacher. So the way you answer cannot make your science teacher look good or bad.

Please just let us know what is true. Thank you!

For each of the items below, please mark a number to indicate your level of agreement.
1—agree strongly
2—agree
3—neutral
4—disagree
5—disagree strongly

		Agree Strongly				Disagree Strongly
*12	Science class this year makes me feel capable.	1	2	3	4	5
*13	Science class this year makes me feel that I belong and my classmates care about me.	1	2	3	4	5
14	My Science teacher gives a lot of time to commenting on students' work.	1	2	3	4	5
15	What I learn in my Science class is useful to me.	1	2	3	4	5
16	The rules in my Science class are fair.	1	2	3	4	5

* Items that were added to the questionnaire at Time 2

		Agree Strongly				Disagree Strongly
17	There's a lot of pressure on you as a student in my Science class.	1	2	3	4	5
*18	Science class this year makes me feel like a good student.	1	2	3	4	5
19	My Science teacher makes sure that everyone follows the rules in class.	1	2	3	4	5
20	Most of my classmates think that I should NOT cheat in Science class.	1	2	3	4	5
*21	The work assigned for Science class this year takes too much time.	1	2	3	4	5
22	If a student breaks the rules in my Science class, the teacher will do something about it.	1	2	3	4	5
23	I understand the purpose of graded assignments in my Science class.	1	2	3	4	5
24	The large amount of work you have to do in my Science class means you can't understand it all completely.	1	2	3	4	5
25	My classmates will look down on me if I cheat in Science class this year.	1	2	3	4	5
26	I understand what is needed in all graded assignments in my Science class.	1	2	3	4	5
27	My Science teacher is very good at explaining things to us.	1	2	3	4	5
*28	Science class this year makes me feel free.	1	2	3	4	5
29	In my Science class, we are usually given enough time to understand the things we have to learn.	1	2	3	4	5
*30	Science class this year makes me feel involved with close friends.	1	2	3	4	5
*31	Science class this year makes me feel pressured.	1	2	3	4	5
*32	The amount of work assigned for Science class this year is not reasonable.	1	2	3	4	5
33	My Science teacher normally gives helpful feedback about how you are doing.	1	2	3	4	5

34	None of my classmates think it is okay to cheat in my Science class this year.	**1**	**2**	**3**	**4**	**5**
35	If a rule is broken in my Science class, students know what the teacher will do about it.	**1**	**2**	**3**	**4**	**5**
36	In my Science class, graded assignments check my understanding of topics.	**1**	**2**	**3**	**4**	**5**
*37	Science class this year makes me feel competent.	**1**	**2**	**3**	**4**	**5**

* Items that were added to the questionnaire at Time 2

		Agree Strongly				**Disagree Strongly**
*38	Science class this year makes me feel that I'm doing what I want to be doing.	**1**	**2**	**3**	**4**	**5**
*39	Science class this year makes me feel like I am able to do well at Science.	**1**	**2**	**3**	**4**	**5**
*40	The workload in Science class this year requires too much effort.	**1**	**2**	**3**	**4**	**5**
41	My Science teacher makes a good effort to understand problems students may be having with their work.	**1**	**2**	**3**	**4**	**5**
42	The amount of work in my Science class is too large.	**1**	**2**	**3**	**4**	**5**
43	In my Science class I learn things that will be useful to me when I leave school.	**1**	**2**	**3**	**4**	**5**
44	Everyone knows the rules for how students should behave in my Science class.	**1**	**2**	**3**	**4**	**5**
*45	Science class this year makes me feel emotionally close to my classmates.	**1**	**2**	**3**	**4**	**5**
46	What I learn in my Science class will be useful to me later on as a student.	**1**	**2**	**3**	**4**	**5**
*47	Science class this year makes me feel that my Science skills are improving.	**1**	**2**	**3**	**4**	**5**
48	If I cheated on a test in Science class this year, most of my classmates would think that's okay.	**1**	**2**	**3**	**4**	**5**
49	The punishment for breaking rules in my Science class is the same no matter who you are.	**1**	**2**	**3**	**4**	**5**

50	In my Science class, graded assignments are useful.	**1**	**2**	**3**	**4**	**5**
51	My Science teacher tells us how we compare to other students.	**1**	**2**	**3**	**4**	**5**
52	My Science teacher motivates students to do their best work.	**1**	**2**	**3**	**4**	**5**
*53	Science class this year makes me feel free to decide for myself what to do.	**1**	**2**	**3**	**4**	**5**
54	What I learn in my Science class will be useful to me in the future.	**1**	**2**	**3**	**4**	**5**
55	Most of my classmates would NOT think it's okay if I cheated in Science class this year.	**1**	**2**	**3**	**4**	**5**
56	I know in advance HOW I will be graded in my Science class.	**1**	**2**	**3**	**4**	**5**
57	My Science teacher works hard to make science interesting.	**1**	**2**	**3**	**4**	**5**
58	My Science teacher doesn't really care about what students have to say.	**1**	**2**	**3**	**4**	**5**

* Items that were added to the questionnaire at Time 2

		Agre Strongly				**Disagree Strongly**
59	My Science teacher points out those students who get good grades as an example to all of us.	**1**	**2**	**3**	**4**	**5**
*60	Science class this year makes me feel free to work in my own way.	**1**	**2**	**3**	**4**	**5**
61	In my Science class, graded assignments test my ability to use what I've learned.	**1**	**2**	**3**	**4**	**5**
*62	Science class this year makes me feel like I'm good at Science.	**1**	**2**	**3**	**4**	**5**
63	My Science teacher makes it obvious when certain students are not doing well on their Science work.	**1**	**2**	**3**	**4**	**5**
64	My Science teacher lets us know which students get the highest scores on a test.	**1**	**2**	**3**	**4**	**5**
65	My Science class really tries to get the best out of all students.	**1**	**2**	**3**	**4**	**5**
66	In my Science class, graded assignments examine my ability to answer important questions.	**1**	**2**	**3**	**4**	**5**

* Items that were added to the questionnaire at Time 2

SECTION D: HOW I LEARN AND ACHIEVE IN SCIENCE CLASS

This last short section lets us know how you learn and achieve grades in Science class, THIS YEAR.

Your answers will help us understand how your experiences relate to your behavior in Science class.

We hope this information will help us understand how to improve student experiences, overall.

Remember, your answers are anonymous and confidential. Thank you!

For each of the items below, please mark a number to indicate your level of agreement.

1—agree strongly
2—agree
3—neutral
4—disagree
5—disagree strongly

		Agree Strongly				**Disagree Strongly**
67	It's reasonable to cheat in my Science class.	**1**	**2**	**3**	**4**	**5**
*68	I try to do the smallest amount of work possible for Science class.	**1**	**2**	**3**	**4**	**5**
69	I ask myself questions when I work on Science to make sure I understand.	**1**	**2**	**3**	**4**	**5**
*70	Cheating on Science work this year would make-up for some things that were wrong with the class.	**1**	**2**	**3**	**4**	**5**
71	I try to connect new work in Science to what I've learned before.	**1**	**2**	**3**	**4**	**5**
72	I sometimes cheat on Science tests, this year.	**1**	**2**	**3**	**4**	**5**
73	When working on a Science problem, I try to see how it connects with something in everyday life.	**1**	**2**	**3**	**4**	**5**
74	Students would have a good reason to cheat on a test in my Science class.	**1**	**2**	**3**	**4**	**5**
75	I study for Science class by skipping parts I do not understand.	**1**	**2**	**3**	**4**	**5**
76	I study for Science class by memorizing things I do not understand.	**1**	**2**	**3**	**4**	**5**
77	I have cheated in Science class this year.	**1**	**2**	**3**	**4**	**5**
*78	I try to complete Science assignments with the smallest effort possible.	**1**	**2**	**3**	**4**	**5**

79	I study for Science class by skipping over parts I think the teacher will not ask questions about.	**1**	**2**	**3**	**4**	**5**

* Items that were added to the questionnaire at Time 2

		Agree Strongly				**Disagree Strongly**
80	I sometimes cheat on my Science class work, this year.	**1**	**2**	**3**	**4**	**5**
81	I take my time to figure out my work in Science.	**1**	**2**	**3**	**4**	**5**
*82	Cheating on a Science exam this year would balance-out some things that were unfair about the class this year.	**1**	**2**	**3**	**4**	**5**
83	When I make mistakes in Science, I try to figure out why.	**1**	**2**	**3**	**4**	**5**
84	I have cheated on Science class work by copying answers from other students this year.	**1**	**2**	**3**	**4**	**5**
85	If I can't solve a Science problem one way, I try to use a different way.	**1**	**2**	**3**	**4**	**5**
*86	I try to finish work for Science class as quickly as possible, even if it means I don't learn very much.	**1**	**2**	**3**	**4**	**5**
87	I study for Science class by skipping parts I do not find important.	**1**	**2**	**3**	**4**	**5**
*88	Cheating on Science assignments this year is fair.	**1**	**2**	**3**	**4**	**5**
89	Students would be justified to cheat on an exam in my Science class.	**1**	**2**	**3**	**4**	**5**

* Items that were added to the questionnaire at Time 2

Appendix E:

Construct acronyms, definitions, and valences

Variable	Acronym	Definition and (valence)
Person (Perceptions of self)		
Subject self-concept	'SUB'	A student's assessment of his/her own strength/ability in a given subject area. (LOWER score = more positive self-concept)
Honesty-trustworthiness self-concept	'HON'	A student's assessment of his/her tendency to be honest/trustworthy. (LOWER score = more positive self-concept)
Context (Perceptions of Science Class)		
Performance goal structure	'PERF'	Perception of the degree to which the Science teacher emphasizes 'performance goals'. E.g. competitive and approval-seeking goals, as contrasted with 'mastery goals'. (LOWER score = more performance goal structure in Science class)
Usefulness of curriculum	'CURUSE'	Perception of the degree to which the curriculum in Science class is useful (LOWER score = curriculum perceived as more useful)
Teacher quality (See 'TEACHER', below)	'GTEACH'	Perception of the pedagogical quality along such lines as feedback, clarity, and supportiveness. (LOWER score = better perception of teacher quality)
Assessment quality (See 'TEACHER', below)	'ASSESS'	Perception of the quality of assessment in terms of authenticity and transparency. (LOWER score = better perception of assessment quality)

TEACHER QUALITY (This 'second-order factor' comprises two first order factors found to be multicollinear: assessment (ASSESS), and good teaching (GTEACH))	'TEACHER'	=ASSESS + GTEACH. This higher-order factor represents a global measure students' experience of a particular teacher. TEACHER combines student perceptions of the quality of assessment, and pedagogical skill in Science class. Both first-order measures appear, statistically, to reflect a single underlying source of variance: the teacher. (LOWER score = better perception teacher)
Peer cheating norms	'PEER'	The degree to which peers in Science class are perceived to condone cheating (LOWER score = perception that cheating is more accepted by a respondent's peers)
Moral obligation:		
Justifiability of cheating	'CHJUST'	The degree to which cheating is viewed as justifiable, for oneself or others, within the specific context of Science class (LOWER score = more justifiable)
Behavior:		
Surface learning strategies	'SURF'	The use of effort-minimizing/corner-cutting strategies such as memorizing, over-reliance on formulae, and generally targeting the production of answers purely to fulfill work requirements. (LOWER score = more use of surface strategies)
Self-Reported cheating	'CHEAT'	The frequency or degree of a student's cheating behaviour during the year, in Science class (LOWER score = more cheating)

Appendix F:

Item descriptive statistics: Time 1 *vs*. Pilot Study

Table F1

Item descriptive statistics: Pilot *vs*. Time 1

	Time 1 sample (*n*=493)						Pilot sample (*n*=96)						Δ = (Time 1) - (Pilot)				
	mean	SE	SD	var.	S	K	mean	SE	SD	var.	S	K	Δmean	ΔSD	Δvar.	ΔS	ΔK
SUB																	
sub2	2.43	0.05	1.04	1.08	0.40	-0.41	2.33	0.10	1.01	1.02	0.59	0.25	0.10	0.03	0.06	-0.19	-0.66
sub3	2.31	0.05	1.01	1.01	0.52	-0.13	2.32	0.10	1.00	1.00	0.47	0.00	-0.01	0.01	0.01	0.05	-0.13
sub5	2.85	0.06	1.23	1.52	0.08	-0.90	2.65	0.12	1.22	1.49	0.39	-0.76	0.20	0.01	0.03	-0.31	-0.14
sub13	2.72	0.05	1.10	1.20	0.17	-0.56	2.58	0.10	1.02	1.05	0.31	-0.25	0.14	0.07	0.16	-0.14	-0.30
sub15	1.88	0.05	1.03	1.06	1.00	0.20	1.79	0.10	0.98	0.97	1.25	1.17	0.09	0.05	0.10	-0.25	-0.97
HON																	
hon 1	1.86	0.03	0.72	0.51	0.62	0.60	2.20	0.08	0.82	0.67	0.69	0.86	-0.34	-0.10	-0.15	-0.07	-0.26
hon 6	2.46	0.04	0.89	0.80	0.39	0.16	2.58	0.09	0.93	0.86	0.36	0.13	-0.12	-0.03	-0.06	0.03	0.04
hon 8	1.90	0.04	0.79	0.63	0.60	0.02	1.92	0.09	0.87	0.75	1.15	2.08	-0.02	-0.07	-0.12	-0.55	-2.06
hon 9	1.98	0.03	0.78	0.60	0.77	1.19	2.10	0.09	0.88	0.77	0.94	1.40	-0.13	-0.10	-0.16	-0.17	-0.21
hon10	1.77	0.04	0.83	0.69	1.01	0.78	1.79	0.09	0.92	0.84	1.35	2.04	-0.02	-0.09	-0.15	-0.34	-1.26
hon11	2.17	0.04	0.98	0.96	0.65	-0.14	2.31	0.11	1.04	1.08	0.72	-0.05	-0.14	-0.06	-0.12	-0.07	-0.10

Note. SE = Standard error; SD = Standard deviation; Var. = variance; S = Skewness; K = Kurtosis; Δ = 'change in'. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level.

Table F1, continued

	Time 1 sample (*n*=493)						Pilot sample (*n*=96)						Δ = (Time 1) - (Pilot)				
	mean	SE	SD	var.	S	K	mean	SE	SD	var.	S	K	Δmean	ΔSD	Δvar.	ΔS	ΔK
GTEACH																	
gteach18	3.12	0.05	1.06	1.11	-0.03	-0.57	3.13	0.10	1.00	0.99	0.00	-0.48	-0.01	0.06	0.12	-0.04	-0.09
gteach33	2.42	0.05	1.18	1.40	0.67	-0.36	2.41	0.13	1.29	1.68	0.63	-0.70	0.02	-0.11	-0.28	0.05	0.34
gteach39	2.52	0.05	1.09	1.19	0.46	-0.45	2.44	0.11	1.06	1.13	0.54	-0.26	0.08	0.03	0.06	-0.08	-0.19
gteach50	2.52	0.05	1.07	1.15	0.58	-0.18	2.48	0.10	1.01	1.01	0.50	0.04	0.04	0.07	0.14	0.08	-0.22
gteach62	2.39	0.05	1.06	1.12	0.68	0.09	2.31	0.12	1.21	1.46	0.65	-0.49	0.08	-0.15	-0.34	0.03	0.57
gteach67	2.40	0.05	1.08	1.16	0.61	-0.11	2.41	0.13	1.31	1.72	0.58	-0.78	-0.01	-0.23	-0.56	0.03	0.67
gteach68	2.26	0.05	1.10	1.21	0.66	-0.18	2.20	0.12	1.18	1.40	0.77	-0.30	0.06	-0.08	-0.19	-0.11	0.11
gteach77	2.48	0.04	1.00	0.99	0.55	0.10	2.44	0.12	1.18	1.39	0.47	-0.42	0.04	-0.18	-0.39	0.08	0.52
PERF																	
perf61	3.49	0.05	1.16	1.36	-0.36	-0.75	3.36	0.11	1.12	1.24	0.07	-0.85	0.13	0.05	0.11	-0.43	0.11
perf69	3.24	0.06	1.24	1.54	-0.15	-0.99	2.89	0.13	1.26	1.58	0.06	-0.98	0.36	-0.02	-0.04	-0.21	-0.01
perf74	3.17	0.06	1.25	1.56	-0.09	-0.98	3.09	0.11	1.10	1.20	-0.04	-0.65	0.07	0.15	0.36	-0.05	-0.33
perf75	3.61	0.06	1.34	1.79	-0.54	-0.94	3.22	0.13	1.27	1.60	-0.23	-0.87	0.39	0.07	0.19	-0.31	-0.07
CURUSE																	
curuse19	2.53	0.05	1.12	1.26	0.50	-0.43	2.41	0.12	1.13	1.28	0.66	-0.02	0.12	-0.01	-0.01	-0.17	-0.41
curuse53	2.61	0.05	1.11	1.23	0.50	-0.36	2.48	0.11	1.12	1.26	0.55	-0.15	0.13	-0.01	-0.03	-0.05	-0.22
curuse56	2.31	0.05	1.02	1.04	0.58	-0.14	2.27	0.11	1.08	1.17	0.56	-0.16	0.04	-0.06	-0.12	0.02	0.03
curuse64	2.48	0.05	1.08	1.16	0.51	-0.25	2.46	0.12	1.22	1.49	0.61	-0.42	0.02	-0.15	-0.34	-0.10	0.16

Table F1, continued

	Time 1 sample (*n*=493)						Pilot sample (*n*=96)						Δ = (Time 1) - (Pilot)				
	mean	SE	SD	var.	S	K	mean	SE	SD	var.	S	K	Δmean	ΔSD	Δvar.	ΔS	ΔK
ASSESS																	
auth44	2.06	0.04	0.86	0.73	0.73	0.68	2.11	0.10	0.94	0.88	0.70	0.45	-0.05	-0.08	-0.15	0.03	0.23
auth60	2.30	0.04	0.94	0.89	0.50	0.13	2.24	0.10	0.98	0.96	0.80	0.80	0.06	-0.04	-0.07	-0.29	-0.67
auth71	2.31	0.04	0.90	0.80	0.40	0.07	2.41	0.09	0.90	0.81	0.42	0.23	-0.10	0.00	-0.01	-0.02	-0.16
auth78	2.15	0.04	0.94	0.89	0.68	0.36	2.19	0.10	0.95	0.91	0.35	-0.80	-0.04	-0.01	-0.03	0.32	1.16
trans28	2.04	0.04	0.90	0.81	0.93	1.01	1.95	0.09	0.91	0.83	1.04	1.36	0.09	-0.01	-0.02	-0.11	-0.35
trans32	2.18	0.04	0.86	0.73	0.57	0.21	2.23	0.08	0.80	0.64	0.31	-0.24	-0.05	0.05	0.09	0.26	0.46
trans66	2.29	0.04	0.95	0.91	0.71	0.45	2.41	0.11	1.05	1.11	0.34	-0.43	-0.11	-0.10	-0.20	0.37	0.88
PEER																	
peer24	2.02	0.04	0.99	0.98	1.01	0.78	3.63	0.12	1.20	1.44	-0.43	-0.75	-1.61	-0.21	-0.46	1.44	1.53
peer31	2.12	0.04	0.96	0.92	0.70	0.10	3.20	0.14	1.40	1.95	-0.15	-1.20	-1.08	-0.44	-1.03	0.85	1.30
peer40	2.10	0.05	1.04	1.09	0.85	0.23	3.28	0.14	1.37	1.89	-0.18	-1.19	-1.19	-0.33	-0.80	1.02	1.43
peer58	2.26	0.05	1.16	1.33	0.72	-0.30	3.67	0.13	1.24	1.53	-0.49	-0.85	-1.41	-0.08	-0.19	1.21	0.55
peer65	3.79	0.05	1.15	1.32	-0.75	-0.15	3.63	0.14	1.37	1.88	-0.57	-0.91	0.17	-0.22	-0.56	-0.19	0.76
SURF																	
surf87	3.99	0.05	1.00	1.01	-0.89	0.22	3.81	0.11	1.10	1.21	-0.79	0.02	0.18	0.89	-0.09	-2.10	1.00
surf88	2.94	0.06	1.27	1.62	0.08	-1.05	2.79	0.12	1.13	1.28	0.33	-0.73	0.15	1.16	0.49	-1.20	-1.39
surf91	3.46	0.05	1.20	1.43	-0.31	-0.85	3.56	0.11	1.03	1.07	-0.43	-0.57	-0.11	1.09	0.40	-1.38	-0.42
surf97	3.60	0.05	1.14	1.30	-0.37	-0.78	3.54	0.11	1.04	1.07	-0.32	-0.62	0.06	1.03	0.26	-1.44	-0.46

Table F1, continued

	Time 1 sample (n=493)						Pilot sample (n=96)						Δ = (Time 1) - (Pilot)				
	mean	SE	SD	var.	S	K	mean	SE	SD	var.	S	K	Δmean	ΔSD	Δvar.	ΔS	ΔK
CHJUST																	
chjust79	4.21	0.04	0.94	0.89	-1.06	0.57	4.35	0.10	0.99	0.99	-1.55	1.80	-0.15	0.84	-0.11	-2.05	2.12
chjust86	3.95	0.05	1.17	1.37	-0.95	-0.02	3.97	0.12	1.18	1.40	-0.87	-0.24	-0.02	1.05	0.19	-2.35	0.85
chjust99	3.97	0.05	1.19	1.42	-1.01	0.16	3.77	0.13	1.30	1.69	-0.59	-0.88	0.20	1.06	0.12	-2.71	0.74
CHEAT																	
cheat84	4.41	0.04	0.95	0.90	-1.58	1.63	4.42	0.11	1.04	1.09	-1.76	2.09	-0.01	0.84	-0.14	-2.66	3.40
cheat92	4.28	0.05	1.02	1.05	-1.26	0.51	4.36	0.10	1.02	1.03	-1.46	1.05	-0.08	0.92	0.03	-2.30	1.97
cheat95	4.16	0.05	1.12	1.25	-1.11	0.09	4.29	0.10	1.01	1.03	-1.11	-0.16	-0.13	1.02	0.24	-2.14	1.21

Appendix G:

Time 1 MIMIC model results estimated with all observed indicators

Table G1

MIMIC results: standardized beta coefficients for covariates

	Gen	Gra	Eng	Mom	Dad
Person					
Subject self-concept	.242***	.005	.092	.107*	.116*
Honesty-trust. self-concept	-.059	-.036	.056	.010	.031
Teaching context					
Performance goal structure	.134*	.012	-.115*	.016	-.089
Teacher	.063	.156**	-.043	.000	.030
Usefulness of curriculum	.060	.136**	.026	-.003	-.022
Peer cheating norms	.229***	-.214***	.017	-.014	.026
Moral obligation					
Justifiability of cheating	.183**	-.157**	-.034	.090	-.086
Behavior					
Surface learning strategies	.024	-.149**	-.134*	.047	-.093
Self-reported cheating	.138**	-.064	-.034	.037	-.068

Note. Model fit: $\chi^2(1289) = 2036$; *RMSEA* = .037, *CIs* = .034 - .040, *pclose* = 1.00; *TLI* = .90; *CFI* = .91; *SRMR* = .054; SCF = .953, and the second model, which included all two-way interactions of these variables, ($\chi^2(1489) = 2266$; *RMSEA* = .035, *CIs* = .032 - .038, *pclose* = 1.00; *TLI* = .90; *CFI* = .91; *SRMR* = .051; SCF = .948). Gen = Gender, Gra = Grade-level, Eng = English proficiency, Mom = Maternal educational attainment, Dad = Paternal educational attainment, *p < .05, **p < .01, ***p < .000

Table G2

MIMIC results: standardized beta coefficients for two-way interaction variables

	GenXGra	GenXEng	GenXMom	GenXDad	GraXEng	GraXMom	GraXDad	EngXMom	EngXDad	MomXDad
Person										
Subject self-concept	.105	.146	.111	-.099	-.057	-.058	.214*	.129*	.017	.021
Honesty-trust. self-concept	-.071	-.087	.132	-.049	.133	-.104	.043	-.038	-.024	.007
Teaching context										
Performance goal structure	.116	-.236**	.092	-.129	.093	-.080	.138	.075	.052	-.162*
Teacher	.077	.090	.078	-.101	-.127	-.099	.135	.033	.068	.051
Usefulness of curriculum	.044	.027	.034	-.198	-.011	-.084	.148	.026	.084	.125*
Peer cheating norms	.058	-.118	-.061	.016	.120	.066	.027	.032	-.001	-.049
Moral obligation										
Justifiability of cheating	.054	-.139	.063	.047	.127	-.007	-.065	.035	.012	-.044
Behavior										
Surface learning strategies	-.029	-.176*	-.002	-.256**	.115	.104	.115	-.023	.045	-.062
Self-reported cheating	.052	-.152	-.012	-.097	.166	.049	.050	.019	.001	-.046

Note. Model fir: $\chi^2(1489) = 2266$; *RMSEA* = .035, *CI*s = .032 - .038, *pclose* = 1.00; *TLI* = .90; *CFI* = .91; *SRMR* = .051; SCF = .948. Gen = Gender, Gra = Grade-level, Eng = English proficiency, Mom = Maternal educational attainment, Dad = Paternal educational attainment, *p < .05, **p < .01, ***p < .000.

Appendix H:

Gender-specific congeneric models

Table H1

Congeneric model results for male respondents at Time 1

N = 201	CFA									
				Loading	RMSEA					
Scale (# items)	χ^2	*p*	*df*	range	value	CIs	CFI	TLI	SRMR	Rho
Subject self-concept (5)	2.2	.82	5	.65-.87	.00	.00-.06	1.0	1.02	.01	.90
Honesty-trust. self-concept (6)	12.6	.18	9	.44-.90	.05	.00-.10	.99	.98	.03	.84
Performance structure (4)	.32	.85	2	.52-.87	.00	.00-.08	1.0	1.05	.01	.72
Good teaching (8)	25.0	.20	20	.45-.77	.04	.00-.07	.99	.98	.03	.86
Usefulness of curriculum (4)	2.9	.23	2	.66-.89	.05	.00-.16	1.0	.99	.01	.88
Assessment quality (7)	20.3	.12	14	.62-.76	.05	.00-.09	.98	.97	.04	.86
Peer norms (5)	7.6	.18	5	.63-.73	.05	.00-.12	.99	.97	.02	.81
Surface learning strategies (4)	5.6	.06	2	.35-.90	.09	.00-.19	.97	.90	.03	.73
Justifiability of cheating (3)	.22	.64	1	.63-.84	.00	.00-.15	1.0	1.03	.01	.78
Self-reported cheating (3)	.56	.45	1	.76-.89	.00	.00-.17	1.0	1.01	.01	.88

Table H2

Congeneric model results for female respondents at Time 1

N = 292	CFA									
				Loading	RMSEA					
Scale (# items)	χ^2	*p*	df	range	value	CIs	CFI	TLI	SRMR	Rho
Subject self-concept (5)	1.79	.88	5	.73-.83	.00	.00-.04	1.00	1.01	.01	.90
Honesty-trust. self-concept (6)	6.72	.67	9	.38-.88	.00	.00-.05	1.00	1.01	.02	.80
Performance structure (4)	3.64	.16	2	.53-.84	.05	.00-.14	.99	.97	.02	.74
Good teaching (8)	37.8	.01	20	.37-.75	.06	.03-.08	.97	.96	.03	.86
Usefulness of curriculum (4)	1.29	.53	2	.72-.93	.00	.00-.10	1.00	1.01	.01	.90
Assessment quality (7)	18.4	.19	14	.58-.71	.03	.00-.07	.99	.98	.03	.83
Peer norms (5)	12.8	.03	5	.46-.71	.07	.02-.12	.96	.92	.03	.76
Surface learning strategies (4)	3.75	.15	2	.31-.86	.06	.00-.14	.99	.97	.02	.70
Justifiability of cheating (3)	.670	.41	1	.44-.79	.00	.00-.14	1.00	1.02	.01	.67
Self-reported cheating (3)	1.00	.32	1	.72-.88	.00	.00-.16	1.00	1.00	.04	.85

Appendix I:

Standardized beta coefficients for gender-specific models estimated with weighted composites

Table I1

Male sample model estimated with weighted composites: standardized beta coefficients

	Predictors							
N = 201	Grade	Sub	Hon	Perf	Curuse	Teacher	Peer	Chjust
Sub	-.019							
Hon	-.001	---						
Perf	.038	-.058	---					
Curuse	.074	.515***	---	---				
Teacher	.038	.470***	---	---	---			
Peer	-.253**	---	-.143	-.104	.417	-.691**		
Chjust	-.122	.172	-.212*	.300**	-.093	-.306	.355**	
Surf	.078	-.328**	---	.242*	-.123	.395	---	.685***
Cheat	.068	-.146	-.267**	-.117	.081	.128	.071	.811***

Table I2

Female sample model estimated with weighted composites: standardized beta coefficients

	Predictors							
N = 292	Grade	Sub	Hon	Perf	Curuse	Teacher	Peer	Chjust
Sub	-.045							
Hon	-.017	---						
Perf	.058	-.034	---					
Curuse	.124*	.480***	---	---				
Teacher	.182**	.462***	---	---	---			
Peer	-.089	---	-.148*	.236**	-.089	-.359**		
Chjust	-.012	.038	-.262**	.175*	-.010	-.400**	.401***	
Surf	-.103	-.197*	---	.047	-.076	.542***	---	.542***
Cheat	-.058	-.339***	-.084	.034	-.003	.227	-.134	.834***

Appendix J: Male sample model correlation matrices

Table J1

CFA: Estimated correlation matrix for the Time 1 male sample

	SUB1	HON1	PERF1	GTEACH1	CURUSE1
SUB1	1.000				
HON1	0.192	1.000			
PERF1	-0.095	-0.060	1.000		
GTEACH1	0.396	0.089	-0.144	1.000	
CURUSE1	0.519	0.122	-0.102	0.694	1.000
ASSESS1	0.396	0.090	-0.144	0.725	0.695
PEER1	-0.086	-0.160	0.026	-0.278	-0.181
SURF1	-0.300	-0.287	0.463	-0.209	-0.266
CHJUST1	-0.144	-0.308	0.355	-0.373	-0.367
CHEAT1	-0.220	-0.496	0.181	-0.218	-0.195
TEACHER1	0.465	0.105	-0.170	0.851	0.816

	ASSESS1	PEER1	SURF1	CHJUST1	CHEAT1
ASSESS1	1.000				
PEER1	-0.278	1.000			
SURF1	-0.209	0.226	1.000		
CHJUST1	-0.374	0.492	0.641	1.000	
CHEAT1	-0.218	0.472	0.516	0.783	1.000
TEACHER1	0.852	-0.327	-0.246	-0.439	-0.256

Table J2

Structural model: Estimated correlation matrix for the Time 1 male sample

	SUB1	HON1	PERF1	GTEACH1	CURUSE1
SUB1	1.000				
HON1	0.194	1.000			
PERF1	-0.097	-0.018	1.000		
GTEACH1	0.392	0.077	-0.145	1.000	
CURUSE1	0.521	0.102	-0.103	0.696	1.000
ASSESS1	0.390	0.076	-0.145	0.724	0.694
PEER1	-0.118	-0.161	0.021	-0.277	-0.177
SURF1	-0.308	-0.225	0.457	-0.205	-0.265
CHJUST1	-0.153	-0.301	0.347	-0.377	-0.364
CHEAT1	-0.230	-0.492	0.163	-0.214	-0.186
TEACHER1	0.459	0.090	-0.171	0.852	0.817
GRADE	-0.010	0.014	0.037	0.032	0.065

	ASSESS1	PEER1	SURF1	CHJUST1	CHEAT1
ASSESS1	1.000				
PEER1	-0.276	1.000			
SURF1	-0.204	0.247	1.000		
CHJUST1	-0.376	0.488	0.644	1.000	
CHEAT1	-0.213	0.477	0.507	0.781	1.000
TEACHER1	0.849	-0.326	-0.240	-0.443	-0.251
GRADE	0.032	-0.254	-0.058	-0.217	-0.105

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level.

Appendix K: Female sample model correlation matrix

Table K1

Estimated correlation matrix of the CFA for the Time 1 female sample

	SUB1	HON1	PERF1	GTEACH1	CURUSE1
SUB1	1.000				
HON1	0.216	1.000			
PERF1	-0.051	-0.072	1.000		
GTEACH1	0.360	0.262	-0.063	1.000	
CURUSE1	0.471	0.166	0.065	0.513	1.000
ASSESS1	0.438	0.319	-0.076	0.777	0.624
PEER1	-0.265	-0.249	0.275	-0.407	-0.328
SURF1	-0.345	-0.253	0.215	-0.350	-0.325
CHJUST1	-0.285	-0.446	0.324	-0.467	-0.347
CHEAT1	-0.463	-0.416	0.266	-0.347	-0.281
TEACHER1	0.450	0.328	-0.078	0.799	0.642

	ASSESS1	PEER1	SURF1	CHJUST1	CHEAT1
ASSESS1	1.000				
PEER1	-0.495	1.000			
SURF1	-0.426	0.497	1.000		
CHJUST1	-0.568	0.681	0.524	1.000	
CHEAT1	-0.423	0.475	0.537	0.729	1.000
TEACHER1	0.973	-0.509	-0.438	-0.584	-0.435

Table K2

Estimated correlation matrix of the structural model for the Time 1 female sample

	SUB1	HON1	PERF1	GTEACH1	CURUSE1
SUB1	1.000				
HON1	0.228	1.000			
PERF1	-0.053	-0.013	1.000		
GTEACH1	0.374	0.082	-0.067	1.000	
CURUSE1	0.473	0.105	0.065	0.525	1.000
ASSESS1	0.441	0.096	-0.079	0.776	0.619
PEER1	-0.254	-0.165	0.273	-0.404	-0.325
SURF1	-0.347	-0.196	0.213	-0.343	-0.321
CHJUST1	-0.294	-0.363	0.318	-0.446	-0.343
CHEAT1	-0.469	-0.379	0.255	-0.313	-0.269
TEACHER1	0.461	0.101	-0.082	0.811	0.647
GRADE	-0.043	-0.033	0.050	0.130	0.103

	ASSESS1	PEER1	SURF1	CHJUST1	CHEAT1
ASSESS1	1.000				
PEER1	-0.476	1.000			
SURF1	-0.404	0.406	1.000		
CHJUST1	-0.526	0.689	0.543	1.000	
CHEAT1	-0.369	0.445	0.532	0.726	1.000
TEACHER1	0.957	-0.497	-0.422	-0.550	-0.386
GRADE	0.154	-0.129	-0.152	-0.098	-0.053

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level.

Appendix L:

Co-ed structural model output, Time 1 (*N* = 493)

```
MODEL FIT INFORMATION

Number of Free Parameters                      197

Loglikelihood

          H0 Value                      -30258.349
          H0 Scaling Correction Factor      1.2206
            for MLR
          H1 Value                      -29110.708
          H1 Scaling Correction Factor      1.1391
            for MLR

Information Criteria

          Akaike (AIC)                   60910.699
          Bayesian (BIC)                 61738.199
          Sample-Size Adjusted BIC       61112.920
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                           2039.400*
          Degrees of Freedom                  1175
          P-Value                           0.0000
          Scaling Correction Factor         1.1255
            for MLR

*   The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used
    for chi-square difference testing in the regular way.  MLM, MLR and WLSM
    chi-square difference testing is described on the Mplus website.  MLMV, WLSMV,
    and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                           0.039
          90 Percent C.I.                    0.036  0.041
          Probability RMSEA <= .05           1.000

CFI/TLI

          CFI                                0.909
          TLI                                0.902

Chi-Square Test of Model Fit for the Baseline Model

          Value                          10789.419
          Degrees of Freedom                  1274
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                              0.056
```

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 SUB1     BY
    SUB2               1.000      0.000    999.000    999.000
    SUB3               0.933      0.044     21.135      0.000
    SUB5               1.160      0.051     22.618      0.000
    SUB13              1.042      0.045     23.215      0.000
    SUB15              0.837      0.045     18.638      0.000

 HON1     BY
    HON_1              1.000      0.000    999.000    999.000
    HON6               2.152      0.278      7.755      0.000
    HON8               1.474      0.216      6.838      0.000
    HON9               2.249      0.284      7.928      0.000
    HON10              2.178      0.281      7.739      0.000
    HON11              1.828      0.281      6.497      0.000

 PERF1    BY
    PERF61             1.000      0.000    999.000    999.000
    PERF69             1.483      0.136     10.901      0.000
    PERF74             1.172      0.138      8.508      0.000
    PERF75             1.240      0.154      8.038      0.000

 GTEACH1  BY
    GTEACH18           1.000      0.000    999.000    999.000
    GTEACH33           2.119      0.254      8.355      0.000
    GTEACH39           1.765      0.193      9.145      0.000
    GTEACH50           1.826      0.206      8.845      0.000
    GTEACH62           1.784      0.197      9.047      0.000
    GTEACH67           1.896      0.227      8.346      0.000
    GTEACH68           1.368      0.194      7.065      0.000
    GTEACH77           1.562      0.213      7.328      0.000

 CURUSE1  BY
    CURUSE19           1.000      0.000    999.000    999.000
    CURUSE53           1.205      0.066     18.294      0.000
    CURUSE56           1.005      0.066     15.160      0.000
    CURUSE64           1.190      0.065     18.257      0.000

 ASSESS1  BY
    TRANS28            1.000      0.000    999.000    999.000
    TRANS32            0.922      0.066     14.030      0.000
    TRANS66            1.042      0.106      9.820      0.000
    AUTH44             1.065      0.087     12.224      0.000
    AUTH60             1.196      0.095     12.606      0.000
    AUTH71             0.978      0.085     11.467      0.000
    AUTH78             1.119      0.108     10.378      0.000

 PEER1    BY
    PEER24             1.000      0.000    999.000    999.000
    PEER31             1.061      0.101     10.455      0.000
    PEER40             1.143      0.086     13.345      0.000
    PEER58             1.172      0.128      9.188      0.000
    PEER65             0.863      0.101      8.552      0.000

 SURF1    BY
    SURF87             1.000      0.000    999.000    999.000
    SURF88             0.820      0.139      5.901      0.000
    SURF91             1.647      0.219      7.532      0.000
    SURF97             1.649      0.210      7.863      0.000
```

```
 CHJUST1  BY
    CHJUST79           1.000      0.000    999.000    999.000
    CHJUST86           1.153      0.077     14.896      0.000
    CHJUST99           0.840      0.081     10.424      0.000

 CHEAT1   BY
    CHEAT84            1.000      0.000    999.000    999.000
    CHEAT92            1.124      0.063     17.981      0.000
    CHEAT95            1.066      0.076     14.006      0.000

 TEACHER1 BY
    GTEACH1            1.000      0.000    999.000    999.000
    ASSESS1            1.481      0.226      6.562      0.000

 CHEAT1   ON
    CHJUST1            0.784      0.112      7.014      0.000
    PERF1             -0.042      0.056     -0.746      0.456
    PEER1              0.062      0.067      0.924      0.356
    SUB1              -0.232      0.058     -3.979      0.000
    HON1              -0.422      0.142     -2.984      0.003
    TEACHER1           0.331      0.204      1.620      0.105
    CURUSE1            0.061      0.065      0.943      0.346

 SURF1    ON
    CHJUST1            0.380      0.082      4.657      0.000
    PERF1              0.133      0.058      2.293      0.022
    SUB1              -0.135      0.045     -2.997      0.003
    TEACHER1           0.122      0.170      0.716      0.474
    CURUSE1           -0.050      0.062     -0.808      0.419

 CHJUST1  ON
    SUB1               0.059      0.053      1.118      0.264
    PEER1              0.382      0.080      4.753      0.000
    PERF1              0.267      0.067      3.963      0.000
    TEACHER1          -0.523      0.209     -2.507      0.012
    CURUSE1           -0.059      0.072     -0.810      0.418
    HON1              -0.595      0.191     -3.114      0.002

 PEER1    ON
    PERF1              0.147      0.073      2.011      0.044
    TEACHER1          -0.826      0.219     -3.776      0.000
    CURUSE1            0.044      0.086      0.513      0.608
    HON1              -0.329      0.144     -2.287      0.022

 PERF1    ON
    SUB1              -0.053      0.046     -1.158      0.247

 TEACHER1 ON
    SUB1               0.196      0.031      6.419      0.000

 CURUSE1  ON
    SUB1               0.470      0.050      9.407      0.000

 CHEAT1   ON
    GENDER             0.046      0.070      0.650      0.516
    GRADE              0.022      0.062      0.354      0.723

 SURF1    ON
    GENDER            -0.058      0.054     -1.058      0.290
    GRADE             -0.054      0.050     -1.073      0.283

 CHJUST1  ON
    GENDER             0.145      0.070      2.086      0.037
    GRADE             -0.084      0.067     -1.252      0.211
```

```
PEER1    ON
   GENDER              0.248      0.074      3.343      0.001
   GRADE              -0.226      0.075     -3.037      0.002

 PERF1    ON
   GENDER              0.240      0.078      3.071      0.002
   GRADE               0.051      0.068      0.745      0.456

 TEACHER1 ON
   GENDER             -0.039      0.036     -1.090      0.276
   GRADE               0.087      0.035      2.498      0.012

 CURUSE1  ON
   GENDER             -0.119      0.070     -1.694      0.090
   GRADE               0.168      0.071      2.376      0.017

 HON1     ON
   GENDER             -0.053      0.033     -1.611      0.107
   GRADE              -0.009      0.029     -0.293      0.769

 SUB1     ON
   GENDER              0.483      0.084      5.742      0.000
   GRADE              -0.053      0.083     -0.642      0.521

 SUB1     WITH
   HON1                0.053      0.017      3.228      0.001

 PERF1    WITH
   TEACHER1           -0.024      0.014     -1.673      0.094
   CURUSE1             0.012      0.028      0.427      0.669

 TEACHER1 WITH
   CURUSE1             0.131      0.024      5.565      0.000

 SURF1    WITH
   CHEAT1              0.015      0.017      0.903      0.366

 Intercepts
   SUB2                1.740      0.175      9.937      0.000
   SUB3                1.663      0.163     10.229      0.000
   SUB5                2.044      0.203     10.061      0.000
   SUB13               2.003      0.185     10.816      0.000
   SUB15               1.302      0.147      8.853      0.000
   HON_1               1.954      0.073     26.624      0.000
   HON6                2.671      0.140     19.069      0.000
   HON8                2.040      0.103     19.729      0.000
   HON9                2.194      0.146     15.002      0.000
   HON10               1.980      0.146     13.532      0.000
   HON11               2.348      0.124     18.916      0.000
   PERF61              3.075      0.155     19.892      0.000
   PERF69              2.625      0.211     12.421      0.000
   PERF74              2.676      0.179     14.935      0.000
   PERF75              3.094      0.191     16.170      0.000
   GTEACH18            2.921      0.093     31.282      0.000
   GTEACH33            2.002      0.171     11.678      0.000
   GTEACH39            2.166      0.146     14.883      0.000
   GTEACH50            2.154      0.147     14.637      0.000
   GTEACH62            2.037      0.146     13.942      0.000
   GTEACH67            2.021      0.154     13.118      0.000
   GTEACH68            1.990      0.118     16.794      0.000
   GTEACH77            2.166      0.128     16.858      0.000
   CURUSE19            2.149      0.170     12.604      0.000
   CURUSE53            2.152      0.200     10.741      0.000
   CURUSE56            1.934      0.167     11.559      0.000
   CURUSE64            2.026      0.197     10.304      0.000
   TRANS28             1.746      0.119     14.665      0.000
   TRANS32             1.907      0.113     16.880      0.000
   TRANS66             1.985      0.126     15.722      0.000
```

AUTH44	1.751	0.126	13.929	0.000
AUTH60	1.950	0.142	13.730	0.000
AUTH71	2.022	0.120	16.894	0.000
AUTH78	1.816	0.135	13.470	0.000
PEER24	3.776	0.166	22.802	0.000
PEER31	3.431	0.178	19.313	0.000
PEER40	3.326	0.189	17.623	0.000
PEER58	3.636	0.192	18.928	0.000
PEER65	3.614	0.148	24.364	0.000
SURF87	4.118	0.133	31.016	0.000
SURF88	3.045	0.122	24.912	0.000
SURF91	3.664	0.214	17.120	0.000
SURF97	3.804	0.213	17.902	0.000
CHJUST79	4.007	0.166	24.144	0.000
CHJUST86	3.715	0.191	19.468	0.000
CHJUST99	3.802	0.146	26.064	0.000
CHEAT84	4.194	0.169	24.787	0.000
CHEAT92	4.044	0.186	21.690	0.000
CHEAT95	3.931	0.178	22.130	0.000
Residual Variances				
SUB2	0.301	0.038	7.824	0.000
SUB3	0.329	0.031	10.593	0.000
SUB5	0.473	0.048	9.753	0.000
SUB13	0.351	0.034	10.238	0.000
SUB15	0.517	0.045	11.574	0.000
HON_1	0.421	0.037	11.524	0.000
HON6	0.380	0.032	12.033	0.000
HON8	0.431	0.029	15.103	0.000
HON9	0.148	0.018	8.008	0.000
HON10	0.262	0.025	10.609	0.000
HON11	0.658	0.053	12.343	0.000
PERF61	0.925	0.070	13.290	0.000
PERF69	0.594	0.084	7.074	0.000
PERF74	0.970	0.089	10.937	0.000
PERF75	1.128	0.099	11.452	0.000
GTEACH18	0.932	0.058	16.174	0.000
GTEACH33	0.590	0.049	11.924	0.000
GTEACH39	0.628	0.049	12.809	0.000
GTEACH50	0.552	0.054	10.244	0.000
GTEACH62	0.547	0.052	10.577	0.000
GTEACH67	0.511	0.044	11.543	0.000
GTEACH68	0.871	0.067	12.910	0.000
GTEACH77	0.554	0.059	9.369	0.000
CURUSE19	0.606	0.059	10.283	0.000
CURUSE53	0.280	0.035	8.044	0.000
CURUSE56	0.383	0.047	8.057	0.000
CURUSE64	0.228	0.032	7.180	0.000
TRANS28	0.477	0.043	11.152	0.000
TRANS32	0.451	0.043	10.598	0.000
TRANS66	0.548	0.057	9.566	0.000
AUTH44	0.362	0.033	10.894	0.000
AUTH60	0.419	0.043	9.668	0.000
AUTH71	0.488	0.046	10.564	0.000
AUTH78	0.472	0.048	9.860	0.000
PEER24	0.773	0.092	8.381	0.000
PEER31	0.921	0.083	11.035	0.000
PEER40	0.809	0.084	9.614	0.000
PEER58	0.656	0.090	7.276	0.000
PEER65	1.057	0.117	9.039	0.000
SURF87	0.697	0.071	9.786	0.000
SURF88	1.413	0.080	17.678	0.000
SURF91	0.587	0.073	7.990	0.000
SURF97	0.455	0.082	5.521	0.000
CHJUST79	0.328	0.038	8.528	0.000
CHJUST86	0.624	0.080	7.840	0.000
CHJUST99	1.023	0.111	9.222	0.000
CHEAT84	0.265	0.048	5.549	0.000

```
    CHEAT92            0.242      0.036      6.678      0.000
    CHEAT95            0.525      0.077      6.776      0.000
    SUB1               0.724      0.062     11.671      0.000
    HON1               0.089      0.022      4.006      0.000
    PERF1              0.414      0.074      5.624      0.000
    GTEACH1            0.056      0.017      3.338      0.001
    CURUSE1            0.485      0.053      9.217      0.000
    ASSESS1            0.058      0.020      2.920      0.003
    PEER1              0.403      0.066      6.131      0.000
    SURF1              0.181      0.034      5.298      0.000
    CHJUST1            0.246      0.041      6.028      0.000
    CHEAT1             0.206      0.036      5.689      0.000
    TEACHER1           0.093      0.025      3.777      0.000


STANDARDIZED MODEL RESULTS


STDYX Standardization

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 SUB1     BY
    SUB2               0.849      0.021     40.819      0.000
    SUB3               0.821      0.019     42.835      0.000
    SUB5               0.830      0.020     41.395      0.000
    SUB13              0.841      0.018     46.226      0.000
    SUB15              0.717      0.027     26.869      0.000

 HON1     BY
    HON_1              0.419      0.050      8.417      0.000
    HON6               0.723      0.030     24.354      0.000
    HON8               0.558      0.038     14.721      0.000
    HON9               0.869      0.021     42.338      0.000
    HON10              0.787      0.025     30.940      0.000
    HON11              0.560      0.040     14.064      0.000

 PERF1    BY
    PERF61             0.563      0.044     12.867      0.000
    PERF69             0.783      0.036     21.723      0.000
    PERF74             0.614      0.044     13.813      0.000
    PERF75             0.607      0.044     13.948      0.000

 GTEACH1  BY
    GTEACH18           0.402      0.047      8.635      0.000
    GTEACH33           0.760      0.024     31.179      0.000
    GTEACH39           0.686      0.031     22.210      0.000
    GTEACH50           0.721      0.032     22.742      0.000
    GTEACH62           0.715      0.031     22.935      0.000
    GTEACH67           0.747      0.028     26.975      0.000
    GTEACH68           0.528      0.046     11.519      0.000
    GTEACH77           0.665      0.042     15.897      0.000

 CURUSE1  BY
    CURUSE19           0.720      0.032     22.822      0.000
    CURUSE53           0.879      0.016     53.880      0.000
    CURUSE56           0.796      0.030     26.709      0.000
    CURUSE64           0.896      0.016     55.841      0.000

 ASSESS1  BY
    TRANS28            0.638      0.034     18.912      0.000
    TRANS32            0.618      0.039     15.701      0.000
    TRANS66            0.628      0.042     14.922      0.000
    AUTH44             0.712      0.031     22.752      0.000
    AUTH60             0.727      0.029     24.859      0.000
    AUTH71             0.626      0.038     16.335      0.000
    AUTH78             0.682      0.034     20.301      0.000
```

```
PEER1    BY
   PEER24             0.640      0.046     13.932      0.000
   PEER31             0.629      0.042     15.136      0.000
   PEER40             0.681      0.040     17.201      0.000
   PEER58             0.727      0.042     17.133      0.000
   PEER65             0.523      0.056      9.354      0.000

SURF1    BY
   SURF87             0.553      0.053     10.419      0.000
   SURF88             0.357      0.051      7.006      0.000
   SURF91             0.766      0.034     22.337      0.000
   SURF97             0.805      0.039     20.602      0.000

CHJUST1  BY
   CHJUST79           0.790      0.029     27.601      0.000
   CHJUST86           0.733      0.035     20.899      0.000
   CHJUST99           0.522      0.050     10.545      0.000

CHEAT1   BY
   CHEAT84            0.838      0.031     26.773      0.000
   CHEAT92            0.875      0.020     43.971      0.000
   CHEAT95            0.759      0.036     20.854      0.000

TEACHER1 BY
   GTEACH1            0.829      0.032     26.006      0.000
   ASSESS1            0.908      0.033     27.732      0.000

CHEAT1   ON
   CHJUST1            0.730      0.083      8.754      0.000
   PERF1             -0.035      0.046     -0.755      0.450
   PEER1              0.057      0.061      0.934      0.350
   SUB1              -0.258      0.063     -4.114      0.000
   HON1              -0.160      0.053     -3.021      0.003
   TEACHER1           0.146      0.089      1.646      0.100
   CURUSE1            0.062      0.066      0.940      0.347

SURF1    ON
   CHJUST1            0.505      0.077      6.541      0.000
   PERF1              0.157      0.065      2.415      0.016
   SUB1              -0.216      0.065     -3.301      0.001
   TEACHER1           0.077      0.106      0.730      0.466
   CURUSE1           -0.073      0.088     -0.827      0.408

CHJUST1  ON
   SUB1               0.071      0.063      1.119      0.263
   PEER1              0.379      0.065      5.847      0.000
   PERF1              0.237      0.057      4.198      0.000
   TEACHER1          -0.249      0.099     -2.526      0.012
   CURUSE1           -0.064      0.079     -0.813      0.416
   HON1              -0.242      0.060     -4.045      0.000

PEER1    ON
   PERF1              0.132      0.063      2.103      0.035
   TEACHER1          -0.396      0.098     -4.051      0.000
   CURUSE1            0.049      0.095      0.512      0.609
   HON1              -0.135      0.052     -2.564      0.010

PERF1    ON
   SUB1              -0.072      0.061     -1.170      0.242

TEACHER1 ON
   SUB1               0.493      0.046     10.638      0.000

CURUSE1  ON
   SUB1               0.514      0.044     11.582      0.000
```

```
CHEAT1   ON
   GENDER              0.028       0.043       0.652       0.514
   GRADE               0.014       0.039       0.354       0.723

SURF1    ON
   GENDER             -0.051       0.048      -1.057       0.290
   GRADE              -0.048       0.046      -1.055       0.291

CHJUST1  ON
   GENDER              0.097       0.046       2.092       0.036
   GRADE              -0.057       0.045      -1.252       0.210

PEER1    ON
   GENDER              0.166       0.050       3.350       0.001
   GRADE              -0.154       0.048      -3.190       0.001

PERF1    ON
   GENDER              0.180       0.057       3.179       0.001
   GRADE               0.039       0.052       0.746       0.456

TEACHER1 ON
   GENDER             -0.054       0.050      -1.098       0.272
   GRADE               0.123       0.046       2.681       0.007

CURUSE1  ON
   GENDER             -0.072       0.043      -1.697       0.090
   GRADE               0.103       0.042       2.462       0.014

HON1     ON
   GENDER             -0.086       0.051      -1.705       0.088
   GRADE              -0.014       0.048      -0.294       0.769

SUB1     ON
   GENDER              0.268       0.045       5.952       0.000
   GRADE              -0.030       0.047      -0.640       0.522

SUB1     WITH
   HON1                0.210       0.050       4.158       0.000

PERF1    WITH
   TEACHER1           -0.122       0.071      -1.715       0.086
   CURUSE1             0.026       0.061       0.431       0.667

TEACHER1 WITH
   CURUSE1             0.618       0.046      13.344       0.000

SURF1    WITH
   CHEAT1              0.079       0.083       0.943       0.345

Intercepts
   SUB2                1.674       0.182       9.186       0.000
   SUB3                1.656       0.175       9.472       0.000
   SUB5                1.657       0.175       9.447       0.000
   SUB13               1.831       0.185       9.887       0.000
   SUB15               1.263       0.152       8.317       0.000
   HON_1               2.733       0.122      22.337       0.000
   HON6                2.993       0.172      17.420       0.000
   HON8                2.577       0.123      20.896       0.000
   HON9                2.828       0.190      14.915       0.000
   HON10               2.388       0.161      14.803       0.000
   HON11               2.399       0.129      18.622       0.000
   PERF61              2.643       0.163      16.244       0.000
   PERF69              2.118       0.184      11.505       0.000
   PERF74              2.144       0.160      13.369       0.000
   PERF75              2.314       0.169      13.664       0.000
   GTEACH18            2.771       0.122      22.784       0.000
   GTEACH33            1.695       0.154      10.995       0.000
   GTEACH39            1.988       0.147      13.571       0.000
```

GTEACH50	2.008	0.154	13.017	0.000
GTEACH62	1.926	0.151	12.766	0.000
GTEACH67	1.879	0.154	12.183	0.000
GTEACH68	1.811	0.118	15.324	0.000
GTEACH77	2.175	0.146	14.860	0.000
CURUSE19	1.915	0.168	11.364	0.000
CURUSE53	1.942	0.189	10.266	0.000
CURUSE56	1.893	0.175	10.809	0.000
CURUSE64	1.886	0.195	9.696	0.000
TRANS28	1.945	0.149	13.093	0.000
TRANS32	2.230	0.148	15.026	0.000
TRANS66	2.087	0.146	14.290	0.000
AUTH44	2.044	0.165	12.389	0.000
AUTH60	2.069	0.158	13.117	0.000
AUTH71	2.258	0.139	16.202	0.000
AUTH78	1.933	0.152	12.688	0.000
PEER24	3.301	0.189	17.450	0.000
PEER31	2.779	0.168	16.534	0.000
PEER40	2.708	0.170	15.957	0.000
PEER58	3.081	0.194	15.913	0.000
PEER65	2.996	0.166	18.099	0.000
SURF87	4.110	0.220	18.669	0.000
SURF88	2.392	0.103	23.151	0.000
SURF91	3.076	0.193	15.962	0.000
SURF97	3.350	0.210	15.987	0.000
CHJUST79	4.291	0.268	16.029	0.000
CHJUST86	3.200	0.214	14.932	0.000
CHJUST99	3.205	0.184	17.375	0.000
CHEAT84	4.436	0.302	14.672	0.000
CHEAT92	3.973	0.261	15.252	0.000
CHEAT95	3.532	0.219	16.113	0.000
Residual Variances				
SUB2	0.278	0.035	7.874	0.000
SUB3	0.326	0.031	10.370	0.000
SUB5	0.311	0.033	9.327	0.000
SUB13	0.293	0.031	9.577	0.000
SUB15	0.486	0.038	12.717	0.000
HON_1	0.824	0.042	19.752	0.000
HON6	0.478	0.043	11.138	0.000
HON8	0.689	0.042	16.273	0.000
HON9	0.246	0.036	6.893	0.000
HON10	0.381	0.040	9.502	0.000
HON11	0.687	0.045	15.423	0.000
PERF61	0.683	0.049	13.878	0.000
PERF69	0.386	0.056	6.840	0.000
PERF74	0.622	0.055	11.388	0.000
PERF75	0.631	0.053	11.941	0.000
GTEACH18	0.838	0.037	22.408	0.000
GTEACH33	0.422	0.037	11.401	0.000
GTEACH39	0.529	0.042	12.466	0.000
GTEACH50	0.480	0.046	10.486	0.000
GTEACH62	0.489	0.045	10.977	0.000
GTEACH67	0.442	0.041	10.684	0.000
GTEACH68	0.722	0.048	14.930	0.000
GTEACH77	0.558	0.056	10.037	0.000
CURUSE19	0.481	0.045	10.581	0.000
CURUSE53	0.228	0.029	7.939	0.000
CURUSE56	0.367	0.047	7.737	0.000
CURUSE64	0.198	0.029	6.884	0.000
TRANS28	0.593	0.043	13.752	0.000
TRANS32	0.618	0.049	12.688	0.000
TRANS66	0.606	0.053	11.486	0.000
AUTH44	0.493	0.045	11.059	0.000
AUTH60	0.472	0.042	11.104	0.000
AUTH71	0.608	0.048	12.689	0.000
AUTH78	0.535	0.046	11.663	0.000
PEER24	0.591	0.059	10.050	0.000

```
    PEER31             0.604      0.052     11.553      0.000
    PEER40             0.536      0.054      9.951      0.000
    PEER58             0.471      0.062      7.639      0.000
    PEER65             0.726      0.059     12.402      0.000
    SURF87             0.694      0.059     11.833      0.000
    SURF88             0.872      0.036     23.966      0.000
    SURF91             0.414      0.053      7.879      0.000
    SURF97             0.353      0.063      5.614      0.000
    CHJUST79           0.376      0.045      8.329      0.000
    CHJUST86           0.463      0.051      9.023      0.000
    CHJUST99           0.727      0.052     14.054      0.000
    CHEAT84            0.297      0.053      5.656      0.000
    CHEAT92            0.234      0.035      6.701      0.000
    CHEAT95            0.424      0.055      7.664      0.000
    SUB1               0.928      0.024     38.379      0.000
    HON1               0.992      0.009    113.341      0.000
    PERF1              0.967      0.020     49.098      0.000
    GTEACH1            0.313      0.053      5.935      0.000
    CURUSE1            0.742      0.044     16.898      0.000
    ASSESS1            0.176      0.059      2.960      0.003
    PEER1              0.752      0.044     17.164      0.000
    SURF1              0.588      0.061      9.678      0.000
    CHJUST1            0.453      0.054      8.457      0.000
    CHEAT1             0.328      0.054      6.030      0.000
    TEACHER1           0.755      0.044     17.288      0.000


R-SQUARE

    Observed                                            Two-Tailed
    Variable        Estimate       S.E.  Est./S.E.    P-Value

    SUB2               0.722      0.035     20.410      0.000
    SUB3               0.674      0.031     21.418      0.000
    SUB5               0.689      0.033     20.698      0.000
    SUB13              0.707      0.031     23.113      0.000
    SUB15              0.514      0.038     13.435      0.000
    HON_1              0.176      0.042      4.209      0.000
    HON6               0.522      0.043     12.177      0.000
    HON8               0.311      0.042      7.361      0.000
    HON9               0.754      0.036     21.169      0.000
    HON10              0.619      0.040     15.470      0.000
    HON11              0.313      0.045      7.032      0.000
    PERF61             0.317      0.049      6.433      0.000
    PERF69             0.614      0.056     10.862      0.000
    PERF74             0.378      0.055      6.907      0.000
    PERF75             0.369      0.053      6.974      0.000
    GTEACH18           0.162      0.037      4.317      0.000
    GTEACH33           0.578      0.037     15.589      0.000
    GTEACH39           0.471      0.042     11.105      0.000
    GTEACH50           0.520      0.046     11.371      0.000
    GTEACH62           0.511      0.045     11.468      0.000
    GTEACH67           0.558      0.041     13.487      0.000
    GTEACH68           0.278      0.048      5.760      0.000
    GTEACH77           0.442      0.056      7.948      0.000
    CURUSE19           0.519      0.045     11.411      0.000
    CURUSE53           0.772      0.029     26.940      0.000
    CURUSE56           0.633      0.047     13.355      0.000
    CURUSE64           0.802      0.029     27.920      0.000
    TRANS28            0.407      0.043      9.456      0.000
    TRANS32            0.382      0.049      7.851      0.000
    TRANS66            0.394      0.053      7.461      0.000
    AUTH44             0.507      0.045     11.376      0.000
    AUTH60             0.528      0.042     12.430      0.000
    AUTH71             0.392      0.048      8.167      0.000
    AUTH78             0.465      0.046     10.151      0.000
    PEER24             0.409      0.059      6.966      0.000
    PEER31             0.396      0.052      7.568      0.000
```

PEER40	0.464	0.054	8.600	0.000
PEER58	0.529	0.062	8.566	0.000
PEER65	0.274	0.059	4.677	0.000
SURF87	0.306	0.059	5.209	0.000
SURF88	0.128	0.036	3.503	0.000
SURF91	0.586	0.053	11.169	0.000
SURF97	0.647	0.063	10.301	0.000
CHJUST79	0.624	0.045	13.800	0.000
CHJUST86	0.537	0.051	10.449	0.000
CHJUST99	0.273	0.052	5.273	0.000
CHEAT84	0.703	0.053	13.386	0.000
CHEAT92	0.766	0.035	21.986	0.000
CHEAT95	0.576	0.055	10.427	0.000

Latent Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SUB1	0.072	0.024	2.973	0.003
HON1	0.008	0.009	0.891	0.373
PERF1	0.033	0.020	1.683	0.092
GTEACH1	0.687	0.053	13.003	0.000
CURUSE1	0.258	0.044	5.879	0.000
ASSESS1	0.824	0.059	13.866	0.000
PEER1	0.248	0.044	5.654	0.000
SURF1	0.412	0.061	6.768	0.000
CHJUST1	0.547	0.054	10.221	0.000
CHEAT1	0.672	0.054	12.339	0.000
TEACHER1	0.245	0.044	5.601	0.000

QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix 0.676E-04
(ratio of smallest to largest eigenvalue)

TECHNICAL 4 OUTPUT

ESTIMATES DERIVED FROM THE MODEL

ESTIMATED MEANS FOR THE LATENT VARIABLES

	SUB1	HON1	PERF1	GTEACH1	CURUSE1
1	0.692	-0.096	0.418	0.199	0.379

ESTIMATED MEANS FOR THE LATENT VARIABLES

	ASSESS1	PEER1	SURF1	CHJUST1	CHEAT1
1	0.295	0.015	-0.126	0.200	0.214

ESTIMATED MEANS FOR THE LATENT VARIABLES

	TEACHER1	GENDER	GRADE
1	0.199	1.592	1.438

S.E. FOR ESTIMATED MEANS FOR THE LATENT VARIABLES

	SUB1	HON1	PERF1	GTEACH1	CURUSE1
1	0.174	0.063	0.142	0.081	0.171

S.E. FOR ESTIMATED MEANS FOR THE LATENT VARIABLES

	ASSESS1	PEER1	SURF1	CHJUST1	CHEAT1
1	0.115	0.157	0.122	0.158	0.158

S.E. FOR ESTIMATED MEANS FOR THE LATENT VARIABLES

	TEACHER1	GENDER	GRADE
1	0.081	0.022	0.022

ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	SUB1	HON1	PERF1	GTEACH1	CURUSE1
SUB1	0.780				
HON1	0.047	0.090			
PERF1	-0.014	-0.006	0.429		
GTEACH1	0.148	0.010	-0.027	0.180	
CURUSE1	0.352	0.023	0.002	0.203	0.653
ASSESS1	0.219	0.014	-0.040	0.183	0.300
PEER1	-0.095	-0.040	0.097	-0.101	-0.147
SURF1	-0.152	-0.035	0.118	-0.066	-0.132
CHJUST1	-0.103	-0.075	0.175	-0.119	-0.192
CHEAT1	-0.211	-0.106	0.125	-0.082	-0.142
TEACHER1	0.148	0.010	-0.027	0.123	0.203
GENDER	0.116	-0.013	0.053	0.015	0.029
GRADE	-0.005	-0.003	0.017	0.020	0.037

ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	ASSESS1	PEER1	SURF1	CHJUST1	CHEAT1
ASSESS1	0.328				
PEER1	-0.150	0.536			
SURF1	-0.097	0.144	0.307		
CHJUST1	-0.176	0.324	0.237	0.544	
CHEAT1	-0.122	0.281	0.225	0.446	0.628
TEACHER1	0.183	-0.101	-0.066	-0.119	-0.082
GENDER	0.022	0.057	0.005	0.075	0.057
GRADE	0.029	-0.063	-0.029	-0.049	-0.026

ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	TEACHER1	GENDER	GRADE
TEACHER1	0.123		
GENDER	0.015	0.241	
GRADE	0.020	0.016	0.246

S.E. FOR ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	SUB1	HON1	PERF1	GTEACH1	CURUSE1
SUB1	0.064				
HON1	0.017	0.023			
PERF1	0.033	0.004	0.076		
GTEACH1	0.027	0.004	0.016	0.045	
CURUSE1	0.051	0.009	0.033	0.035	0.075
ASSESS1	0.037	0.006	0.024	0.030	0.040
PEER1	0.028	0.012	0.032	0.023	0.035
SURF1	0.034	0.009	0.031	0.016	0.030
CHJUST1	0.042	0.015	0.037	0.024	0.036
CHEAT1	0.043	0.019	0.030	0.018	0.033
TEACHER1	0.027	0.004	0.016	0.032	0.035
GENDER	0.020	0.008	0.017	0.009	0.019
GRADE	0.021	0.007	0.017	0.009	0.020

S.E. FOR ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	ASSESS1	PEER1	SURF1	CHJUST1	CHEAT1
ASSESS1	0.049				
PEER1	0.030	0.082			
SURF1	0.023	0.029	0.064		
CHJUST1	0.030	0.039	0.043	0.064	
CHEAT1	0.026	0.041	0.046	0.053	0.078
TEACHER1	0.030	0.023	0.016	0.024	0.018
GENDER	0.013	0.019	0.014	0.019	0.020
GRADE	0.013	0.020	0.014	0.019	0.019

S.E. FOR ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	TEACHER1	GENDER	GRADE
TEACHER1	0.032		
GENDER	0.009	0.004	
GRADE	0.009	0.011	0.003

ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	SUB1	HON1	PERF1	GTEACH1	CURUSE1
SUB1	1.000				
HON1	0.179	1.000			
PERF1	-0.024	-0.029	1.000		
GTEACH1	0.395	0.075	-0.098	1.000	
CURUSE1	0.494	0.096	0.004	0.591	1.000
ASSESS1	0.433	0.082	-0.107	0.752	0.648
PEER1	-0.146	-0.181	0.202	-0.325	-0.249
SURF1	-0.312	-0.210	0.325	-0.280	-0.295
CHJUST1	-0.158	-0.341	0.362	-0.380	-0.322
CHEAT1	-0.301	-0.447	0.240	-0.245	-0.222
TEACHER1	0.477	0.090	-0.118	0.829	0.714
GENDER	0.266	-0.087	0.163	0.071	0.072
GRADE	-0.012	-0.020	0.052	0.094	0.092

ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	ASSESS1	PEER1	SURF1	CHJUST1	CHEAT1
ASSESS1	1.000				
PEER1	-0.357	1.000			
SURF1	-0.307	0.355	1.000		
CHJUST1	-0.416	0.600	0.581	1.000	
CHEAT1	-0.268	0.484	0.512	0.763	1.000
TEACHER1	0.908	-0.393	-0.338	-0.458	-0.295
GENDER	0.077	0.159	0.020	0.206	0.145
GRADE	0.103	-0.173	-0.107	-0.134	-0.065

ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	TEACHER1	GENDER	GRADE
TEACHER1	1.000		
GENDER	0.085	1.000	
GRADE	0.113	0.067	1.000

S.E. FOR ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	SUB1	HON1	PERF1	GTEACH1	CURUSE1
SUB1	0.000				
HON1	0.053	0.000			
PERF1	0.057	0.017	0.000		
GTEACH1	0.040	0.024	0.058	0.000	
CURUSE1	0.044	0.029	0.061	0.039	0.000
ASSESS1	0.046	0.027	0.063	0.034	0.039
PEER1	0.039	0.051	0.064	0.049	0.053
SURF1	0.053	0.041	0.060	0.051	0.053
CHJUST1	0.062	0.059	0.058	0.053	0.052
CHEAT1	0.054	0.050	0.050	0.043	0.049
TEACHER1	0.045	0.029	0.069	0.032	0.037
GENDER	0.045	0.050	0.052	0.043	0.047
GRADE	0.048	0.048	0.052	0.042	0.047

S.E. FOR ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	ASSESS1	PEER1	SURF1	CHJUST1	CHEAT1
ASSESS1	0.000				
PEER1	0.051	0.000			
SURF1	0.057	0.047	0.000		
CHJUST1	0.053	0.052	0.053	0.000	
CHEAT1	0.048	0.052	0.053	0.036	0.000
TEACHER1	0.033	0.055	0.061	0.058	0.051
GENDER	0.047	0.051	0.052	0.051	0.048
GRADE	0.046	0.050	0.052	0.052	0.048

S.E. FOR ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	TEACHER1	GENDER	GRADE
TEACHER1	0.000		
GENDER	0.052	0.000	
GRADE	0.050	0.045	0.000

Appendix M:

Indirect effects (standardized) in Model 3 for the co-ed sample, Time 1

```
STANDARDIZED TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS

STDYX Standardization

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Effects from CURUSE1 to CHEAT1

  Total                0.032      0.075      0.424      0.671
  Total indirect      -0.031      0.061     -0.498      0.618

  Specific indirect

    CHEAT1
    PEER1
    CURUSE1            0.003      0.006      0.444      0.657

    CHEAT1
    CHJUST1
    CURUSE1           -0.047      0.058     -0.815      0.415

    CHEAT1
    CHJUST1
    PEER1
    CURUSE1            0.013      0.026      0.513      0.608

  Direct
    CHEAT1
    CURUSE1            0.062      0.066      0.940      0.347


Effects from PERF1 to CHEAT1

  Total                0.183      0.049      3.708      0.000
  Total indirect       0.217      0.049      4.436      0.000

  Specific indirect

    CHEAT1
    PEER1
    PERF1              0.008      0.009      0.846      0.398

    CHEAT1
    CHJUST1
    PERF1              0.173      0.047      3.715      0.000

    CHEAT1
    CHJUST1
    PEER1
    PERF1              0.037      0.019      1.973      0.049

  Direct
    CHEAT1
    PERF1             -0.035      0.046     -0.755      0.450
```

```
Effects from PEER1 to CHEAT1

  Total                0.334      0.059      5.664      0.000
  Total indirect       0.277      0.057      4.898      0.000

  Specific indirect

    CHEAT1
    CHJUST1
    PEER1              0.277      0.057      4.898      0.000

  Direct
    CHEAT1
    PEER1              0.057      0.061      0.934      0.350


Effects from SUB1 to CHEAT1

  Total               -0.286      0.054     -5.302      0.000
  Total indirect      -0.028      0.050     -0.555      0.579

  Specific indirect

    CHEAT1
    PERF1
    SUB1               0.002      0.004      0.620      0.535

    CHEAT1
    CURUSE1
    SUB1               0.032      0.034      0.939      0.348

    CHEAT1
    CHJUST1
    SUB1               0.052      0.048      1.082      0.279

    CHEAT1
    TEACHER1
    SUB1               0.072      0.045      1.616      0.106

    CHEAT1
    PEER1
    PERF1
    SUB1              -0.001      0.001     -0.700      0.484

    CHEAT1
    PEER1
    CURUSE1
    SUB1               0.001      0.003      0.444      0.657

    CHEAT1
    PEER1
    TEACHER1
    SUB1              -0.011      0.013     -0.884      0.377

    CHEAT1
    CHJUST1
    PERF1
    SUB1              -0.012      0.011     -1.108      0.268

    CHEAT1
    CHJUST1
    CURUSE1
    SUB1              -0.024      0.030     -0.811      0.417
```

```
CHEAT1
    CHJUST1
    TEACHER1
    SUB1              -0.090      0.039     -2.287      0.022

    CHEAT1
    CHJUST1
    PEER1
    PERF1
    SUB1              -0.003      0.003     -0.970      0.332

 CHEAT1
    CHJUST1
    PEER1
    CURUSE1
    SUB1               0.007      0.013      0.515      0.607

    CHEAT1
    CHJUST1
    PEER1
    TEACHER1
    SUB1              -0.054      0.018     -3.066      0.002

  Direct
    CHEAT1
    SUB1              -0.258      0.063     -4.114      0.000


Effects from HON1 to CHEAT1

  Total               -0.381      0.055     -6.964      0.000
  Total indirect      -0.221      0.052     -4.268      0.000

  Specific indirect

    CHEAT1
    PEER1
    HON1              -0.008      0.009     -0.877      0.380

    CHEAT1
    CHJUST1
    HON1              -0.176      0.050     -3.527      0.000

    CHEAT1
    CHJUST1
    PEER1
    HON1              -0.037      0.017     -2.203      0.028

  Direct
    CHEAT1
    HON1              -0.160      0.053     -3.021      0.003


Effects from TEACHER1 to CHEAT1

  Total               -0.168      0.080     -2.101      0.036
  Total indirect      -0.314      0.082     -3.853      0.000

  Specific indirect

    CHEAT1
    PEER1
    TEACHER1          -0.023      0.025     -0.899      0.369

    CHEAT1
    CHJUST1
    TEACHER1          -0.182      0.079     -2.308      0.021
```

```
    CHEAT1
    CHJUST1
    PEER1
    TEACHER1          -0.110      0.034     -3.215      0.001

  Direct
    CHEAT1
    TEACHER1           0.146      0.089      1.646      0.100


Effects from CURUSE1 to SURF1

  Total               -0.096      0.090     -1.062      0.288
  Total indirect      -0.023      0.041     -0.561      0.575

  Specific indirect

    SURF1
    CHJUST1
    CURUSE1           -0.032      0.040     -0.812      0.417

    SURF1
    CHJUST1
    PEER1
    CURUSE1            0.009      0.018      0.512      0.609

  Direct
    SURF1
    CURUSE1           -0.073      0.088     -0.827      0.408


Effects from PERF1 to SURF1

  Total                0.303      0.061      4.974      0.000
  Total indirect       0.145      0.039      3.744      0.000

  Specific indirect

    SURF1
    CHJUST1
    PERF1              0.120      0.036      3.340      0.001

    SURF1
    CHJUST1
    PEER1
    PERF1              0.025      0.013      1.932      0.053

  Direct
    SURF1
    PERF1              0.157      0.065      2.415      0.016


Effects from SUB1 to SURF1

  Total               -0.313      0.055     -5.690      0.000
  Total indirect      -0.097      0.049     -1.980      0.048

  Specific indirect

    SURF1
    PERF1
    SUB1              -0.011      0.011     -1.035      0.301

    SURF1
    CURUSE1
    SUB1              -0.037      0.045     -0.824      0.410
```

```
SURF1
    CHJUST1
    SUB1                0.036      0.033      1.087      0.277

    SURF1
    TEACHER1
    SUB1                0.038      0.053      0.723      0.469

    SURF1
    CHJUST1
    PERF1
    SUB1               -0.009      0.008     -1.118      0.264

    SURF1
    CHJUST1
    CURUSE1
    SUB1               -0.017      0.021     -0.808      0.419

    SURF1
    CHJUST1
    TEACHER1
    SUB1               -0.062      0.027     -2.279      0.023

    SURF1
    CHJUST1
    PEER1
    PERF1
    SUB1               -0.002      0.002     -0.980      0.327

    SURF1
    CHJUST1
    PEER1
    CURUSE1
    SUB1                0.005      0.009      0.514      0.607

    SURF1
    CHJUST1
    PEER1
    TEACHER1
    SUB1               -0.038      0.013     -2.922      0.003

  Direct
    SURF1
    SUB1               -0.216      0.065     -3.301      0.001


Effects from TEACHER1 to SURF1

  Total                -0.125      0.099     -1.261      0.207
  Total indirect       -0.202      0.058     -3.466      0.001

  Specific indirect

    SURF1
    CHJUST1
    TEACHER1           -0.126      0.054     -2.311      0.021

    SURF1
    CHJUST1
    PEER1
    TEACHER1           -0.076      0.025     -3.075      0.002

  Direct
    SURF1
    TEACHER1            0.077      0.106      0.730      0.466
```

```
Effects from SUB1 to CHJUST1

  Total               -0.170      0.059     -2.879      0.004
  Total indirect      -0.241      0.044     -5.530      0.000

  Specific indirect

    CHJUST1
    PERF1
    SUB1              -0.017      0.015     -1.136      0.256

    CHJUST1
    CURUSE1
    SUB1              -0.033      0.041     -0.810      0.418

    CHJUST1
    TEACHER1
    SUB1              -0.123      0.049     -2.485      0.013

    CHJUST1
    PEER1
    PERF1
    SUB1              -0.004      0.004     -0.988      0.323

    CHJUST1
    PEER1
    CURUSE1
    SUB1               0.009      0.018      0.517      0.605

    CHJUST1
    PEER1
    TEACHER1
    SUB1              -0.074      0.023     -3.241      0.001

  Direct
    CHJUST1
    SUB1               0.071      0.063      1.119      0.263


Effects from CURUSE1 to CHJUST1

  Total               -0.046      0.082     -0.560      0.576
  Total indirect       0.018      0.036      0.514      0.607

  Specific indirect

    CHJUST1
    PEER1
    CURUSE1            0.018      0.036      0.514      0.607

  Direct
    CHJUST1
    CURUSE1           -0.064      0.079     -0.813      0.416


Effects from PERF1 to CHJUST1

  Total                0.287      0.057      5.014      0.000
  Total indirect       0.050      0.025      2.031      0.042

  Specific indirect

    CHJUST1
    PEER1
    PERF1              0.050      0.025      2.031      0.042
```

```
 Direct
   CHJUST1
   PERF1              0.237      0.057      4.198      0.000


Effects from TEACHER1 to CHJUST1

  Total              -0.400      0.095     -4.227      0.000
  Total indirect     -0.150      0.044     -3.454      0.001

  Specific indirect

    CHJUST1
    PEER1
    TEACHER1         -0.150      0.044     -3.454      0.001

  Direct
    CHJUST1
    TEACHER1         -0.249      0.099     -2.526      0.012


Effects from HON1 to CHJUST1

  Total              -0.293      0.061     -4.791      0.000
  Total indirect     -0.051      0.022     -2.299      0.022

  Specific indirect

    CHJUST1
    PEER1
    HON1             -0.051      0.022     -2.299      0.022

  Direct
    CHJUST1
    HON1             -0.242      0.060     -4.045      0.000

Effects from CURUSE1 to PEER1

  Total               0.049      0.095      0.512      0.609
  Total indirect      0.000      0.000      0.000      1.000

  Direct
    PEER1
    CURUSE1           0.049      0.095      0.512      0.609


Effects from HON1 to PEER1

  Total              -0.135      0.052     -2.564      0.010
  Total indirect      0.000      0.000      0.000      1.000

  Direct
    PEER1
    HON1             -0.135      0.052     -2.564      0.010
```

Appendix N:

Stepwise regression of predictors of Time 1 variables

Table N1

Standardized coefficients of stepwise regression on Peer norms and Justifiability of cheating, Time 1

	Peer norms						Justifiability of cheating					
	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 1	Step 2	Step 3	Step 4	Step5	Step 6
	β	β	β	β	β	β	β	β	β	β	β	β
Curuse	-.250***	-.249***	---	-.225***	-.228***	.050	-.329***	-.334***	-.303***	-.284***	-.297***	-.050
Gender		.180***	---	.168**	.150**	.165**		.235***	.251***	.222***	.182***	.187***
Grade		-.163**	---	-.167**	-.173***	-.157**		-.129**	-.132**	-.133**	-.143**	-.125**
Sub			---	---	---	---			-.062	-.018	.000	.061
Hon				-.171**	-.162**	-.113*				-.308***	-.293***	-.250***
Perf					.161*	.117					.323***	.280***
Teacher						-.401***						-.401***

Table N2

Standardized coefficients of stepwise regression on Surface learning strategies and Self-reported cheating, Time 1

	Surface learning strategies						Self-reported cheating					
	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 1	Step 2	Step 3	Step 4	Step5	Step 6
	β	β	β	β	β	β	β	β	β	β	β	β
Curuse	-.184***	-.282***	-.106*	---	-.177**	-.079	-.231***	-.237***	-.087	-.066	-.073	.024
Gender		.039	.098	---	.054	.054		.161**	.225***	.182***	.159***	.161***
Grade		-.081	-.099*	---	-.106*	-.099*		-.055	-.072	-.079	-.085	-.077
Sub			-.144***	---	-.225**	-.198**			-.305***	-.241***	-.232***	-.209***
Hon				---	---	---				-.395***	-.386***	-.307***
Perf					.307***	.292***					.189***	.174***
Teacher						-.158						-.156*

Note. All models satisfied fit requirements for multivariate models (see Table 5.1); CURUSE = Usefulness of curriculum; Grade = Grade-level; SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance structure; TEACHER = Teacher quality; **p* < .05, ***p* < .01, ****p* < .001

Appendix O:

Standardized beta coefficients for the Time 1 Co-ed model, estimated with composite scores

Table O1

Co-ed sample model: standardized beta coefficients estimated with composite scores, Time 1

					Predictors				
N = 493	Grade	Gender	Sub	Hon	Perf	Curuse	Teacher	Peer	Chjust
Sub	-.033	.278***							
Hon	-.008	-.061	---						
Perf	.046	.166**	-.049	---					
Curuse	.108*	-.077	.515***	---	---				
Teacher	.120**	-.045	.495***	---	---	---			
Peer	-.158**	.170**	---	-.150**	.098	.041	-.392***		
Chjust	-.061	.117*	.103	-.235***	.243***	-.063	-.321**	.373***	
Surf	-.041	-.074	-.248***	---	.135*	-.086	.147	---	.588***
Cheat	.005	.016	-.255***	-.174**	-.047	.027	.188	.005	.805***

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level.

Appendix P: Equivalent model 1, Time 1

Usefulness of curriculum positioned as a predictor of *Teacher quality*

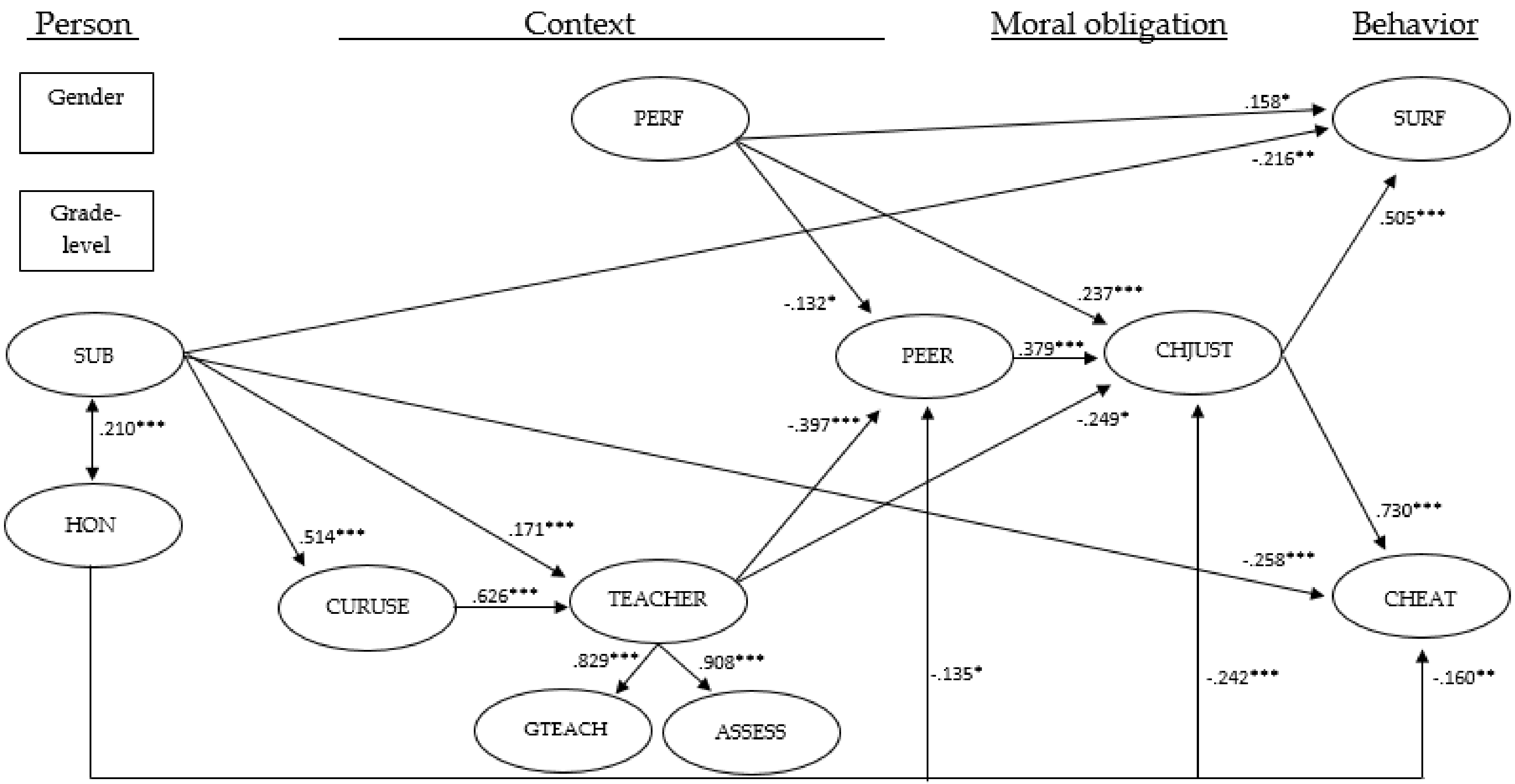

Figure P1. Equivalent Model 1, Time 1: *Usefulness of curriculum* positioned as a correlate of *Teacher quality* (*N* = 493). $\chi^2(1176) = 1962$; *RMSEA* = .037, *CI*s = .034 - .040, *pclose* = 1.00; *TLI* = .91; *CFI* = .92; *SRMR* = .059; *N:q* = 2.5; SCF = .962. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level

Appendix Q: Equivalent model 2, Time 1

Peer norms positioned as a correlate of class context

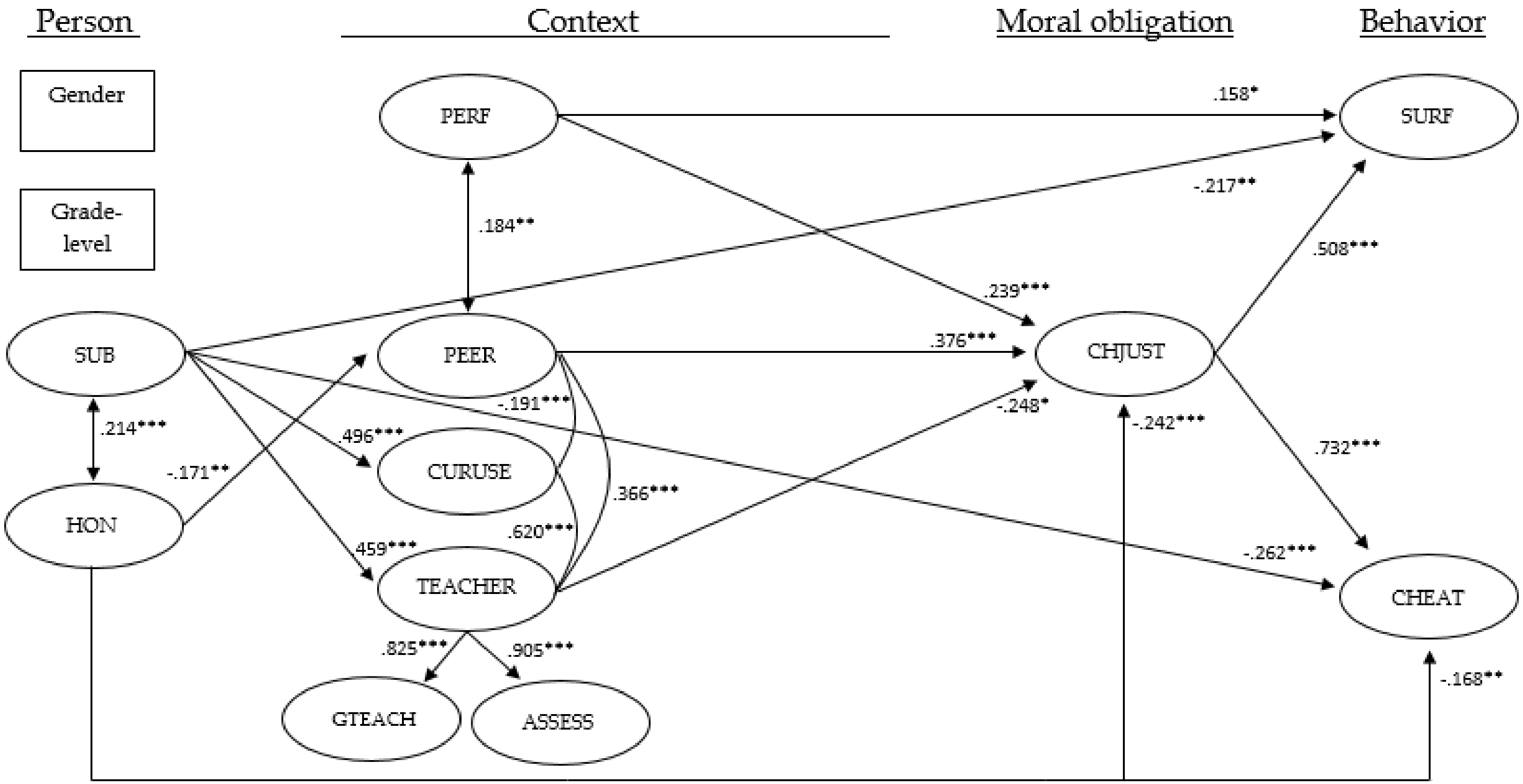

Figure Q1. Equivalent Model 2, Time 1: *Peer norms* positioned as a correlate of class context (*N* = 493). $\chi^2(1175) = 1971$; *RMSEA* = .037, *CI*s = .034 - .040, *pclose* = 1.00; *TLI* = .91; *CFI* = .92; *SRMR* = .059; *N:q* = 2.5; SCF = .962. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level.

Appendix R: Equivalent model 3, Time 1

***Peer norms* positioned as a predictor of class context**

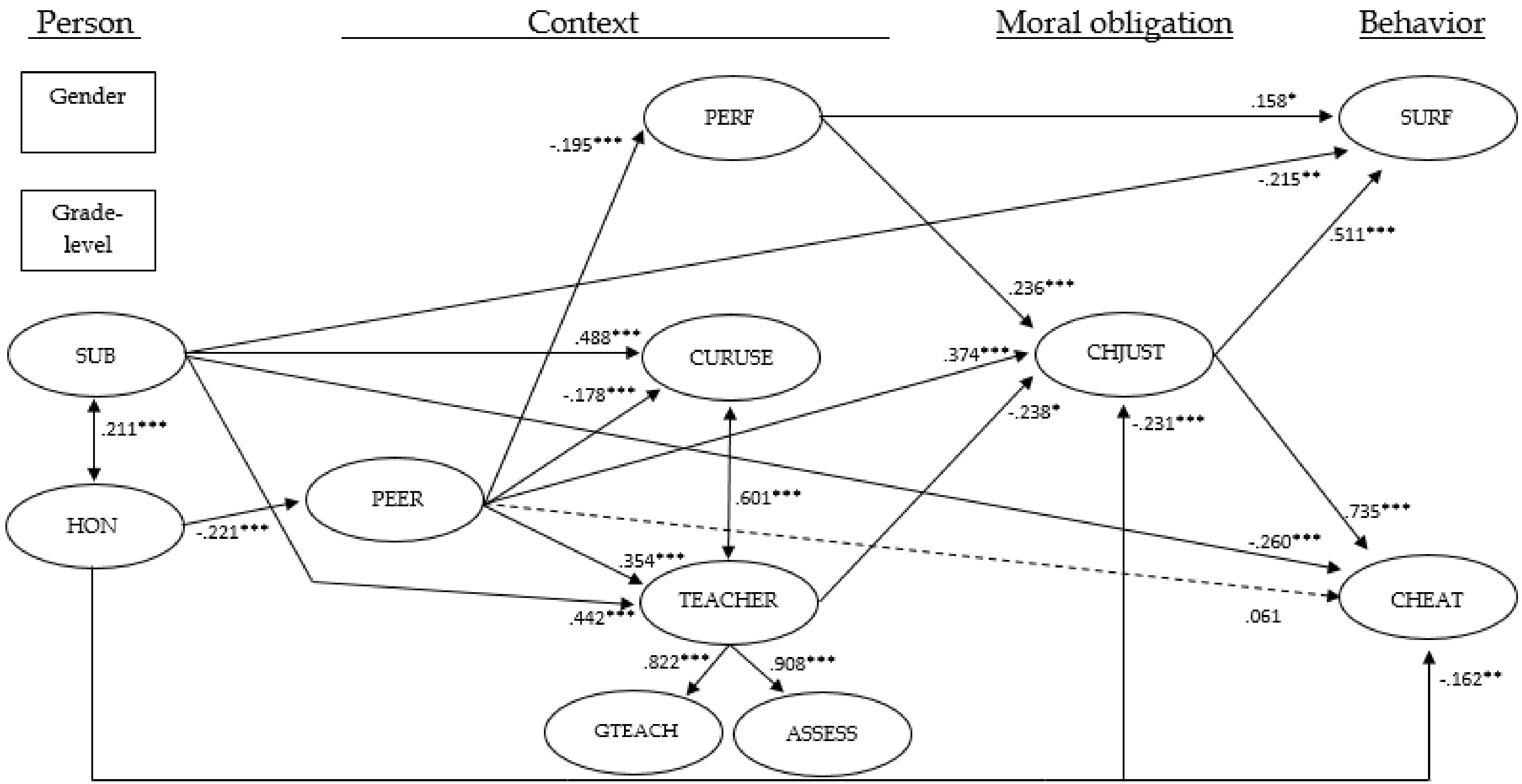

Figure R1. Equivalent Model 3, Time 1: *Peer norms* positioned as a predictor of class context (*N* = 493). $\chi^2(1175) = 1965$; *RMSEA* = .037, *CI*s = .034 - .042, *pclose* = 1.00; *TLI* = .91; *CFI* = .92; *SRMR* = .056; *N:q* = 2.5; SCF = .962. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level.

Appendix S:

Item descriptive statistics: Time 2 *vs*. Time 1

Table S1

Item descriptive statistics: Time 2 *vs*. Time 1

	Time 2 sample (*n*=297)						Time 1 sample (*n*=493)						Δ = (Time 2) - (Time 1)				
	Mean	SE	SD	Var.	S	K	Mean	SE	SD	Var.	S	K	ΔMean	ΔSD	ΔVar.	ΔS	ΔK
SUB																	
sub2	2.29	0.06	1.06	1.12	0.43	-0.55	2.43	0.05	1.04	1.08	0.40	-0.41	-0.14	0.02	0.03	0.03	-0.15
sub3	2.12	0.06	1.04	1.07	0.83	-0.22	2.31	0.05	1.01	1.01	0.52	-0.13	-0.19	0.03	0.06	0.31	-0.09
sub4	2.55	0.08	1.34	1.79	0.37	-1.10	2.85	0.06	1.23	1.52	0.08	-0.90	-0.30	0.10	0.26	0.28	-0.20
sub10	2.52	0.06	1.09	1.20	0.32	-0.45	2.72	0.05	1.10	1.20	0.17	-0.56	-0.20	0.00	0.00	0.15	0.10
sub11	1.97	0.06	1.09	1.19	1.08	0.42	1.88	0.05	1.03	1.06	1.00	0.20	0.09	0.06	0.12	0.08	0.22
HON																	
hon1	1.81	0.05	0.79	0.62	1.05	1.87	1.86	0.03	0.72	0.51	0.62	0.60	-0.05	0.07	0.11	0.44	1.27
hon6	2.40	0.05	0.91	0.83	0.54	0.37	2.46	0.04	0.89	0.80	0.39	0.16	-0.06	0.02	0.03	0.15	0.21
hon8	1.88	0.05	0.86	0.74	0.97	0.98	1.90	0.04	0.79	0.63	0.60	0.02	-0.02	0.07	0.11	0.37	0.96
hon9	2.00	0.05	0.87	0.75	0.97	1.49	1.98	0.03	0.78	0.60	0.77	1.19	0.02	0.09	0.15	0.20	0.30
hon10	1.79	0.06	0.98	0.95	1.36	1.66	1.77	0.04	0.83	0.69	1.01	0.78	0.02	0.15	0.26	0.35	0.88
hon11	2.39	0.06	1.01	1.02	0.55	-0.16	2.17	0.04	0.98	0.96	0.65	-0.14	0.22	0.03	0.06	-0.10	-0.02

Note. SE = Standard error; SD = Standard deviation; Var. = variance; S = Skewness; K = Kurtosis; Δ = 'change in'. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level.

Table S1, continued

	Time 2 sample (*n*=297)						Time 1 sample (*n*=493)						Δ = (Time 2) - (Time 1)				
	Mean	SE	SD	Var.	S	K	Mean	SE	SD	Var.	S	K	ΔMean	ΔSD	ΔVar.	ΔS	ΔK
GTEACH																	
gteach18	2.67	0.07	1.13	1.27	0.29	-0.67	3.12	0.05	1.06	1.11	-0.03	-0.57	-0.45	0.07	0.16	0.32	-0.10
gteach33	2.54	0.07	1.21	1.47	0.49	-0.59	2.42	0.05	1.18	1.40	0.67	-0.36	0.12	0.03	0.07	-0.19	-0.23
gteach39	2.51	0.06	1.00	1.01	0.49	-0.19	2.52	0.05	1.09	1.19	0.46	-0.45	-0.01	-0.09	-0.18	0.03	0.26
gteach50	2.46	0.06	1.03	1.06	0.53	-0.25	2.52	0.05	1.07	1.15	0.58	-0.18	-0.06	-0.04	-0.09	-0.05	-0.07
gteach62	2.32	0.06	1.02	1.04	0.50	-0.57	2.39	0.05	1.06	1.12	0.68	0.09	-0.07	-0.04	-0.08	-0.19	-0.66
gteach67	2.32	0.06	1.11	1.23	0.72	-0.05	2.40	0.05	1.08	1.16	0.61	-0.11	-0.08	0.03	0.07	0.11	0.05
gteach68	2.29	0.07	1.11	1.24	0.72	-0.19	2.26	0.05	1.10	1.21	0.66	-0.18	0.03	0.01	0.03	0.06	0.00
gteach77	2.45	0.06	0.99	0.97	0.23	-0.54	2.48	0.04	1.00	0.99	0.55	0.10	-0.03	-0.01	-0.02	-0.32	-0.64
PERF																	
perf61	3.46	0.07	1.13	1.28	-0.38	-0.63	3.49	0.05	1.16	1.36	-0.36	-0.75	-0.03	-0.03	-0.08	-0.02	0.12
perf69	3.37	0.07	1.21	1.47	-0.22	-0.95	3.24	0.06	1.24	1.54	-0.15	-0.99	0.13	-0.03	-0.07	-0.07	0.05
perf74	3.49	0.07	1.15	1.32	-0.36	-0.68	3.17	0.06	1.25	1.56	-0.09	-0.98	0.32	-0.10	-0.24	-0.27	0.30
perf75	3.88	0.07	1.17	1.36	-0.83	-0.26	3.61	0.06	1.34	1.79	-0.54	-0.94	0.27	-0.17	-0.43	-0.28	0.68
CURUSE																	
curuse19	2.39	0.06	1.10	1.21	0.62	-0.21	2.53	0.05	1.12	1.26	0.50	-0.43	-0.14	-0.02	-0.05	0.13	0.22
curuse53	2.48	0.06	1.04	1.08	0.43	-0.36	2.61	0.05	1.11	1.23	0.50	-0.36	-0.13	-0.07	-0.15	-0.07	0.00
curuse56	2.39	0.06	1.04	1.09	0.62	-0.12	2.31	0.05	1.02	1.04	0.58	-0.14	0.08	0.02	0.05	0.04	0.02
curuse64	2.47	0.06	1.08	1.16	0.52	-0.27	2.48	0.05	1.08	1.16	0.51	-0.25	-0.01	0.00	0.00	0.01	-0.02

Table S1, continued

	Time 2 sample (n=297)						Time 1 sample (n=493)						Δ = (Time 2) - (Time 1)				
	Mean	SE	SD	Var.	S	K	Mean	SE	SD	Var.	S	K	ΔMean	ΔSD	ΔVar.	ΔS	ΔK
ASSESS																	
auth44	2.10	0.05	0.85	0.72	0.72	0.65	2.06	0.04	0.86	0.73	0.73	0.68	0.04	-0.01	-0.02	-0.01	-0.03
auth60	2.19	0.05	0.88	0.77	0.56	0.02	2.30	0.04	0.94	0.89	0.50	0.13	-0.11	-0.06	-0.12	0.06	-0.11
auth71	2.14	0.05	0.84	0.71	0.45	-0.12	2.31	0.04	0.90	0.80	0.40	0.07	-0.17	-0.06	-0.10	0.04	-0.19
auth78	2.31	0.05	0.90	0.81	0.51	0.31	2.15	0.04	0.94	0.89	0.68	0.36	0.16	-0.04	-0.08	-0.17	-0.05
trans28	1.97	0.05	0.85	0.73	0.79	0.52	2.04	0.04	0.90	0.81	0.93	1.01	-0.07	-0.04	-0.08	-0.15	-0.50
trans32	2.15	0.05	0.91	0.83	0.79	0.70	2.18	0.04	0.86	0.73	0.57	0.21	-0.03	0.05	0.09	0.22	0.49
trans66	2.18	0.06	0.95	0.90	0.72	0.21	2.29	0.04	0.95	0.91	0.71	0.45	-0.11	-0.01	-0.01	0.01	-0.24
PEER																	
peer24	1.97	0.06	1.01	1.02	0.91	0.41	2.02	0.04	0.99	0.98	1.01	0.78	-0.05	0.02	0.04	-0.10	-0.37
peer31	2.41	0.07	1.12	1.25	0.51	-0.47	2.12	0.04	0.96	0.92	0.70	0.10	0.29	0.16	0.33	-0.19	-0.57
peer40	2.55	0.07	1.21	1.47	0.40	-0.76	2.10	0.05	1.04	1.09	0.85	0.23	0.45	0.17	0.38	-0.45	-1.00
peer58	2.39	0.07	1.14	1.29	0.39	-0.74	2.26	0.05	1.16	1.33	0.72	-0.30	0.13	-0.02	-0.04	-0.33	-0.44
peer65	2.24	0.06	1.02	1.04	0.55	-0.26	3.79	0.05	1.15	1.32	-0.08	-0.02	-1.55	-0.13	-0.28	0.62	-0.24
SURF																	
surf87	3.96	0.06	1.05	1.10	-1.03	0.64	3.99	0.05	1.00	1.01	-0.89	0.22	-0.03	0.04	0.09	-0.14	0.42
surf88	2.88	0.07	1.22	0.05	0.21	-0.89	2.94	0.06	1.27	1.62	0.08	-1.05	-0.06	-0.06	-1.58	0.13	0.17
surf91	3.37	0.07	1.18	1.39	-0.32	-0.78	3.46	0.05	1.20	1.43	-0.31	-0.85	-0.09	-0.02	-0.04	-0.01	0.07
surf97	3.60	0.07	1.13	1.28	-0.58	-0.42	3.60	0.05	1.14	1.30	-0.37	-0.78	0.00	-0.01	-0.02	-0.21	0.36

Table S1, continued

	Time 2 sample (*n*=297)						Time 1 sample (*n*=493)						Δ = (Time 2) - (Time 1)				
	Mean	SE	SD	Var.	S	K	Mean	SE	SD	Var.	S	K	ΔMean	ΔSD	ΔVar.	ΔS	ΔK
CHJUST																	
chjust79	4.22	0.05	0.92	0.85	-1.27	1.61	4.21	0.04	0.94	0.89	-1.06	0.57	0.01	-0.02	-0.04	-0.21	1.04
chjust86	3.94	0.07	1.14	1.14	-0.88	-0.15	3.95	0.05	1.17	1.37	-0.95	-0.02	-0.01	-0.03	-0.23	0.07	-0.13
chjust99	4.14	0.06	1.08	1.08	-1.18	0.68	3.97	0.05	1.19	1.42	-1.01	0.16	0.17	-0.12	-0.34	-0.16	0.52
CHEAT																	
cheat84	4.37	0.06	0.98	0.95	-1.59	1.95	4.41	0.04	0.95	0.90	-1.58	1.63	-0.04	0.02	0.05	-0.02	0.31
cheat92	4.29	0.06	1.08	1.17	-1.44	1.10	4.28	0.05	1.02	1.05	-1.26	0.51	0.01	0.06	0.12	-0.18	0.59
cheat95	4.16	0.06	1.11	1.24	-1.15	0.25	4.16	0.05	1.12	1.25	-1.11	0.09	0.00	-0.01	-0.02	-0.04	0.16

Appendix T:

Gender-specific congeneric model results, Time 2

Table T1
Congeneric model results for male respondents at Time 2

N = 115	CFA									
				Loading	RMSEA					
Scale (# items)	χ^2	*p*	df	range	value	CIs	CFI	TLI	SRMR	Rho
Subject self-concept (5)	.70	.98	5	.48-.90	.000	.000-.000	1.00	1.05	.007	.90
Honesty-trust. self-concept (6)	15.2	.09	9	.40-.87	.077	.000-.143	.97	.94	.041	.83
Performance structure (4)	1.4	.51	2	.57-.86	.000	.000-.165	1.00	1.03	.018	.78
Good teaching (8)	31.8	.046	20	.19-.85	.072	.010-.117	.96	.94	.044	.86
Usefulness of curriculum (4)	1.3	.53	2	.74-.91	.000	.000-.162	1.00	1.02	.009	.91
Assessment quality	30.6	.01	14	.70-.80	.102	.052-.151	.93	.90	.046	.91
Peer norms (5)	1.1	.96	5	.63-.92	.000	.000-.000	1.00	1.05	.012	.85
Surface learning strategies (4)	.26	.61	1	.46-.85	.000	.000-.197	1.00	1.05	.007	.81
Justifiability of cheating (3)	.07	.80	1	.62-.81	.000	.000-.158	1.00	1.08	0.01	.75
Self-reported cheating (3)	.33	.56	1	.68-.92	.000	.000-.205	1.00	1.04	.008	.87

Table T2
Congeneric model results for female respondents at Time 2

(*N* = 182)	CFA									
				Loading	RMSEA					
Scale (# items)	χ^2	*p*	df	range	value	CIs	CFI	TLI	SRMR	Rho
Subject self-concept (5)	6.6	.25	5	.78-.91	.042	.000-.117	1.00	.99	.013	.93
Honesty-trust. self-concept (6)	8.6	.47	9	.54-.89	.000	.000-.081	1.00	1.00	.024	.88
Performance structure (4)	11.9	.003	2	.48-.87	.165	.083-.261	.92	.76	.040	.77
Good teaching (8)	35.1	.02	20	.44-.82	.064	.026-.099	.97	.96	.035	.88
Usefulness of curriculum (4)	1.4	.51	2	.79-.94	.000	.000-.131	1.00	1.01	.007	.93
Assessment quality	32.9	.003	14	.57-.79	.086	.048-.125	.94	.92	.047	.86
Peer norms (5)	9.1	.10	5	.54-.85	.067	.000-.136	.98	.95	.07	.85
Surface learning strategies (4)	3.3	.07	1	.39-.88	.112	.000-.256	.98	.90	.022	.72
Justifiability of cheating (3)	.001	.97	1	.65-.83	.000	.000-.000	1.00	1.05	.001	.78
Self-reported cheating (3)	2.3	.13	1	.65-.96	.083	.000-.233	.99	.96	.113	.85

Appendix U: Male sample model correlation matrices, Time 2

Table U1

CFA: Estimated correlation matrix for the Time 2 male sample

	SUB2	HON2	PERF2	GTEACH2	ASSESS2
SUB2	1.000				
HON2	0.330	1.000			
PERF2	0.282	-0.090	1.000		
GTEACH2	0.421	0.362	0.225	1.000	
ASSESS2	0.457	0.393	0.244	0.921	1.000
CURUSE2	0.531	0.230	0.243	0.651	0.707
PEER2	-0.330	-0.328	-0.267	-0.583	-0.633
SURF2	-0.279	-0.315	0.304	-0.045	-0.049
CHJUST2	-0.178	-0.439	0.326	-0.262	-0.284
CHEAT2	-0.367	-0.450	-0.115	-0.426	-0.462
TEACHER2	0.457	0.393	0.244	0.921	1.000

	CURUSE2	PEER2	SURF2	CHJUST2	CHEAT2
CURUSE2	1.000				
PEER2	-0.532	1.000			
SURF2	-0.127	0.099	1.000		
CHJUST2	-0.210	0.351	0.765	1.000	
CHEAT2	-0.423	0.577	0.390	0.786	1.000
TEACHER2	0.707	-0.633	-0.049	-0.284	-0.462

Table U2

Structural model: Estimated correlation matrix for the Time 2 male sample

	SUB2	HON2	PERF2	GTEACH2	ASSESS2
SUB2	1.000				
HON2	0.336	1.000			
PERF2	0.272	0.088	1.000		
GTEACH2	0.427	0.142	0.225	1.000	
ASSESS2	0.464	0.154	0.244	0.922	1.000
CURUSE2	0.532	0.177	0.243	0.652	0.707
PEER2	-0.386	-0.240	-0.285	-0.567	-0.615
SURF2	-0.290	-0.306	0.274	-0.001	-0.001
CHJUST2	-0.200	-0.346	0.287	-0.212	-0.230
CHEAT2	-0.391	-0.403	-0.159	-0.377	-0.409
TEACHER2	0.464	0.154	0.244	0.922	1.000

	CURUSE2	PEER2	SURF2	CHJUST2	CHEAT2
CURUSE2	1.000				
PEER2	-0.529	1.000			
SURF2	-0.115	0.112	1.000		
CHJUST2	-0.200	0.324	0.759	1.000	
CHEAT2	-0.414	0.561	0.389	0.777	1.000
TEACHER2	0.707	-0.615	-0.001	-0.230	-0.409

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level.

Appendix V: Female sample model correlation matrix, Time 2

Table V1

Estimated correlation matrix of the CFA for the Time 2 female sample

	SUB2	HON2	PERF2	GTEACH2	ASSESS2
SUB2	1.000				
HON2	0.245	1.000			
PERF2	0.022	-0.099	1.000		
GTEACH2	0.419	0.358	-0.015	1.000	
ASSESS2	0.502	0.429	-0.018	0.730	1.000
CURUSE2	0.543	0.230	0.061	0.517	0.619
PEER2	-0.066	-0.324	0.125	-0.318	-0.381
SURF2	-0.501	-0.226	0.218	-0.522	-0.626
CHJUST2	-0.368	-0.195	0.277	-0.455	-0.545
CHEAT2	-0.366	-0.252	0.188	-0.359	-0.430
TEACHER2	0.537	0.458	-0.019	0.781	0.935

	CURUSE2	PEER2	SURF2	CHJUST2	CHEAT2
CURUSE2	1.000				
PEER2	-0.283	1.000			
SURF2	-0.494	0.275	1.000		
CHJUST2	-0.407	0.566	0.655	1.000	
CHEAT2	-0.347	0.364	0.603	0.713	1.000
TEACHER2	0.662	-0.407	-0.669	-0.583	-0.460

Table V2

Estimated correlation matrix of the structural model for the Time 2 female sample

	SUB2	HON2	PERF2	GTEACH2	ASSESS2
SUB2	1.000				
HON2	0.246	1.000			
PERF2	0.025	0.001	1.000		
GTEACH2	0.421	0.324	0.007	1.000	
ASSESS2	0.498	0.384	0.008	0.732	1.000
CURUSE2	0.544	0.129	0.061	0.504	0.597
PEER2	-0.218	-0.296	0.100	-0.295	-0.349
SURF2	-0.519	-0.238	0.208	-0.510	-0.603
CHJUST2	-0.417	-0.140	0.270	-0.446	-0.528
CHEAT2	-0.396	-0.229	0.175	-0.349	-0.413
TEACHER2	0.535	0.412	0.009	0.787	0.931

	CURUSE2	PEER2	SURF2	CHJUST2	CHEAT2
CURUSE2	1.000				
PEER2	-0.256	1.000			
SURF2	-0.484	0.375	1.000		
CHJUST2	-0.398	0.566	0.643	1.000	
CHEAT2	-0.331	0.396	0.608	0.710	1.000
TEACHER2	0.641	-0.375	-0.648	-0.567	-0.444

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Experience of assessment; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; Grade = Grade-level.

Appendix W:

Full output for Model 4 co-ed data (*N* = 297)

estimated with observed indicator variables, Time 2

```
MODEL FIT INFORMATION

Number of Free Parameters                        199

Loglikelihood

          H0 Value                        -17300.886
          H0 Scaling Correction Factor        1.3399
            for MLR
          H1 Value                        -16224.500
          H1 Scaling Correction Factor        1.1412
            for MLR

Information Criteria

          Akaike (AIC)                     34999.772
          Bayesian (BIC)                   35734.824
          Sample-Size Adjusted BIC         35103.727
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                             1943.855*
          Degrees of Freedom                    1173
          P-Value                             0.0000
          Scaling Correction Factor           1.1075
            for MLR

*   The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used
    for chi-square difference testing in the regular way.  MLM, MLR and WLSM
    chi-square difference testing is described on the Mplus website.  MLMV, WLSMV,
    and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                             0.047
          90 Percent C.I.                      0.043  0.051
          Probability RMSEA <= .05             0.907

CFI/TLI

          CFI                                  0.894
          TLI                                  0.885

Chi-Square Test of Model Fit for the Baseline Model

          Value                             8537.683
          Degrees of Freedom                    1274
          P-Value                             0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                                0.069
```

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SUB2 BY				
SUB2_2	1.000	0.000	999.000	999.000
SUB3_2	0.994	0.044	22.708	0.000
SUB4_2	1.261	0.049	25.743	0.000
SUB10_2	1.014	0.049	20.721	0.000
SUB11_2	0.791	0.069	11.474	0.000
HON2 BY				
HON1_2	1.000	0.000	999.000	999.000
HON5_2	1.367	0.196	6.957	0.000
HON6_2	1.235	0.133	9.258	0.000
HON7_2	1.634	0.174	9.402	0.000
HON8_2	1.740	0.190	9.160	0.000
HON9_2	1.129	0.174	6.481	0.000
PERF2 BY				
PERF51_2	1.000	0.000	999.000	999.000
PERF59_2	1.486	0.174	8.544	0.000
PERF63_2	1.064	0.140	7.597	0.000
PERF64_2	1.227	0.177	6.929	0.000
CURUSE2 BY				
CURU15_2	1.000	0.000	999.000	999.000
CURU43_2	1.060	0.066	16.169	0.000
CURU46_2	1.045	0.069	15.180	0.000
CURU54_2	1.135	0.068	16.599	0.000
PEER2 BY				
PEER20_2	1.000	0.000	999.000	999.000
PEER25_2	1.323	0.136	9.703	0.000
PEER34_2	1.334	0.138	9.696	0.000
PEER48_2	1.248	0.146	8.567	0.000
PEER55_2	1.435	0.140	10.232	0.000
SURF2 BY				
SURF75_2	1.000	0.000	999.000	999.000
SURF76_2	0.646	0.116	5.553	0.000
SURF79_2	0.984	0.118	8.353	0.000
SURF87_2	1.041	0.100	10.387	0.000
CHJUST2 BY				
CHJU67_2	1.000	0.000	999.000	999.000
CHJU74_2	1.256	0.145	8.685	0.000
CHJU89_2	1.057	0.147	7.190	0.000
CHEAT2 BY				
CHEA72_2	1.000	0.000	999.000	999.000
CHEA80_2	1.408	0.161	8.771	0.000
CHEA84_2	1.242	0.151	8.215	0.000
ASSESS2 BY				
TRAN23_2	1.000	0.000	999.000	999.000
TRAN26_2	1.044	0.097	10.797	0.000
TRAN56_2	1.122	0.094	11.926	0.000
AUTH36_2	0.977	0.096	10.164	0.000
AUTH50_2	1.134	0.093	12.142	0.000
AUTH61_2	1.072	0.093	11.509	0.000
AUTH66_2	1.144	0.104	10.958	0.000
GTEACH2 BY				
GTEA14_2	1.000	0.000	999.000	999.000
GTEA27_2	1.128	0.105	10.713	0.000

```
    GTEA33_2           0.984      0.081     12.157      0.000
    GTEA41_2           1.136      0.091     12.508      0.000
    GTEA52_2           1.061      0.096     11.104      0.000
    GTEA57_2           1.103      0.109     10.095      0.000
    GTEA58_2           0.495      0.112      4.426      0.000
    GTEA65_2           1.013      0.096     10.594      0.000

 TEACHER2 BY
    GTEACH2            1.000      0.000    999.000    999.000
    ASSESS2            0.934      0.117      7.954      0.000

 CHEAT2   ON
    CURUSE2           -0.075      0.078     -0.963      0.336
    CHJUST2            0.708      0.155      4.584      0.000
    PERF2             -0.161      0.091     -1.776      0.076
    PEER2              0.098      0.100      0.986      0.324
    SUB2              -0.049      0.063     -0.787      0.431
    HON2              -0.079      0.113     -0.704      0.481
    TEACHER2           0.073      0.133      0.551      0.581

 SURF2    ON
    CURUSE2           -0.107      0.096     -1.115      0.265
    CHJUST2            0.678      0.159      4.264      0.000
    PERF2              0.111      0.092      1.213      0.225
    SUB2              -0.164      0.079     -2.080      0.038
    TEACHER2           0.079      0.175      0.450      0.653

 CHJUST2  ON
    SUB2              -0.107      0.067     -1.597      0.110
    CURUSE2           -0.007      0.092     -0.077      0.939
    PEER2              0.322      0.120      2.681      0.007
    PERF2              0.396      0.087      4.552      0.000
    TEACHER2          -0.260      0.146     -1.781      0.075
    HON2              -0.002      0.122     -0.013      0.990

 PEER2    ON
    CURUSE2           -0.077      0.072     -1.068      0.286
    PERF2             -0.037      0.062     -0.602      0.547
    TEACHER2          -0.428      0.123     -3.484      0.000
    HON2              -0.127      0.095     -1.328      0.184

 PERF2    ON
    SUB2               0.082      0.049      1.656      0.098

 TEACHER2 ON
    SUB2               0.240      0.047      5.150      0.000
    HON2               0.394      0.087      4.513      0.000

 CURUSE2  ON
    SUB2               0.489      0.071      6.884      0.000

 CHEAT2   ON
    GENDER             0.093      0.080      1.167      0.243
    GRADE              0.062      0.069      0.890      0.373

 SURF2    ON
    GENDER            -0.005      0.111     -0.046      0.963
    GRADE              0.127      0.087      1.459      0.145

 CHJUST2  ON
    GENDER             0.210      0.091      2.313      0.021
    GRADE              0.061      0.075      0.818      0.413

 PEER2    ON
    GENDER             0.306      0.080      3.836      0.000
    GRADE             -0.196      0.076     -2.570      0.010
```

```
PERF2    ON
    GENDER             0.355      0.106      3.367      0.001
    GRADE             -0.184      0.084     -2.184      0.029

 TEACHER2 ON
    GENDER             0.025      0.074      0.333      0.739
    GRADE             -0.214      0.075     -2.843      0.004

 CURUSE2  ON
    GENDER            -0.016      0.094     -0.171      0.864
    GRADE             -0.180      0.090     -1.988      0.047

 HON2     ON
    GENDER            -0.027      0.059     -0.453      0.651
    GRADE              0.012      0.058      0.210      0.834

 SUB2     ON
    GENDER             0.461      0.113      4.084      0.000
    GRADE             -0.183      0.110     -1.667      0.096

 SUB2     WITH
    HON2               0.126      0.036      3.457      0.001

 PERF2    WITH
    TEACHER2           0.022      0.026      0.839      0.401
    CURUSE2            0.025      0.033      0.762      0.446

 TEACHER2 WITH
    CURUSE2            0.196      0.040      4.930      0.000

 SURF2    WITH
    CHEAT2            -0.015      0.033     -0.470      0.638

 SURF87_2 WITH
    SURF79_2           0.219      0.084      2.612      0.009

 Intercepts
    SUB2_2             1.817      0.243      7.481      0.000
    SUB3_2             1.651      0.237      6.968      0.000
    SUB4_2             1.960      0.306      6.408      0.000
    SUB10_2            2.039      0.248      8.236      0.000
    SUB11_2            1.602      0.190      8.412      0.000
    HON1_2             1.840      0.134     13.717      0.000
    HON5_2             2.438      0.191     12.783      0.000
    HON6_2             1.909      0.162     11.751      0.000
    HON7_2             2.040      0.217      9.413      0.000
    HON8_2             1.838      0.232      7.908      0.000
    HON9_2             2.415      0.162     14.945      0.000
    PERF51_2           3.124      0.212     14.709      0.000
    PERF59_2           2.866      0.289      9.929      0.000
    PERF63_2           3.130      0.215     14.562      0.000
    PERF64_2           3.465      0.244     14.207      0.000
    GTEA14_2           2.844      0.181     15.734      0.000
    GTEA27_2           2.743      0.202     13.558      0.000
    GTEA33_2           2.687      0.179     14.979      0.000
    GTEA41_2           2.660      0.197     13.469      0.000
    GTEA52_2           2.508      0.188     13.364      0.000
    GTEA57_2           2.519      0.192     13.136      0.000
    GTEA58_2           2.378      0.104     22.966      0.000
    GTEA65_2           2.631      0.177     14.855      0.000
    CURU15_2           2.459      0.210     11.692      0.000
    CURU43_2           2.554      0.221     11.542      0.000
    CURU46_2           2.459      0.218     11.283      0.000
    CURU54_2           2.545      0.237     10.722      0.000
    TRAN23_2           2.132      0.171     12.473      0.000
    TRAN26_2           2.325      0.176     13.218      0.000
    TRAN56_2           2.365      0.189     12.501      0.000
```

```
    AUTH36_2           2.260      0.163     13.844      0.000
    AUTH50_2           2.373      0.184     12.878      0.000
    AUTH61_2           2.323      0.175     13.283      0.000
    AUTH66_2           2.496      0.187     13.334      0.000
    PEER20_2           3.758      0.178     21.087      0.000
    PEER25_2           3.232      0.228     14.171      0.000
    PEER34_2           3.084      0.226     13.646      0.000
    PEER48_2           3.273      0.214     15.317      0.000
    PEER55_2           3.370      0.240     14.035      0.000
    SURF75_2           3.388      0.257     13.182      0.000
    SURF76_2           2.507      0.172     14.559      0.000
    SURF79_2           2.807      0.250     11.245      0.000
    SURF87_2           2.997      0.269     11.141      0.000
    CHJU67_2           3.567      0.211     16.926      0.000
    CHJU74_2           3.125      0.250     12.519      0.000
    CHJU89_2           3.460      0.231     14.977      0.000
    CHEA72_2           3.726      0.213     17.465      0.000
    CHEA80_2           3.382      0.285     11.864      0.000
    CHEA84_2           3.357      0.252     13.311      0.000

 Residual Variances
    SUB2_2             0.248      0.057      4.349      0.000
    SUB3_2             0.213      0.028      7.665      0.000
    SUB4_2             0.406      0.057      7.167      0.000
    SUB10_2            0.304      0.039      7.877      0.000
    SUB11_2            0.644      0.107      6.026      0.000
    HON1_2             0.395      0.063      6.301      0.000
    HON5_2             0.415      0.050      8.255      0.000
    HON6_2             0.403      0.049      8.301      0.000
    HON7_2             0.159      0.024      6.622      0.000
    HON8_2             0.283      0.039      7.348      0.000
    HON9_2             0.737      0.082      9.037      0.000
    PERF51_2           0.819      0.089      9.176      0.000
    PERF59_2           0.465      0.092      5.069      0.000
    PERF63_2           0.801      0.115      6.987      0.000
    PERF64_2           0.677      0.097      6.985      0.000
    GTEA14_2           0.717      0.071     10.044      0.000
    GTEA27_2           0.762      0.107      7.148      0.000
    GTEA33_2           0.473      0.057      8.370      0.000
    GTEA41_2           0.349      0.050      6.996      0.000
    GTEA52_2           0.422      0.049      8.569      0.000
    GTEA57_2           0.562      0.081      6.944      0.000
    GTEA58_2           1.101      0.114      9.635      0.000
    GTEA65_2           0.405      0.045      9.079      0.000
    CURU15_2           0.453      0.066      6.899      0.000
    CURU43_2           0.229      0.033      6.901      0.000
    CURU46_2           0.261      0.038      6.952      0.000
    CURU54_2           0.179      0.039      4.556      0.000
    TRAN23_2           0.371      0.044      8.490      0.000
    TRAN26_2           0.435      0.059      7.385      0.000
    TRAN56_2           0.448      0.051      8.840      0.000
    AUTH36_2           0.376      0.045      8.341      0.000
    AUTH50_2           0.314      0.046      6.778      0.000
    AUTH61_2           0.295      0.039      7.535      0.000
    AUTH66_2           0.341      0.040      8.432      0.000
    PEER20_2           0.622      0.087      7.110      0.000
    PEER25_2           0.549      0.079      6.919      0.000
    PEER34_2           0.759      0.108      7.007      0.000
    PEER48_2           0.669      0.113      5.927      0.000
    PEER55_2           0.219      0.070      3.129      0.002
    SURF75_2           0.462      0.079      5.863      0.000
    SURF76_2           1.217      0.106     11.530      0.000
    SURF79_2           0.774      0.100      7.731      0.000
    SURF87_2           0.595      0.090      6.592      0.000
    CHJU67_2           0.376      0.055      6.781      0.000
    CHJU74_2           0.556      0.099      5.590      0.000
    CHJU89_2           0.625      0.101      6.190      0.000
    CHEA72_2           0.447      0.093      4.803      0.000
```

```
    CHEA80_2           0.170      0.052      3.258      0.001
    CHEA84_2           0.459      0.076      6.041      0.000
    SUB2               0.806      0.078     10.357      0.000
    HON2               0.220      0.053      4.196      0.000
    PERF2              0.401      0.083      4.839      0.000
    CURUSE2            0.534      0.067      7.954      0.000
    PEER2              0.273      0.052      5.235      0.000
    SURF2              0.311      0.083      3.736      0.000
    CHJUST2            0.271      0.055      4.957      0.000
    CHEAT2             0.213      0.051      4.155      0.000
    ASSESS2            0.029      0.024      1.225      0.221
    GTEACH2            0.176      0.040      4.422      0.000
    TEACHER2           0.241      0.046      5.242      0.000

STANDARDIZED MODEL RESULTS

STDYX Standardization

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 SUB2     BY
    SUB2_2             0.882      0.028     31.788      0.000
    SUB3_2             0.895      0.016     54.621      0.000
    SUB4_2             0.879      0.018     49.495      0.000
    SUB10_2            0.863      0.021     41.938      0.000
    SUB11_2            0.676      0.054     12.626      0.000

 HON2     BY
    HON1_2             0.599      0.061      9.755      0.000
    HON5_2             0.706      0.035     19.953      0.000
    HON6_2             0.675      0.053     12.725      0.000
    HON7_2             0.887      0.023     38.946      0.000
    HON8_2             0.838      0.029     28.912      0.000
    HON9_2             0.525      0.056      9.300      0.000

 PERF2    BY
    PERF51_2           0.597      0.056     10.575      0.000
    PERF59_2           0.826      0.037     22.181      0.000
    PERF63_2           0.625      0.063      9.986      0.000
    PERF64_2           0.708      0.047     14.974      0.000

 CURUSE2  BY
    CURU15_2           0.791      0.034     23.265      0.000
    CURU43_2           0.888      0.018     48.135      0.000
    CURU46_2           0.872      0.021     41.852      0.000
    CURU54_2           0.919      0.019     49.662      0.000

 PEER2    BY
    PEER20_2           0.623      0.054     11.461      0.000
    PEER25_2           0.746      0.040     18.428      0.000
    PEER34_2           0.693      0.047     14.857      0.000
    PEER48_2           0.692      0.054     12.753      0.000
    PEER55_2           0.888      0.037     24.280      0.000

 SURF2    BY
    SURF75_2           0.760      0.049     15.574      0.000
    SURF76_2           0.422      0.070      5.985      0.000
    SURF79_2           0.665      0.054     12.279      0.000
    SURF87_2           0.731      0.050     14.611      0.000

 CHJUST2  BY
    CHJU67_2           0.746      0.043     17.426      0.000
    CHJU74_2           0.756      0.046     16.301      0.000
    CHJU89_2           0.676      0.055     12.258      0.000
 CHEAT2   BY
    CHEA72_2           0.727      0.060     12.137      0.000
    CHEA80_2           0.924      0.025     36.729      0.000
    CHEA84_2           0.792      0.039     20.144      0.000
```

```
 ASSESS2  BY
    TRAN23_2           0.696      0.039     18.018      0.000
    TRAN26_2           0.683      0.046     14.778      0.000
    TRAN56_2           0.703      0.039     17.854      0.000
    AUTH36_2           0.685      0.042     16.361      0.000
    AUTH50_2           0.767      0.035     21.766      0.000
    AUTH61_2           0.759      0.036     20.835      0.000
    AUTH66_2           0.757      0.032     24.002      0.000

 GTEACH2  BY
    GTEA14_2           0.656      0.042     15.650      0.000
    GTEA27_2           0.689      0.049     14.136      0.000
    GTEA33_2           0.725      0.037     19.647      0.000
    GTEA41_2           0.817      0.029     28.485      0.000
    GTEA52_2           0.769      0.031     24.662      0.000
    GTEA57_2           0.735      0.039     18.635      0.000
    GTEA58_2           0.328      0.071      4.588      0.000
    GTEA65_2           0.760      0.030     25.296      0.000

 TEACHER2 BY
    GTEACH2            0.822      0.035     23.263      0.000
    ASSESS2            0.957      0.035     27.065      0.000

 CHEAT2   ON
    CURUSE2           -0.093      0.098     -0.946      0.344
    CHJUST2            0.687      0.121      5.693      0.000
    PERF2             -0.153      0.086     -1.779      0.075
    PEER2              0.087      0.090      0.969      0.332
    SUB2              -0.065      0.083     -0.780      0.436
    HON2              -0.053      0.075     -0.704      0.482
    TEACHER2           0.063      0.114      0.551      0.581

 SURF2    ON
    CURUSE2           -0.116      0.102     -1.145      0.252
    CHJUST2            0.586      0.114      5.155      0.000
    PERF2              0.094      0.080      1.179      0.238
    SUB2              -0.192      0.091     -2.103      0.036
    TEACHER2           0.060      0.133      0.453      0.651

 CHJUST2  ON
    SUB2              -0.144      0.089     -1.631      0.103
    CURUSE2           -0.009      0.116     -0.077      0.939
    PEER2              0.295      0.091      3.245      0.001
    PERF2              0.388      0.085      4.591      0.000
    TEACHER2          -0.229      0.129     -1.769      0.077
    HON2              -0.001      0.084     -0.013      0.990

 PEER2    ON
    CURUSE2           -0.107      0.100     -1.076      0.282
    PERF2             -0.040      0.066     -0.601      0.548
    TEACHER2          -0.412      0.105     -3.914      0.000
    HON2              -0.095      0.071     -1.336      0.181

 PERF2    ON
    SUB2               0.113      0.067      1.683      0.092

 TEACHER2 ON
    SUB2               0.369      0.064      5.801      0.000
    HON2               0.306      0.060      5.084      0.000

 CURUSE2  ON
    SUB2               0.523      0.059      8.886      0.000

 CHEAT2   ON
    GENDER             0.064      0.054      1.179      0.239
    GRADE              0.044      0.049      0.889      0.374
```

```
 SURF2    ON
    GENDER            -0.003      0.068     -0.046      0.963
    GRADE              0.080      0.055      1.464      0.143

 CHJUST2  ON
    GENDER             0.149      0.064      2.315      0.021
    GRADE              0.045      0.055      0.814      0.416

 PEER2    ON
    GENDER             0.238      0.055      4.318      0.000
    GRADE             -0.156      0.055     -2.823      0.005

 PERF2    ON
    GENDER             0.257      0.064      4.026      0.000
    GRADE             -0.136      0.060     -2.260      0.024

 TEACHER2 ON
    GENDER             0.020      0.060      0.335      0.737
    GRADE             -0.177      0.057     -3.136      0.002

 CURUSE2  ON
    GENDER            -0.009      0.053     -0.171      0.864
    GRADE             -0.103      0.052     -1.977      0.048

 HON2     ON
    GENDER            -0.028      0.062     -0.445      0.656
    GRADE              0.013      0.062      0.210      0.834

 SUB2     ON
    GENDER             0.241      0.057      4.236      0.000
    GRADE             -0.099      0.059     -1.673      0.094

 SUB2     WITH
    HON2               0.299      0.063      4.713      0.000

 PERF2    WITH
    TEACHER2           0.071      0.082      0.867      0.386
    CURUSE2            0.055      0.071      0.765      0.444

 TEACHER2 WITH
    CURUSE2            0.546      0.068      8.064      0.000

 SURF2    WITH
    CHEAT2            -0.060      0.127     -0.469      0.639

 SURF87_2 WITH
    SURF79_2           0.322      0.094      3.442      0.001

 Intercepts
    SUB2_2             1.722      0.239      7.192      0.000
    SUB3_2             1.599      0.245      6.538      0.000
    SUB4_2             1.468      0.232      6.323      0.000
    SUB10_2            1.867      0.243      7.672      0.000
    SUB11_2            1.472      0.190      7.735      0.000
    HON1_2             2.345      0.214     10.948      0.000
    HON5_2             2.681      0.221     12.145      0.000
    HON6_2             2.220      0.224      9.916      0.000
    HON7_2             2.360      0.275      8.597      0.000
    HON8_2             1.884      0.248      7.607      0.000
    HON9_2             2.393      0.165     14.498      0.000
    PERF51_2           2.770      0.231     11.995      0.000
    PERF59_2           2.369      0.257      9.218      0.000
    PERF63_2           2.730      0.217     12.577      0.000
    PERF64_2           2.973      0.257     11.546      0.000
    GTEA14_2           2.535      0.164     15.506      0.000
    GTEA27_2           2.277      0.170     13.397      0.000
    GTEA33_2           2.691      0.184     14.635      0.000
    GTEA41_2           2.599      0.202     12.882      0.000
```

GTEA52_2	2.470	0.185	13.353	0.000
GTEA57_2	2.281	0.181	12.618	0.000
GTEA58_2	2.140	0.103	20.845	0.000
GTEA65_2	2.684	0.193	13.939	0.000
CURU15_2	2.237	0.194	11.505	0.000
CURU43_2	2.460	0.226	10.871	0.000
CURU46_2	2.359	0.222	10.627	0.000
CURU54_2	2.372	0.230	10.331	0.000
TRAN23_2	2.513	0.182	13.847	0.000
TRAN26_2	2.575	0.194	13.296	0.000
TRAN56_2	2.512	0.194	12.933	0.000
AUTH36_2	2.685	0.196	13.706	0.000
AUTH50_2	2.718	0.218	12.479	0.000
AUTH61_2	2.785	0.205	13.597	0.000
AUTH66_2	2.797	0.218	12.844	0.000
PEER20_2	3.728	0.266	14.002	0.000
PEER25_2	2.903	0.241	12.043	0.000
PEER34_2	2.552	0.214	11.918	0.000
PEER48_2	2.889	0.219	13.208	0.000
PEER55_2	3.319	0.273	12.149	0.000
SURF75_2	3.238	0.331	9.783	0.000
SURF76_2	2.060	0.159	12.941	0.000
SURF79_2	2.384	0.240	9.946	0.000
SURF87_2	2.650	0.286	9.263	0.000
CHJU67_2	3.877	0.363	10.685	0.000
CHJU74_2	2.742	0.271	10.127	0.000
CHJU89_2	3.224	0.305	10.572	0.000
CHEA72_2	3.829	0.370	10.338	0.000
CHEA80_2	3.138	0.358	8.757	0.000
CHEA84_2	3.026	0.298	10.144	0.000
Residual Variances				
SUB2_2	0.223	0.049	4.555	0.000
SUB3_2	0.199	0.029	6.794	0.000
SUB4_2	0.228	0.031	7.293	0.000
SUB10_2	0.255	0.036	7.159	0.000
SUB11_2	0.543	0.072	7.510	0.000
HON1_2	0.642	0.073	8.738	0.000
HON5_2	0.502	0.050	10.038	0.000
HON6_2	0.545	0.072	7.622	0.000
HON7_2	0.213	0.040	5.258	0.000
HON8_2	0.298	0.049	6.131	0.000
HON9_2	0.724	0.059	12.205	0.000
PERF51_2	0.644	0.067	9.566	0.000
PERF59_2	0.317	0.062	5.154	0.000
PERF63_2	0.610	0.078	7.794	0.000
PERF64_2	0.498	0.067	7.437	0.000
GTEA14_2	0.570	0.055	10.367	0.000
GTEA27_2	0.525	0.067	7.813	0.000
GTEA33_2	0.474	0.053	8.869	0.000
GTEA41_2	0.333	0.047	7.113	0.000
GTEA52_2	0.409	0.048	8.534	0.000
GTEA57_2	0.460	0.058	7.949	0.000
GTEA58_2	0.892	0.047	19.034	0.000
GTEA65_2	0.422	0.046	9.226	0.000
CURU15_2	0.375	0.054	6.979	0.000
CURU43_2	0.212	0.033	6.481	0.000
CURU46_2	0.240	0.036	6.624	0.000
CURU54_2	0.155	0.034	4.563	0.000
TRAN23_2	0.516	0.054	9.607	0.000
TRAN26_2	0.534	0.063	8.469	0.000
TRAN56_2	0.506	0.055	9.128	0.000
AUTH36_2	0.530	0.057	9.236	0.000
AUTH50_2	0.412	0.054	7.630	0.000
AUTH61_2	0.425	0.055	7.688	0.000
AUTH66_2	0.427	0.048	8.959	0.000
PEER20_2	0.612	0.068	9.043	0.000
PEER25_2	0.443	0.060	7.329	0.000

PEER34_2	0.520	0.065	8.035	0.000
PEER48_2	0.522	0.075	6.955	0.000
PEER55_2	0.212	0.065	3.270	0.001
SURF75_2	0.422	0.074	5.694	0.000
SURF76_2	0.822	0.059	13.824	0.000
SURF79_2	0.558	0.072	7.762	0.000
SURF87_2	0.465	0.073	6.349	0.000
CHJU67_2	0.444	0.064	6.952	0.000
CHJU74_2	0.428	0.070	6.097	0.000
CHJU89_2	0.543	0.075	7.284	0.000
CHEA72_2	0.472	0.087	5.424	0.000
CHEA80_2	0.147	0.046	3.154	0.002
CHEA84_2	0.373	0.062	5.993	0.000
SUB2	0.932	0.030	30.772	0.000
HON2	0.999	0.004	265.150	0.000
PERF2	0.885	0.039	22.705	0.000
CURUSE2	0.707	0.061	11.564	0.000
PEER2	0.693	0.057	12.067	0.000
SURF2	0.491	0.089	5.527	0.000
CHJUST2	0.575	0.075	7.664	0.000
CHEAT2	0.426	0.085	5.045	0.000
ASSESS2	0.084	0.068	1.244	0.213
GTEACH2	0.325	0.058	5.596	0.000
TEACHER2	0.660	0.058	11.328	0.000

R-SQUARE

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SUB2_2	0.777	0.049	15.894	0.000
SUB3_2	0.801	0.029	27.310	0.000
SUB4_2	0.772	0.031	24.748	0.000
SUB10_2	0.745	0.036	20.969	0.000
SUB11_2	0.457	0.072	6.313	0.000
HON1_2	0.358	0.073	4.878	0.000
HON5_2	0.498	0.050	9.977	0.000
HON6_2	0.455	0.072	6.363	0.000
HON7_2	0.787	0.040	19.473	0.000
HON8_2	0.702	0.049	14.456	0.000
HON9_2	0.276	0.059	4.650	0.000
PERF51_2	0.356	0.067	5.287	0.000
PERF59_2	0.683	0.062	11.090	0.000
PERF63_2	0.390	0.078	4.993	0.000
PERF64_2	0.502	0.067	7.487	0.000
GTEA14_2	0.430	0.055	7.825	0.000
GTEA27_2	0.475	0.067	7.068	0.000
GTEA33_2	0.526	0.053	9.823	0.000
GTEA41_2	0.667	0.047	14.243	0.000
GTEA52_2	0.591	0.048	12.331	0.000
GTEA57_2	0.540	0.058	9.318	0.000
GTEA58_2	0.108	0.047	2.294	0.022
GTEA65_2	0.578	0.046	12.648	0.000
CURU15_2	0.625	0.054	11.633	0.000
CURU43_2	0.788	0.033	24.068	0.000
CURU46_2	0.760	0.036	20.926	0.000
CURU54_2	0.845	0.034	24.831	0.000
TRAN23_2	0.484	0.054	9.009	0.000
TRAN26_2	0.466	0.063	7.389	0.000
TRAN56_2	0.494	0.055	8.927	0.000
AUTH36_2	0.470	0.057	8.180	0.000
AUTH50_2	0.588	0.054	10.883	0.000
AUTH61_2	0.575	0.055	10.418	0.000
AUTH66_2	0.573	0.048	12.001	0.000
PEER20_2	0.388	0.068	5.731	0.000
PEER25_2	0.557	0.060	9.214	0.000
PEER34_2	0.480	0.065	7.428	0.000

```
PEER48_2            0.478      0.075      6.377      0.000
PEER55_2            0.788      0.065     12.140      0.000
SURF75_2            0.578      0.074      7.787      0.000
SURF76_2            0.178      0.059      2.993      0.003
SURF79_2            0.442      0.072      6.140      0.000
SURF87_2            0.535      0.073      7.305      0.000
CHJU67_2            0.556      0.064      8.713      0.000
CHJU74_2            0.572      0.070      8.151      0.000
CHJU89_2            0.457      0.075      6.129      0.000
CHEA72_2            0.528      0.087      6.069      0.000
CHEA80_2            0.853      0.046     18.364      0.000
CHEA84_2            0.627      0.062     10.072      0.000

 Latent                                          Two-Tailed
Variable        Estimate       S.E.  Est./S.E.    P-Value

SUB2                0.068      0.030      2.248      0.025
HON2                0.001      0.004      0.247      0.805
PERF2               0.115      0.039      2.940      0.003
CURUSE2             0.293      0.061      4.782      0.000
PEER2               0.307      0.057      5.355      0.000
SURF2               0.509      0.089      5.723      0.000
CHJUST2             0.425      0.075      5.673      0.000
CHEAT2              0.574      0.085      6.786      0.000
ASSESS2             0.916      0.068     13.532      0.000
GTEACH2             0.675      0.058     11.632      0.000
TEACHER2            0.340      0.058      5.842      0.000

 ESTIMATES DERIVED FROM THE MODEL


      ESTIMATED MEANS FOR THE LATENT VARIABLES
         SUB2          HON2          PERF2         CURUSE2       PEER2
         ________      ________      ________      ________      ________
 1          0.469        -0.025         0.337        -0.065         0.273


      ESTIMATED MEANS FOR THE LATENT VARIABLES
         SURF2         CHJUST2       CHEAT2        ASSESS2       GTEACH2
         ________      ________      ________      ________      ________
 1          0.575         0.648         0.645        -0.166        -0.178


      ESTIMATED MEANS FOR THE LATENT VARIABLES
         TEACHER2      GENDER        GRADE
         ________      ________      ________
 1         -0.178         1.613         1.495


      S.E. FOR ESTIMATED MEANS FOR THE LATENT VARIABLES
         SUB2          HON2          PERF2         CURUSE2       PEER2
         ________      ________      ________      ________      ________
 1          0.240         0.129         0.199         0.204         0.168


      S.E. FOR ESTIMATED MEANS FOR THE LATENT VARIABLES
         SURF2         CHJUST2       CHEAT2        ASSESS2       GTEACH2
         ________      ________      ________      ________      ________
 1          0.242         0.195         0.192         0.154         0.163


      S.E. FOR ESTIMATED MEANS FOR THE LATENT VARIABLES
         TEACHER2      GENDER        GRADE
         ________      ________      ________
 1          0.163         0.028         0.029
```

ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	SUB2	HON2	PERF2	CURUSE2	PEER2
SUB2	0.865				
HON2	0.122	0.221			
PERF2	0.118	0.007	0.453		
CURUSE2	0.429	0.059	0.090	0.755	
PEER2	-0.125	-0.085	-0.015	-0.185	0.394
SURF2	-0.254	-0.063	0.129	-0.229	0.153
CHJUST2	-0.139	-0.069	0.161	-0.157	0.192
CHEAT2	-0.187	-0.079	0.037	-0.201	0.195
ASSESS2	0.251	0.108	0.062	0.316	-0.166
GTEACH2	0.269	0.115	0.066	0.338	-0.178
TEACHER2	0.269	0.115	0.066	0.338	-0.178
GENDER	0.109	-0.006	0.093	0.050	0.054
GRADE	-0.046	0.003	-0.050	-0.067	-0.015

ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	SURF2	CHJUST2	CHEAT2	ASSESS2	GTEACH2
SURF2	0.632				
CHJUST2	0.366	0.471			
CHEAT2	0.268	0.348	0.500		
ASSESS2	-0.147	-0.145	-0.149	0.348	
GTEACH2	-0.157	-0.155	-0.160	0.341	0.541
TEACHER2	-0.157	-0.155	-0.160	0.341	0.366
GENDER	0.045	0.084	0.066	0.028	0.030
GRADE	0.044	0.013	0.033	-0.059	-0.063

ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	TEACHER2	GENDER	GRADE
TEACHER2	0.366		
GENDER	0.030	0.237	
GRADE	-0.063	0.000	0.250

S.E. FOR ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	SUB2	HON2	PERF2	CURUSE2	PEER2
SUB2	0.081				
HON2	0.038	0.052			
PERF2	0.045	0.009	0.098		
CURUSE2	0.070	0.020	0.044	0.097	
PEER2	0.031	0.026	0.033	0.051	0.078
SURF2	0.061	0.020	0.059	0.060	0.030
CHJUST2	0.047	0.025	0.053	0.046	0.045
CHEAT2	0.052	0.030	0.038	0.051	0.044
ASSESS2	0.038	0.025	0.032	0.052	0.033
GTEACH2	0.051	0.028	0.036	0.057	0.033
TEACHER2	0.051	0.028	0.036	0.057	0.033
GENDER	0.027	0.014	0.025	0.026	0.020
GRADE	0.028	0.014	0.022	0.026	0.020

S.E. FOR ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	SURF2	CHJUST2	CHEAT2	ASSESS2	GTEACH2
SURF2	0.115				
CHJUST2	0.059	0.086			
CHEAT2	0.054	0.071	0.104		
ASSESS2	0.045	0.036	0.032	0.056	
GTEACH2	0.049	0.039	0.036	0.046	0.087
TEACHER2	0.049	0.039	0.036	0.046	0.070
GENDER	0.028	0.024	0.023	0.018	0.021
GRADE	0.027	0.023	0.021	0.018	0.022

S.E. FOR ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	TEACHER2	GENDER	GRADE
TEACHER2	0.070		
GENDER	0.021	0.006	
GRADE	0.022	0.014	0.000

ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	SUB2	HON2	PERF2	CURUSE2	PEER2
SUB2	1.000				
HON2	0.280	1.000			
PERF2	0.188	0.023	1.000		
CURUSE2	0.531	0.145	0.154	1.000	
PEER2	-0.215	-0.287	-0.035	-0.340	1.000
SURF2	-0.343	-0.169	0.241	-0.332	0.306
CHJUST2	-0.217	-0.215	0.348	-0.263	0.447
CHEAT2	-0.285	-0.237	0.077	-0.327	0.440
ASSESS2	0.457	0.389	0.156	0.616	-0.449
GTEACH2	0.392	0.334	0.134	0.529	-0.385
TEACHER2	0.478	0.407	0.163	0.643	-0.469
GENDER	0.242	-0.028	0.285	0.117	0.175
GRADE	-0.099	0.013	-0.148	-0.155	-0.049

ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	SURF2	CHJUST2	CHEAT2	ASSESS2	GTEACH2
SURF2	1.000				
CHJUST2	0.670	1.000			
CHEAT2	0.477	0.717	1.000		
ASSESS2	-0.313	-0.357	-0.358	1.000	
GTEACH2	-0.269	-0.307	-0.308	0.786	1.000
TEACHER2	-0.327	-0.373	-0.374	0.957	0.822
GENDER	0.117	0.252	0.190	0.097	0.083
GRADE	0.112	0.036	0.094	-0.201	-0.173

ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	TEACHER2	GENDER	GRADE
TEACHER2	1.000		
GENDER	0.101	1.000	
GRADE	-0.210	-0.001	1.000

S.E. FOR ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	SUB2	HON2	PERF2	CURUSE2	PEER2
SUB2	0.000				
HON2	0.066	0.000			
PERF2	0.063	0.028	0.000		
CURUSE2	0.057	0.040	0.069	0.000	
PEER2	0.044	0.062	0.078	0.069	0.000
SURF2	0.076	0.052	0.092	0.081	0.059
CHJUST2	0.070	0.070	0.084	0.074	0.084
CHEAT2	0.068	0.078	0.078	0.067	0.062
ASSESS2	0.058	0.055	0.073	0.055	0.055
GTEACH2	0.054	0.051	0.067	0.050	0.051
TEACHER2	0.060	0.059	0.078	0.053	0.057
GENDER	0.057	0.062	0.061	0.060	0.060
GRADE	0.061	0.062	0.062	0.059	0.061

S.E. FOR ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	SURF2	CHJUST2	CHEAT2	ASSESS2	GTEACH2
SURF2	0.000				
CHJUST2	0.067	0.000			
CHEAT2	0.073	0.065	0.000		
ASSESS2	0.090	0.081	0.059	0.000	
GTEACH2	0.079	0.069	0.052	0.033	0.000
TEACHER2	0.094	0.083	0.061	0.035	0.035
GENDER	0.071	0.066	0.061	0.064	0.057
GRADE	0.067	0.066	0.061	0.059	0.055

S.E. FOR ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	TEACHER2	GENDER	GRADE
TEACHER2	0.000		
GENDER	0.068	0.000	
GRADE	0.064	0.058	0.000

Appendix X:

Indirect effects (standardized) in Model 4 for the Co-ed sample, Time 2

```
STANDARDIZED TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS

STDYX Standardization

                                                        Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Effects from CURUSE2 to CHEAT2

  Total               -0.130      0.096     -1.358      0.175
  Total indirect      -0.037      0.082     -0.451      0.652

  Specific indirect

    CHEAT2
    PEER2
    CURUSE2           -0.009      0.014     -0.669      0.503

    CHEAT2
    CHJUST2
    CURUSE2           -0.006      0.080     -0.077      0.938

    CHEAT2
    CHJUST2
    PEER2
    CURUSE2           -0.022      0.023     -0.963      0.335

  Direct
    CHEAT2
    CURUSE2           -0.093      0.098     -0.946      0.344

Effects from PERF2 to CHEAT2

  Total                0.102      0.074      1.384      0.166
  Total indirect       0.255      0.082      3.099      0.002

  Specific indirect

    CHEAT2
    PEER2
    PERF2             -0.003      0.007     -0.503      0.615

    CHEAT2
    CHJUST2
    PERF2              0.267      0.081      3.310      0.001

    CHEAT2
    CHJUST2
    PEER2
    PERF2             -0.008      0.014     -0.582      0.561

  Direct
    CHEAT2
    PERF2             -0.153      0.086     -1.779      0.075
```

```
Effects from PEER2 to CHEAT2

  Total                 0.290      0.074      3.901      0.000
  Total indirect        0.203      0.080      2.530      0.011

  Specific indirect

    CHEAT2
    CHJUST2
    PEER2               0.203      0.080      2.530      0.011

  Direct
    CHEAT2
    PEER2               0.087      0.090      0.969      0.332

Effects from SUB2 to CHEAT2

  Total                -0.299      0.072     -4.137      0.000
  Total indirect       -0.235      0.066     -3.543      0.000

  Specific indirect

    CHEAT2
    PERF2
    SUB2               -0.017      0.014     -1.233      0.218

    CHEAT2
    CURUSE2
    SUB2               -0.048      0.053     -0.918      0.359

    CHEAT2
    CHJUST2
    SUB2               -0.099      0.065     -1.525      0.127

    CHEAT2
    TEACHER2
    SUB2                0.023      0.042      0.549      0.583

    CHEAT2
    PEER2
    PERF2
    SUB2                0.000      0.001     -0.475      0.634

    CHEAT2
    PEER2
    CURUSE2
    SUB2               -0.005      0.007     -0.666      0.505

    CHEAT2
    PEER2
    TEACHER2
    SUB2               -0.013      0.014     -0.938      0.348

    CHEAT2
    CHJUST2
    PERF2
    SUB2                0.030      0.020      1.529      0.126

    CHEAT2
    CHJUST2
    CURUSE2
    SUB2               -0.003      0.042     -0.077      0.938

    CHEAT2
    CHJUST2
    TEACHER2
    SUB2               -0.058      0.037     -1.559      0.119
```

```
    CHEAT2
    CHJUST2
    PEER2
    PERF2
    SUB2                -0.001      0.002     -0.562      0.574

    CHEAT2
    CHJUST2
    PEER2
    CURUSE2
    SUB2                -0.011      0.012     -0.946      0.344

    CHEAT2
    CHJUST2
    PEER2
    TEACHER2
    SUB2                -0.031      0.014     -2.186      0.029

  Direct
    CHEAT2
    SUB2                -0.065      0.083     -0.780      0.436


Effects from HON2 to CHEAT2

  Total                 -0.146      0.084     -1.746      0.081
  Total indirect        -0.094      0.060     -1.553      0.120

  Specific indirect

    CHEAT2
    PEER2
    HON2                -0.008      0.010     -0.823      0.411

    CHEAT2
    CHJUST2
    HON2                -0.001      0.057     -0.013      0.990

    CHEAT2
    TEACHER2
    HON2                 0.019      0.035      0.549      0.583

    CHEAT2
    PEER2
    TEACHER2
    HON2                -0.011      0.012     -0.916      0.360

    CHEAT2
    CHJUST2
    PEER2
    HON2                -0.019      0.016     -1.215      0.224

    CHEAT2
    CHJUST2
    TEACHER2
    HON2                -0.048      0.030     -1.587      0.113

    CHEAT2
    CHJUST2
    PEER2
    TEACHER2
    HON2                -0.026      0.013     -2.003      0.045

  Direct
    CHEAT2
    HON2                -0.053      0.075     -0.704      0.482
```

```
Effects from TEACHER2 to CHEAT2

  Total                -0.214      0.108     -1.977      0.048
  Total indirect       -0.277      0.099     -2.784      0.005

  Specific indirect

    CHEAT2
    PEER2
    TEACHER2           -0.036      0.038     -0.941      0.347

    CHEAT2
    CHJUST2
    TEACHER2           -0.157      0.098     -1.606      0.108

    CHEAT2
    CHJUST2
    PEER2
    TEACHER2           -0.083      0.038     -2.204      0.027

  Direct
    CHEAT2
    TEACHER2            0.063      0.114      0.551      0.581


Effects from CURUSE2 to SURF2

  Total                -0.140      0.133     -1.057      0.290
  Total indirect       -0.024      0.068     -0.349      0.727

  Specific indirect

    SURF2
    CHJUST2
    CURUSE2            -0.005      0.068     -0.077      0.939

    SURF2
    CHJUST2
    PEER2
    CURUSE2            -0.018      0.018     -1.009      0.313

  Direct
    SURF2
    CURUSE2            -0.116      0.102     -1.145      0.252


Effects from PERF2 to SURF2

  Total                 0.315      0.089      3.519      0.000
  Total indirect        0.221      0.069      3.186      0.001

  Specific indirect

    SURF2
    CHJUST2
    PERF2               0.227      0.070      3.272      0.001

    SURF2
    CHJUST2
    PEER2
    PERF2              -0.007      0.012     -0.583      0.560

  Direct
    SURF2
    PERF2               0.094      0.080      1.179      0.238
```

```
Effects from SUB2 to SURF2

  Total                -0.368      0.074     -4.983      0.000
  Total indirect       -0.176      0.065     -2.724      0.006

  Specific indirect

    SURF2
    PERF2
    SUB2                0.011      0.011      0.955      0.340

    SURF2
    CURUSE2
    SUB2               -0.061      0.052     -1.165      0.244

    SURF2
    CHJUST2
    SUB2               -0.085      0.051     -1.658      0.097

    SURF2
    TEACHER2
    SUB2                0.022      0.049      0.448      0.654

    SURF2
    CHJUST2
    PERF2
    SUB2                0.026      0.018      1.461      0.144

    SURF2
    CHJUST2
    CURUSE2
    SUB2               -0.003      0.036     -0.077      0.939

    SURF2
    CHJUST2
    TEACHER2
    SUB2               -0.049      0.030     -1.669      0.095

    SURF2
    CHJUST2
    PEER2
    PERF2
    SUB2               -0.001      0.001     -0.561      0.575

    SURF2
    CHJUST2
    PEER2
    CURUSE2
    SUB2               -0.010      0.010     -0.994      0.320

    SURF2
    CHJUST2
    PEER2
    TEACHER2
    SUB2               -0.026      0.011     -2.499      0.012

  Direct
    SURF2
    SUB2               -0.192      0.091     -2.103      0.036



Effects from TEACHER2 to SURF2

  Total                -0.145      0.134     -1.081      0.280
  Total indirect       -0.205      0.080     -2.566      0.010
```

```
  Specific indirect

    SURF2
    CHJUST2
    TEACHER2          -0.134      0.077     -1.733      0.083

    SURF2
    CHJUST2
    PEER2
    TEACHER2          -0.071      0.028     -2.544      0.011

  Direct
    SURF2
    TEACHER2           0.060      0.133      0.453      0.651


Effects from SUB2 to CHJUST2

  Total               -0.253      0.072     -3.516      0.000
  Total indirect      -0.108      0.065     -1.656      0.098

  Specific indirect

    CHJUST2
    PERF2
    SUB2               0.044      0.028      1.540      0.124

    CHJUST2
    CURUSE2
    SUB2              -0.005      0.061     -0.077      0.939

    CHJUST2
    TEACHER2
    SUB2              -0.085      0.049     -1.715      0.086

    CHJUST2
    PEER2
    PERF2
    SUB2              -0.001      0.002     -0.559      0.576

    CHJUST2
    PEER2
    CURUSE2
    SUB2              -0.016      0.017     -0.971      0.332

    CHJUST2
    PEER2
    TEACHER2
    SUB2              -0.045      0.017     -2.614      0.009

  Direct
    CHJUST2
    SUB2              -0.144      0.089     -1.631      0.103


Effects from CURUSE2 to CHJUST2

  Total               -0.041      0.115     -0.354      0.724
  Total indirect      -0.032      0.032     -0.987      0.323

  Specific indirect

    CHJUST2
    PEER2
    CURUSE2           -0.032      0.032     -0.987      0.323
```

```
Direct
    CHJUST2
    CURUSE2            -0.009      0.116     -0.077      0.939


Effects from PERF2 to CHJUST2

  Total                 0.377      0.085      4.448      0.000
  Total indirect       -0.012      0.020     -0.582      0.561

  Specific indirect

    CHJUST2
    PEER2
    PERF2              -0.012      0.020     -0.582      0.561

  Direct
    CHJUST2
    PERF2               0.388      0.085      4.591      0.000


Effects from TEACHER2 to CHJUST2

  Total                -0.350      0.131     -2.672      0.008
  Total indirect       -0.121      0.046     -2.617      0.009

  Specific indirect

    CHJUST2
    PEER2
    TEACHER2           -0.121      0.046     -2.617      0.009

  Direct
    CHJUST2
    TEACHER2           -0.229      0.129     -1.769      0.077


Effects from HON2 to CHJUST2

  Total                -0.136      0.074     -1.844      0.065
  Total indirect       -0.135      0.049     -2.740      0.006

  Specific indirect

    CHJUST2
    PEER2
    HON2               -0.028      0.022     -1.278      0.201

    CHJUST2
    TEACHER2
    HON2               -0.070      0.040     -1.729      0.084

    CHJUST2
    PEER2
    TEACHER2
    HON2               -0.037      0.016     -2.264      0.024

  Direct
    CHJUST2
    HON2               -0.001      0.084     -0.013      0.990


Effects from CURUSE2 to PEER2

  Total                -0.107      0.100     -1.076      0.282
  Total indirect        0.000      0.000      0.000      1.000
```

```
  Direct
    PEER2
    CURUSE2           -0.107      0.100     -1.076      0.282


Effects from HON2 to PEER2

  Total               -0.221      0.066     -3.343      0.001
  Total indirect      -0.126      0.040     -3.146      0.002

  Specific indirect

    PEER2
    TEACHER2
    HON2              -0.126      0.040     -3.146      0.002

  Direct
    PEER2
    HON2              -0.095      0.071     -1.336      0.181
```

Appendix Y: Stepwise regression of predictors of Time 2 variables

Table Y1

Standardized coefficients of stepwise regression on Peer norms and Justifiability of cheating, Time 2

	Peer norms						Justifiability of cheating					
	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 1	Step 2	Step 3	Step 4	Step5	Step 6
	β	β	β	β	β	β	β	β	β	β	β	β
Curuse	-.358***	-.394***	---	-.343***	-.334***	-.117	-.274***	-.304***	-.230***	-.213*	-.244**	-.066
Gender		.215***	---	.204***	.216***	.224***		.274***	.298***	.287***	.220***	.228
Grade		-.105	---	-.096	-.102	-.137**		-.007	-.010	-.003	.032	.001
Sub			---	---	---	---			-.144	-.106	-.154	-.117
Hon				-.221**	-.227**	-.103				-.157*	-.112	-.013
Perf					-.061	-.039					.356***	.366***
Teacher						-.384***						-.349**

Table Y2

Standardized coefficients of stepwise regression on Surface learning strategies and Self-reported cheating, Time 2

	Surface learning strategies						Self-reported cheating					
	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 1	Step 2	Step 3	Step 4	Step5	Step 6
	β	β	β	β	β	β	β	β	β	β	β	β
Curuse	-.328***	-.337***	-.196*	---	-.211*	-.134	-.338***	-.354***	-.273**	-.256**	.261**	-.160
Gender		.159*	.201***	---	.141*	.140*		.224***	-.159	.242***	.231***	.233***
Grade		.062	.055	---	.087	.073		.049	.252***	.050	.055	.036
Sub			-.260**	---	-.295**	-.269**			.043	-.128	-.137	-.119
Hon				---	---	---				-.143	-.137	-.082
Perf					.298***	.297**					.060	.069
Teacher						-.146						-.194

Note. All models satisfied fit requirements for multivariate models (see Table 5.1); CURUSE = Usefulness of curriculum; Grade = Grade-level; SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance structure; TEACHER = Teacher quality; **p* < .05, ***p* < .01, ****p* < .001.

Appendix Z:

Standardized beta coefficients for Model 4 co-ed data (*N* = 297)

estimated with composite scores, Time 2

Table Z1

Model 4, Time 2: Standardized beta coefficients estimated with composite scores

	Predictors								
N = 297	Grade	Gender	Sub	Hon	Perf	Curuse	Teacher	Peer	Chjust
Sub	-.099	.245***							
Hon	.014	-.033	---						
Perf	-.128*	.264***	.095	---					
Curuse	-.105*	-.016	.540***	---	---				
Teacher	-.173**	.041	.394***	.294***	---	---			
Peer	-.150**	.249***	---	-.103	-.008	-.056	-.456***		
Chjust	.059	.134*	-.144	-.049	.372***	-.028	-.201	.290**	
Surf	.088	-.022	-.253**	---	.142	-.061	.058	---	.596***
Cheat	.035	.060	-.113	-.070	-.139	-.033	.024	.093	.695***

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; **p* < .05, ***p* < .01, ****p* < .001.

Appendix AA: Equivalent model 1, Time 2

Usefulness of curriculum* positioned as a predictor of *Teacher quality

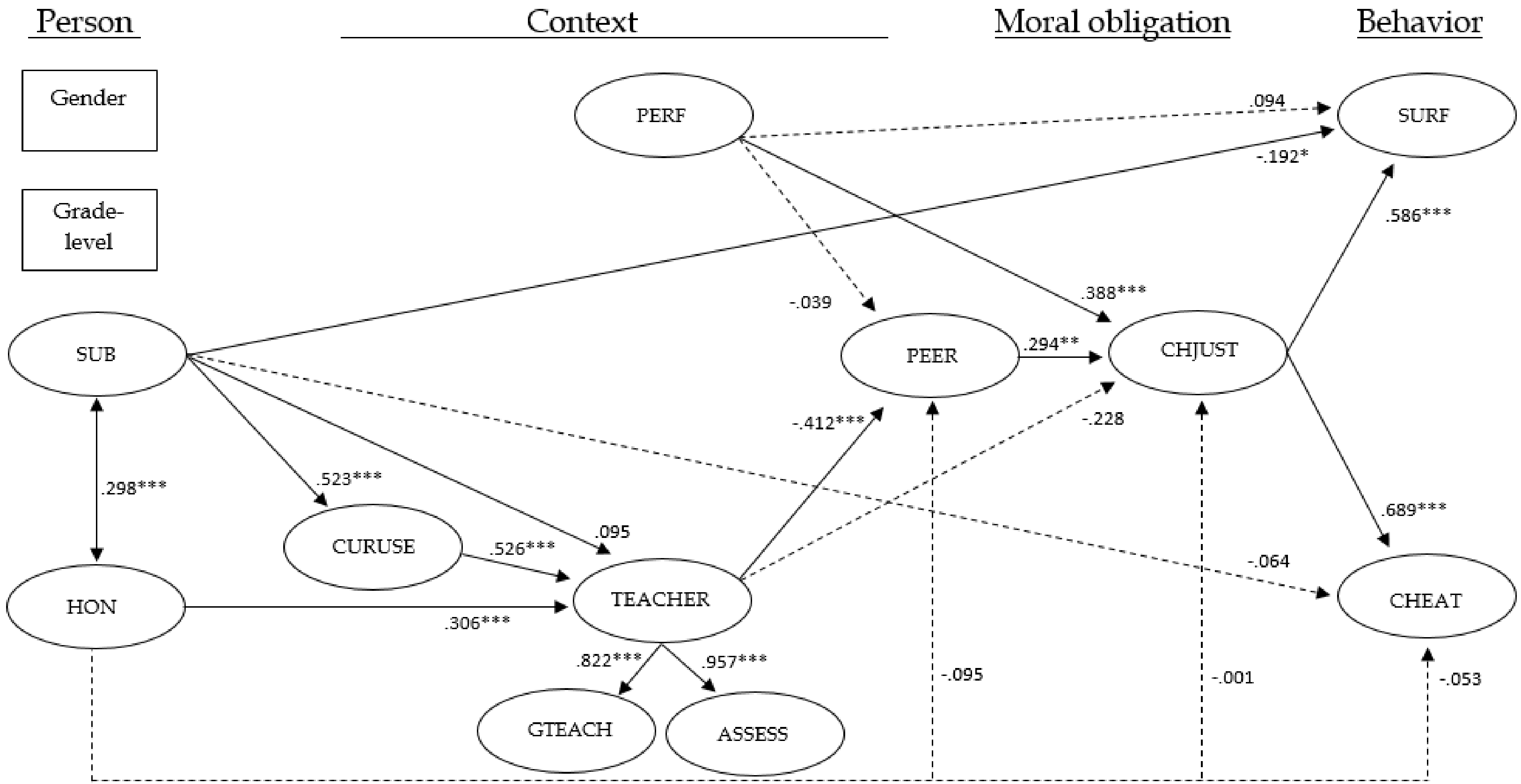

Figure AA1. Equivalent Model 1, Time 2: *Usefulness of curriculum* positioned as a predictor of *Teacher quality* (N = 297). $\chi^2(1174) = 1822$; *RMSEA* = .043, *CI*s = .039 - .047, *pclose* = 1.00; *TLI* = .90; *CFI* = .91; *SRMR* = .069; *N:q* = 1.5; SCF = .937. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; *p < .05, **p < .01, ***p < .001.

Appendix AB: Equivalent model 2, Time 2

***Peer norms* positioned as a correlate of class context**

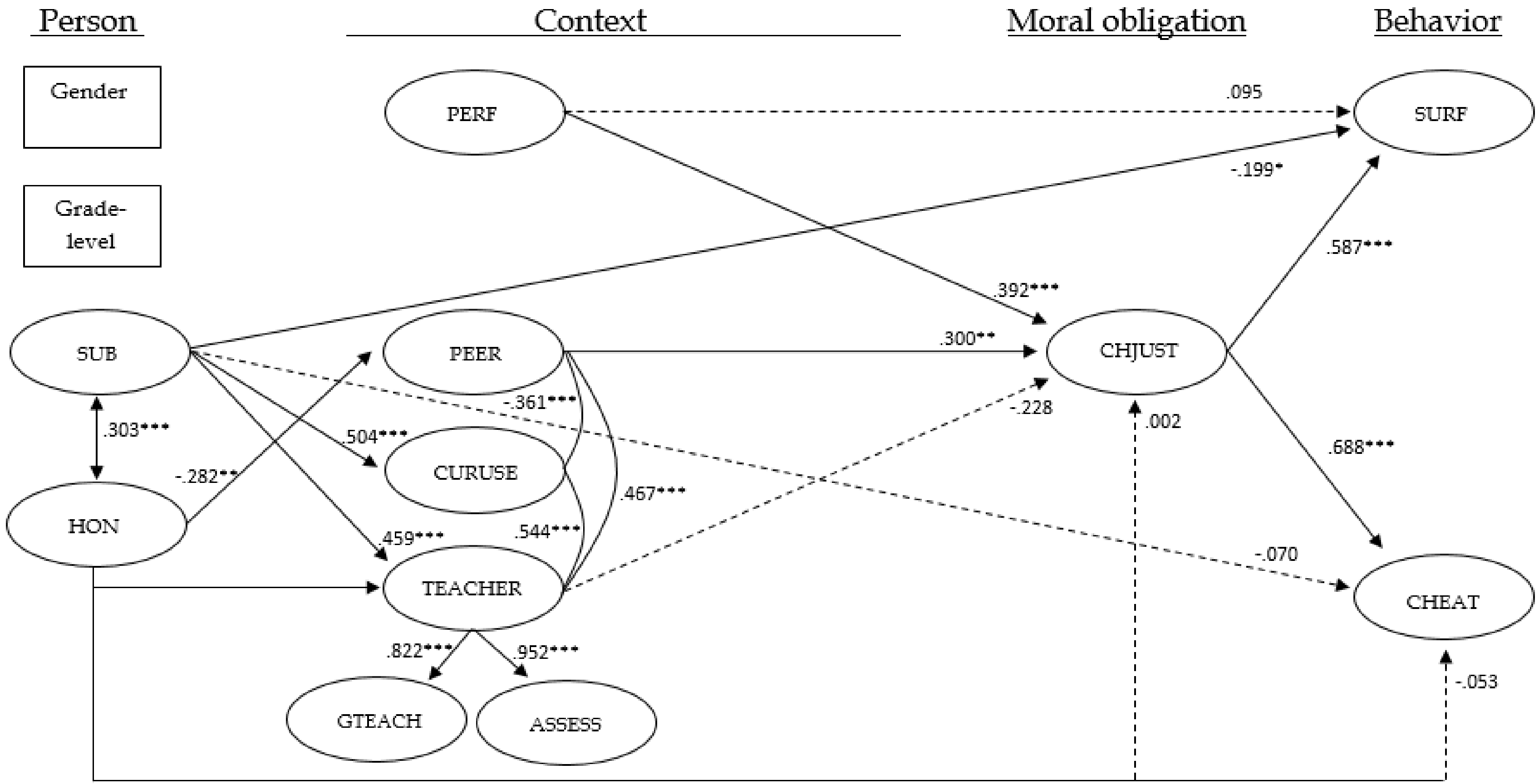

Figure AB1. Equivalent Model 2, Time 2: *Peer norms* positioned as a correlate of class context (N = 297). $\chi^2(1173) = 1818$; *RMSEA* = .043, *CIs* = .039 - .047, *pclose* = .999; *TLI* = .90; *CFI* = .91; *SRMR* = .068; *N:q* = 1.5; SCF = .937. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; **p* < .05, ***p* < .01, ****p* < .001.

Appendix AC: Equivalent model 3, Time 2

Peer norms positioned as a predictor of class context

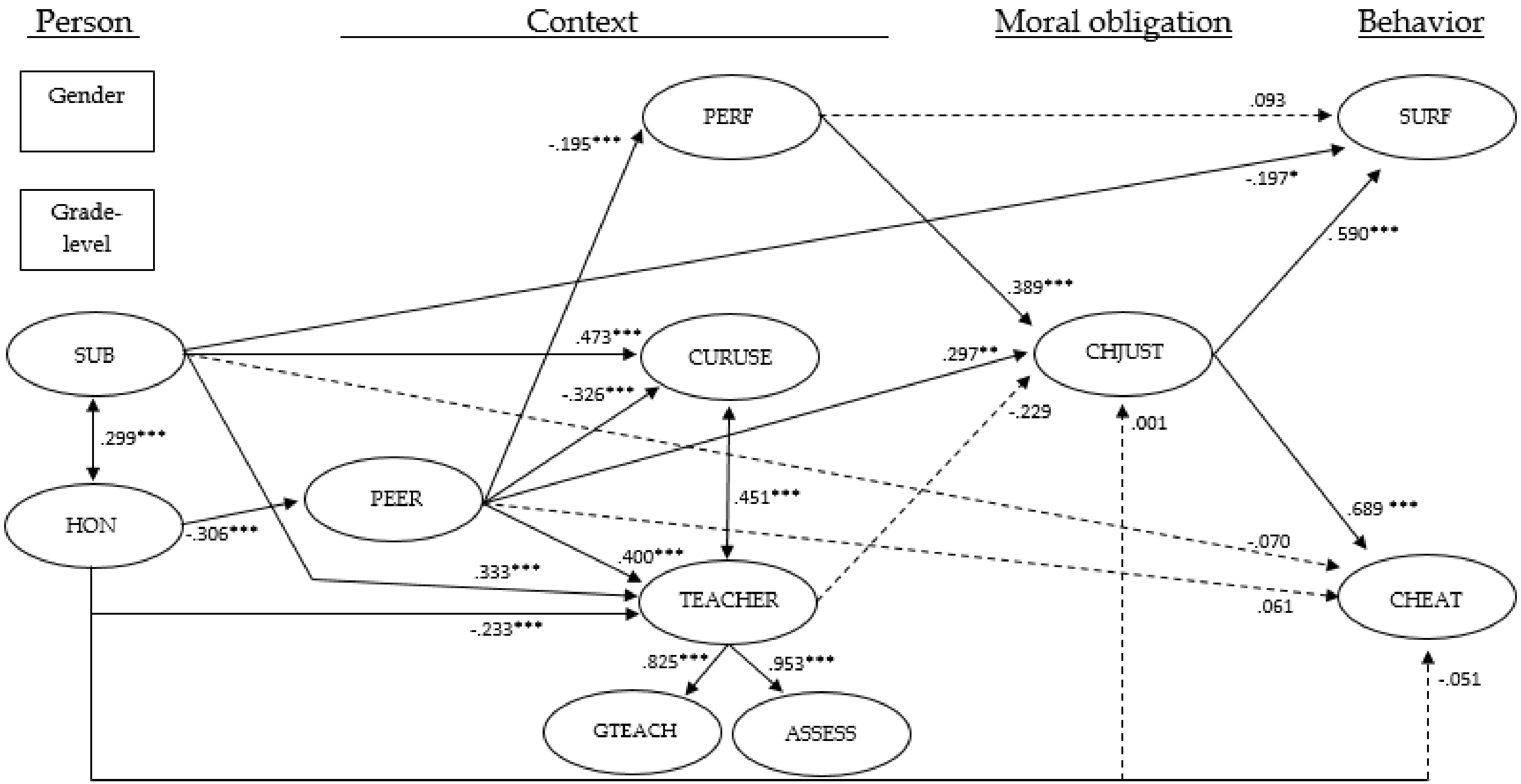

Figure AC1. Equivalent Model 3, Time 2: *Peer norms* positioned as a predictor of class context (N = 297). $\chi^2(1173) = 1820$; *RMSEA* = .043, *CIs* = .039 - .047, *pclose* = .999; *TLI* = .90; *CFI* = .91; *SRMR* = .068; *N:q* = 1.5; SCF = .937. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating; *p < .05, **p < .01, ***p < .001.

Appendix AD:

Paired-samples *t*-tests, longitudinal matched samples, Time 2

Table AD1

Two-tailed paired-samples t-tests, longitudinal matched samples

	Class of 2016							Class of 2015						
	Grade 8 → Grade 9 transition (*N* = 123)							Grade 9 → Grade 10 transition (*N* = 102)						
	Grade 8		Grade 9					Grade 9		Grade 10				
	M	*SD*	*M*	*SD*	*t*	Sig.	Interpretation	*M*	*SD*	*M*	*SD*	*t*	Sig.	Interpretation
Sub	2.43	.919	2.44	.916	-.082	NS		2.52	1.02	2.22	1.07	4.45	**.000**	Better self-concept at G10
Hon	2.01	.676	1.96	.709	1.04	NS		1.96	.548	1.93	.682	.491	NS	
Perf	3.38	.911	3.65	.940	-2.65	**.009**	More prevalent at G9	3.49	.983	3.52	1.04	-.246	NS	
Curuse	2.47	.946	2.61	1.00	-1.46	NS		2.57	1.06	2.26	1.03	2.96	**.004**	More useful at G10
Gteach	2.42	.747	2.60	.866	-2.04	**.044**	Worse evaluations at G9	2.65	.894	2.24	.765	3.92	**.000**	Better evaluations at G10
Assess	2.20	.628	2.21	.685	-.133	NS		2.26	.665	2.00	.620	4.04	**.000**	Better evaluations at G10
Peer	3.68	.789	3.83	.810	-1.74	NS		3.55	.856	3.67	1.00	-.1.35	NS	
Chjust	4.13	.930	4.06	.848	.889	NS		4.08	.933	4.16	.949	-.736	NS	
Surf	3.69	0.95	3.45	.953	2.67	**.009**	More prevalent at G9	3.54	.960	3.66	.979	-1.13	NS	
Cheat	4.39	.884	4.23	.993	1.57	NS		4.40	.876	4.45	.962	-.489	NS	

Note. SUB = Subject self-concept; HON = Honesty-trustworthiness self-concept; PERF = Performance goal structure; TEACHER = Teacher quality, ASSESS = Assessment quality; GTEACH = Good teaching; CURUSE = Usefulness of curriculum; PEER = Peer norms for cheating; CHJUST = Justifiability of cheating; SURF = Surface learning strategies; CHEAT = Self-reported cheating.

Appendix AE:

Longitudinal model output

```
MODEL FIT INFORMATION

Number of Free Parameters                        151

Loglikelihood

          H0 Value                         -5218.936
          H1 Value                         -5128.920

Information Criteria

          Akaike (AIC)                     10739.871
          Bayesian (BIC)                   11255.703
          Sample-Size Adjusted BIC         10777.153
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                              180.031
          Degrees of Freedom                     119
          P-Value                             0.0003

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                             0.048
          90 Percent C.I.                      0.033  0.061
          Probability RMSEA <= .05             0.590

CFI/TLI

          CFI                                  0.967
          TLI                                  0.937

Chi-Square Test of Model Fit for the Baseline Model

          Value                             2088.462
          Degrees of Freedom                     230
          P-Value                             0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                                0.043

MODEL RESULTS

                                                  Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 SUB1     BY
    NSUB               0.959      0.000    999.000    999.000

 HON1     BY
    NHON               0.906      0.000    999.000    999.000

 PERF1    BY
    NPERF              0.849      0.000    999.000    999.000
```

```
GTEACH1  BY
   NGTEACH            0.933      0.000    999.000    999.000

ASSESS1  BY
   NASSESS            0.917      0.000    999.000    999.000

CURUSE1  BY
   NCURUSE            0.954      0.000    999.000    999.000

PEER1    BY
   NPEER              0.860      0.000    999.000    999.000

SURF1    BY
   NSURF              0.843      0.000    999.000    999.000

CHJUST1  BY
   NCHJUST            0.889      0.000    999.000    999.000

CHEAT1   BY
   NCHEAT             0.927      0.000    999.000    999.000

SUB2     BY
   NSUB2              0.959      0.000    999.000    999.000

HON2     BY
   NHON2              0.922      0.000    999.000    999.000

PERF2    BY
   NPERF2             0.889      0.000    999.000    999.000

GTEACH2  BY
   NGTEA2             0.938      0.000    999.000    999.000

ASSESS2  BY
   NASSESS2           0.943      0.000    999.000    999.000

CURUSE2  BY
   NCURU2             0.964      0.000    999.000    999.000

PEER2    BY
   NPEER2             0.922      0.000    999.000    999.000

SURF2    BY
   NSURF2             0.872      0.000    999.000    999.000

CHJUST2  BY
   NCHJUST2           0.877      0.000    999.000    999.000

CHEAT2   BY
   NCHEAT2            0.917      0.000    999.000    999.000

TEACHER1 BY
   ASSESS1            1.000      0.000    999.000    999.000
   GTEACH1            0.948      0.081     11.634      0.000

TEACHER2 BY
   ASSESS2            1.000      0.000    999.000    999.000
   GTEACH2            0.800      0.072     11.185      0.000
```

CHEAT2 ON				
CHJUST2	0.596	0.110	5.399	0.000
PERF2	-0.081	0.076	-1.065	0.287
PEER2	0.133	0.081	1.654	0.098
SUB2	-0.116	0.075	-1.556	0.120
HON2	-0.028	0.067	-0.426	0.670
TEACHER2	0.088	0.098	0.898	0.369
CURUSE2	-0.051	0.088	-0.585	0.559
CHEAT1	0.054	0.088	0.618	0.537
SURF2 ON				
CHJUST2	0.513	0.117	4.379	0.000
PERF2	0.184	0.084	2.203	0.028
SUB2	-0.025	0.087	-0.290	0.772
TEACHER2	-0.077	0.100	-0.764	0.445
CURUSE2	-0.126	0.100	-1.259	0.208
SURF1	0.231	0.087	2.667	0.008
CHJUST2 ON				
SUB2	-0.102	0.082	-1.243	0.214
PEER2	0.219	0.087	2.523	0.012
PERF2	0.300	0.074	4.036	0.000
TEACHER2	-0.070	0.110	-0.641	0.522
CURUSE2	-0.133	0.098	-1.360	0.174
HON2	-0.037	0.072	-0.508	0.611
CHJUST1	0.276	0.087	3.183	0.001
PEER2 ON				
PERF2	-0.105	0.070	-1.507	0.132
TEACHER2	-0.358	0.107	-3.354	0.001
CURUSE2	-0.051	0.092	-0.559	0.576
HON2	-0.105	0.072	-1.458	0.145
PEER1	0.379	0.073	5.201	0.000
PERF2 ON				
SUB2	0.016	0.080	0.202	0.840
PERF1	0.313	0.086	3.662	0.000
TEACHER2 ON				
SUB2	0.281	0.069	4.069	0.000
TEACHER1	0.357	0.072	4.988	0.000
HON2	0.210	0.057	3.671	0.000
CURUSE2 ON				
SUB2	0.416	0.069	6.027	0.000
CURUSE1	0.282	0.062	4.524	0.000
SUB2 ON				
SUB1	0.773	0.052	14.972	0.000
HON2 ON				
HON1	0.739	0.065	11.422	0.000
CHEAT1 ON				
CHJUST1	0.432	0.132	3.284	0.001
PERF1	0.017	0.079	0.220	0.826
PEER1	0.069	0.098	0.707	0.480
SUB1	-0.196	0.067	-2.907	0.004
HON1	-0.210	0.063	-3.339	0.001
TEACHER1	-0.081	0.120	-0.669	0.503
CURUSE1	0.113	0.099	1.144	0.253

```
 SURF1    ON
    CHJUST1            0.785      0.133      5.913      0.000
    PERF1              0.028      0.102      0.273      0.785
    SUB1              -0.254      0.094     -2.716      0.007
    TEACHER1           0.116      0.156      0.740      0.459
    CURUSE1           -0.058      0.126     -0.459      0.646

 CHJUST1  ON
    SUB1               0.021      0.076      0.279      0.780
    PEER1              0.401      0.088      4.549      0.000
    PERF1              0.296      0.074      3.997      0.000
    TEACHER1          -0.049      0.138     -0.352      0.725
    CURUSE1           -0.177      0.107     -1.653      0.098
    HON1              -0.171      0.066     -2.594      0.010

 PEER1    ON
    PERF1              0.123      0.091      1.347      0.178
    TEACHER1          -0.591      0.159     -3.720      0.000
    CURUSE1            0.165      0.130      1.275      0.202
    HON1              -0.222      0.082     -2.699      0.007

 PERF1    ON
    SUB1               0.088      0.086      1.018      0.309

 TEACHER1 ON
    SUB1               0.413      0.073      5.623      0.000
    HON1               0.160      0.059      2.703      0.007

 CURUSE1  ON
    SUB1               0.546      0.068      8.057      0.000

 CHEAT2   ON
    GENDER            -0.002      0.127     -0.012      0.991
    GRADE2             0.121      0.110      1.100      0.271

 SURF2    ON
    GENDER             0.042      0.144      0.290      0.772
    GRADE2             0.156      0.123      1.271      0.204

 CHJUST2  ON
    GENDER             0.162      0.144      1.123      0.261
    GRADE2             0.126      0.124      1.017      0.309

 PEER2    ON
    GENDER             0.384      0.134      2.862      0.004
    GRADE2            -0.274      0.122     -2.253      0.024

 PERF2    ON
    GENDER             0.357      0.158      2.266      0.023
    GRADE2            -0.195      0.142     -1.366      0.172

 TEACHER2 ON
    GENDER             0.279      0.124      2.261      0.024
    GRADE2            -0.372      0.114     -3.268      0.001

 CURUSE2  ON
    GENDER             0.079      0.122      0.647      0.517
    GRADE2            -0.279      0.111     -2.515      0.012

 SUB2     ON
    GENDER             0.111      0.102      1.093      0.275
    GRADE2            -0.288      0.091     -3.165      0.002
```

```
HON2     ON
   GENDER              0.106      0.123      0.862      0.389
   GRADE2             -0.034      0.115     -0.294      0.768


CHEAT1   ON
   GENDER              0.154      0.114      1.351      0.177
   GRADE2             -0.043      0.097     -0.439      0.661

SURF1    ON
   GENDER             -0.174      0.160     -1.090      0.276
   GRADE2              0.015      0.136      0.109      0.914

CHJUST1  ON
   GENDER              0.165      0.128      1.285      0.199
   GRADE2              0.004      0.111      0.041      0.968

PEER1    ON
   GENDER              0.287      0.149      1.920      0.055
   GRADE2             -0.141      0.140     -1.009      0.313

PERF1    ON
   GENDER              0.209      0.172      1.219      0.223
   GRADE2              0.113      0.153      0.739      0.460

TEACHER1 ON
   GENDER             -0.147      0.136     -1.076      0.282
   GRADE2              0.181      0.121      1.495      0.135

CURUSE1  ON
   GENDER             -0.272      0.134     -2.029      0.042
   GRADE2              0.099      0.120      0.821      0.411

SUB1     ON
   GENDER              0.593      0.138      4.292      0.000
   GRADE2              0.027      0.129      0.210      0.834

HON1     ON
   GENDER             -0.142      0.155     -0.916      0.360
   GRADE2             -0.021      0.146     -0.143      0.886

SUB2     WITH
   HON2                0.057      0.038      1.528      0.126

PERF2    WITH
   TEACHER2            0.070      0.057      1.212      0.226
   CURUSE2             0.080      0.056      1.416      0.157

TEACHER2 WITH
   CURUSE2             0.287      0.048      5.930      0.000

SURF2    WITH
   CHEAT2              0.047      0.047      0.997      0.319

SUB1     WITH
   HON1                0.239      0.070      3.412      0.001

PERF1    WITH
   TEACHER1           -0.151      0.069     -2.193      0.028
   CURUSE1             0.052      0.068      0.766      0.443

TEACHER1 WITH
   CURUSE1             0.389      0.061      6.396      0.000
```

```
SURF1    WITH
   CHEAT1             0.072      0.048      1.526      0.127

Intercepts
   NSUB              -1.193      1.181     -1.010      0.313
   NHON               0.398      1.255      0.317      0.751
   NPERF             -1.306      1.241     -1.052      0.293
   NGTEACH           -1.693      1.137     -1.489      0.137
   NCURUSE           -1.099      1.233     -0.892      0.372
   NASSESS           -1.753      1.167     -1.502      0.133
   NPEER              1.291      1.240      1.041      0.298
   NSURF              0.245      1.254      0.196      0.845
   NCHJUST           -0.012      1.127     -0.010      0.992
   NCHEAT             0.325      1.068      0.304      0.761
   NSUB2              1.514      1.169      1.295      0.195
   NHON2              0.431      1.252      0.344      0.730
   NPERF2             0.689      1.215      0.567      0.571
   NGTEA2             2.176      0.969      2.246      0.025
   NCURU2             2.739      1.189      2.305      0.021
   NASSESS2           2.739      1.180      2.322      0.020
   NPEER2             1.096      1.218      0.900      0.368
   NSURF2            -2.470      1.216     -2.031      0.042
   NCHJUST2          -1.544      1.140     -1.354      0.176
   NCHEAT2           -2.012      1.073     -1.876      0.061

Residual Variances
   NSUB               0.080      0.000    999.000    999.000
   NHON               0.180      0.000    999.000    999.000
   NPERF              0.280      0.000    999.000    999.000
   NGTEACH            0.130      0.000    999.000    999.000
   NCURUSE            0.090      0.000    999.000    999.000
   NASSESS            0.160      0.000    999.000    999.000
   NPEER              0.260      0.000    999.000    999.000
   NSURF              0.290      0.000    999.000    999.000
   NCHJUST            0.210      0.000    999.000    999.000
   NCHEAT             0.140      0.000    999.000    999.000
   NSUB2              0.080      0.000    999.000    999.000
   NHON2              0.150      0.000    999.000    999.000
   NPERF2             0.210      0.000    999.000    999.000
   NGTEA2             0.120      0.000    999.000    999.000
   NCURU2             0.070      0.000    999.000    999.000
   NASSESS2           0.110      0.000    999.000    999.000
   NPEER2             0.150      0.000    999.000    999.000
   NSURF2             0.240      0.000    999.000    999.000
   NCHJUST2           0.230      0.000    999.000    999.000
   NCHEAT2            0.160      0.000    999.000    999.000
   SUB1               0.844      0.088      9.614      0.000
   HON1               0.958      0.111      8.636      0.000
   PERF1              0.916      0.123      7.446      0.000
   GTEACH1            0.245      0.059      4.129      0.000
   ASSESS1            0.147      0.061      2.420      0.016
   CURUSE1            0.673      0.075      8.990      0.000
   PEER1              0.600      0.096      6.248      0.000
   SURF1              0.416      0.091      4.568      0.000
   CHJUST1            0.285      0.059      4.861      0.000
   CHEAT1             0.273      0.047      5.772      0.000
   TEACHER1           0.597      0.090      6.653      0.000
   SUB2               0.319      0.043      7.451      0.000
   HON2               0.432      0.068      6.373      0.000
   PERF2              0.789      0.102      7.703      0.000
   GTEACH2            0.374      0.062      6.059      0.000
   ASSESS2            0.011      0.059      0.190      0.849
```

```
CURUSE2            0.558      0.062      9.056      0.000
PEER2              0.492      0.069      7.175      0.000
SURF2              0.354      0.073      4.887      0.000
CHJUST2            0.391      0.066      5.886      0.000
CHEAT2             0.262      0.054      4.862      0.000
TEACHER2           0.518      0.080      6.475      0.000
```

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SUB1 BY				
NSUB	0.956	0.004	224.528	0.000
HON1 BY				
NHON	0.902	0.010	93.151	0.000
PERF1 BY				
NPERF	0.841	0.016	51.417	0.000
GTEACH1 BY				
NGTEACH	0.931	0.007	138.660	0.000
ASSESS1 BY				
NASSESS	0.913	0.009	107.124	0.000
CURUSE1 BY				
NCURUSE	0.951	0.005	199.375	0.000
PEER1 BY				
NPEER	0.855	0.015	58.283	0.000
SURF1 BY				
NSURF	0.836	0.017	49.936	0.000
CHJUST1 BY				
NCHJUST	0.860	0.014	61.393	0.000
CHEAT1 BY				
NCHEAT	0.897	0.010	88.031	0.000
SUB2 BY				
NSUB2	0.955	0.004	221.997	0.000
HON2 BY				
NHON2	0.919	0.008	115.446	0.000
PERF2 BY				
NPERF2	0.882	0.012	74.235	0.000
GTEACH2 BY				
NGTEA2	0.936	0.006	151.213	0.000
ASSESS2 BY				
NASSESS2	0.939	0.006	162.563	0.000
CURUSE2 BY				
NCURU2	0.961	0.004	257.057	0.000
PEER2 BY				
NPEER2	0.916	0.008	111.241	0.000
SURF2 BY				
NSURF2	0.861	0.014	62.354	0.000
CHJUST2 BY				

```
   NCHJUST2           0.848      0.015     55.558      0.000

CHEAT2   BY
   NCHEAT2            0.882      0.012     74.672      0.000

TEACHER1 BY
   ASSESS1            0.920      0.034     26.773      0.000
   GTEACH1            0.865      0.035     24.474      0.000

TEACHER2 BY
   ASSESS2            0.994      0.032     30.813      0.000
   GTEACH2            0.781      0.040     19.610      0.000

CHEAT2   ON
   CHJUST2            0.639      0.108      5.923      0.000
   PERF2             -0.096      0.090     -1.067      0.286
   PEER2              0.157      0.094      1.664      0.096
   SUB2              -0.135      0.087     -1.561      0.119
   HON2              -0.034      0.080     -0.425      0.671
   TEACHER2           0.103      0.114      0.903      0.366
   CURUSE2           -0.060      0.103     -0.584      0.559
   CHEAT1             0.054      0.088      0.618      0.537

SURF2    ON
   CHJUST2            0.472      0.102      4.642      0.000
   PERF2              0.187      0.084      2.212      0.027
   SUB2              -0.025      0.088     -0.290      0.771
   TEACHER2          -0.077      0.101     -0.766      0.444
   CURUSE2           -0.127      0.101     -1.259      0.208
   SURF1              0.237      0.089      2.668      0.008

CHJUST2  ON
   SUB2              -0.111      0.089     -1.247      0.212
   PEER2              0.240      0.094      2.566      0.010
   PERF2              0.331      0.080      4.142      0.000
   TEACHER2          -0.077      0.119     -0.644      0.520
   CURUSE2           -0.145      0.107     -1.359      0.174
   HON2              -0.041      0.081     -0.508      0.612
   CHJUST1            0.274      0.083      3.290      0.001

PEER2    ON
   PERF2             -0.106      0.070     -1.516      0.130
   TEACHER2          -0.357      0.101     -3.532      0.000
   CURUSE2           -0.051      0.091     -0.559      0.576
   HON2              -0.107      0.073     -1.462      0.144
   PEER1              0.386      0.070      5.554      0.000

PERF2    ON
   SUB2               0.016      0.079      0.202      0.840
   PERF1              0.316      0.082      3.868      0.000

TEACHER2 ON
   SUB2               0.280      0.067      4.183      0.000
   TEACHER1           0.335      0.063      5.325      0.000
   HON2               0.214      0.059      3.630      0.000


CURUSE2  ON
   SUB2               0.415      0.065      6.419      0.000
   CURUSE1            0.285      0.061      4.637      0.000

SUB2     ON
   SUB1               0.779      0.036     21.663      0.000
```

```
HON2     ON
   HON1              0.742      0.045     16.578      0.000

CHEAT1   ON
   CHJUST1           0.459      0.135      3.404      0.001
   PERF1             0.021      0.094      0.220      0.826
   PEER1             0.083      0.117      0.707      0.479
   SUB1             -0.230      0.079     -2.927      0.003
   HON1             -0.252      0.074     -3.383      0.001
   TEACHER1         -0.089      0.132     -0.669      0.503
   CURUSE1           0.133      0.116      1.146      0.252

SURF1    ON
   CHJUST1           0.700      0.105      6.657      0.000
   PERF1             0.028      0.102      0.273      0.785
   SUB1             -0.250      0.091     -2.741      0.006
   TEACHER1          0.107      0.144      0.743      0.458
   CURUSE1          -0.057      0.125     -0.459      0.646

CHJUST1  ON
   SUB1              0.023      0.084      0.279      0.780
   PEER1             0.452      0.094      4.829      0.000
   PERF1             0.331      0.080      4.143      0.000
   TEACHER1         -0.050      0.143     -0.352      0.725
   CURUSE1          -0.197      0.119     -1.655      0.098
   HON1             -0.194      0.074     -2.610      0.009

PEER1    ON
   PERF1             0.122      0.090      1.352      0.176
   TEACHER1         -0.544      0.135     -4.024      0.000
   CURUSE1           0.163      0.127      1.279      0.201
   HON1             -0.223      0.081     -2.738      0.006

PERF1    ON
   SUB1              0.087      0.085      1.022      0.307

TEACHER1 ON
   SUB1              0.441      0.068      6.451      0.000
   HON1              0.174      0.064      2.717      0.007

CURUSE1  ON
   SUB1              0.544      0.058      9.397      0.000

CHEAT2   ON
   GENDER           -0.001      0.072     -0.012      0.991
   GRADE2            0.074      0.067      1.101      0.271

SURF2    ON
   GENDER            0.020      0.070      0.290      0.772
   GRADE2            0.082      0.064      1.271      0.204

CHJUST2  ON
   GENDER            0.086      0.077      1.124      0.261
   GRADE2            0.072      0.071      1.019      0.308


PEER2    ON
   GENDER            0.187      0.065      2.879      0.004
   GRADE2           -0.142      0.063     -2.260      0.024

PERF2    ON
   GENDER            0.173      0.075      2.291      0.022
```

```
   GRADE2             -0.100      0.073     -1.371      0.170

TEACHER2 ON
   GENDER              0.136      0.060      2.252      0.024
   GRADE2             -0.193      0.060     -3.242      0.001

CURUSE2  ON
   GENDER              0.039      0.060      0.648      0.517
   GRADE2             -0.146      0.058     -2.527      0.011

SUB2     ON
   GENDER              0.054      0.050      1.092      0.275
   GRADE2             -0.150      0.048     -3.152      0.002


HON2     ON
   GENDER              0.050      0.058      0.862      0.389
   GRADE2             -0.017      0.058     -0.294      0.768

CHEAT1   ON
   GENDER              0.088      0.065      1.352      0.176
   GRADE2             -0.026      0.059     -0.439      0.661

SURF1    ON
   GENDER             -0.083      0.077     -1.091      0.275
   GRADE2              0.008      0.070      0.109      0.914

CHJUST1  ON
   GENDER              0.088      0.069      1.286      0.198
   GRADE2              0.003      0.063      0.041      0.968

PEER1    ON
   GENDER              0.137      0.071      1.930      0.054
   GRADE2             -0.072      0.071     -1.010      0.312

PERF1    ON
   GENDER              0.100      0.082      1.224      0.221
   GRADE2              0.058      0.079      0.740      0.459

TEACHER1 ON
   GENDER             -0.076      0.071     -1.080      0.280
   GRADE2              0.100      0.067      1.494      0.135

CURUSE1  ON
   GENDER             -0.132      0.065     -2.039      0.041
   GRADE2              0.051      0.062      0.822      0.411

SUB1     ON
   GENDER              0.288      0.064      4.480      0.000
   GRADE2              0.014      0.067      0.210      0.834

HON1     ON
   GENDER             -0.068      0.074     -0.918      0.359
   GRADE2             -0.011      0.074     -0.143      0.886

SUB2     WITH
   HON2                0.155      0.099      1.562      0.118

PERF2    WITH
   TEACHER2            0.109      0.090      1.213      0.225
   CURUSE2             0.120      0.083      1.439      0.150

TEACHER2 WITH
```

```
   CURUSE2            0.535      0.066      8.152      0.000

SURF2    WITH
   CHEAT2             0.155      0.147      1.049      0.294

SUB1     WITH
   HON1               0.266      0.071      3.739      0.000

PERF1    WITH
   TEACHER1          -0.204      0.089     -2.279      0.023
   CURUSE1            0.066      0.086      0.770      0.441

TEACHER1 WITH
   CURUSE1            0.613      0.059     10.385      0.000

SURF1    WITH
   CHEAT1             0.215      0.131      1.645      0.100



Intercepts
   NSUB              -1.239      1.223     -1.013      0.311
   NHON               0.404      1.274      0.317      0.751
   NPERF             -1.334      1.263     -1.056      0.291
   NGTEACH           -1.715      1.144     -1.500      0.134
   NCURUSE           -1.137      1.271     -0.894      0.371
   NASSESS           -1.787      1.180     -1.513      0.130
   NPEER              1.313      1.257      1.045      0.296
   NSURF              0.250      1.277      0.196      0.845
   NCHJUST           -0.013      1.255     -0.010      0.992
   NCHEAT             0.385      1.264      0.305      0.761
   NSUB2              1.583      1.216      1.302      0.193
   NHON2              0.440      1.276      0.345      0.730
   NPERF2             0.709      1.250      0.567      0.570
   NGTEA2             2.215      0.971      2.282      0.022
   NCURU2             2.864      1.222      2.343      0.019
   NASSESS2           2.833      1.200      2.361      0.018
   NPEER2             1.137      1.260      0.902      0.367
   NSURF2            -2.564      1.245     -2.059      0.039
   NCHJUST2          -1.706      1.252     -1.362      0.173
   NCHEAT2           -2.372      1.250     -1.897      0.058

Residual Variances
   NSUB               0.086      0.008     10.606      0.000
   NHON               0.186      0.017     10.613      0.000
   NPERF              0.292      0.028     10.604      0.000
   NGTEACH            0.133      0.012     10.676      0.000
   NCURUSE            0.096      0.009     10.607      0.000
   NASSESS            0.166      0.016     10.682      0.000
   NPEER              0.269      0.025     10.732      0.000
   NSURF              0.301      0.028     10.732      0.000
   NCHJUST            0.261      0.024     10.814      0.000
   NCHEAT             0.196      0.018     10.735      0.000
   NSUB2              0.087      0.008     10.642      0.000
   NHON2              0.156      0.015     10.660      0.000
   NPERF2             0.222      0.021     10.618      0.000
   NGTEA2             0.124      0.012     10.730      0.000
   NCURU2             0.077      0.007     10.654      0.000
   NASSESS2           0.118      0.011     10.846      0.000
   NPEER2             0.161      0.015     10.709      0.000
   NSURF2             0.259      0.024     10.876      0.000
   NCHJUST2           0.281      0.026     10.852      0.000
   NCHEAT2            0.222      0.021     10.673      0.000
```

SUB1	0.916	0.037	24.647	0.000
HON1	0.995	0.010	97.691	0.000
PERF1	0.973	0.026	37.598	0.000
GTEACH1	0.253	0.061	4.136	0.000
ASSESS1	0.154	0.063	2.433	0.015
CURUSE1	0.725	0.058	12.574	0.000
PEER1	0.629	0.076	8.306	0.000
SURF1	0.438	0.083	5.273	0.000
CHJUST1	0.378	0.069	5.443	0.000
CHEAT1	0.408	0.065	6.318	0.000
TEACHER1	0.739	0.061	12.169	0.000
SUB2	0.352	0.048	7.324	0.000
HON2	0.452	0.066	6.855	0.000
PERF2	0.850	0.058	14.728	0.000
GTEACH2	0.390	0.062	6.260	0.000
ASSESS2	0.012	0.064	0.190	0.849
CURUSE2	0.614	0.056	10.962	0.000
PEER2	0.537	0.064	8.361	0.000
SURF2	0.392	0.070	5.620	0.000
CHJUST2	0.511	0.070	7.327	0.000
CHEAT2	0.393	0.074	5.302	0.000
TEACHER2	0.565	0.061	9.338	0.000

R-SQUARE

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
NSUB	0.914	0.008	112.264	0.000
NHON	0.814	0.017	46.575	0.000
NPERF	0.708	0.028	25.708	0.000
NGTEACH	0.867	0.012	69.330	0.000
NCURUSE	0.904	0.009	99.688	0.000
NASSESS	0.834	0.016	53.562	0.000
NPEER	0.731	0.025	29.142	0.000
NSURF	0.699	0.028	24.968	0.000
NCHJUST	0.739	0.024	30.696	0.000
NCHEAT	0.804	0.018	44.015	0.000
NSUB2	0.913	0.008	110.998	0.000
NHON2	0.844	0.015	57.723	0.000
NPERF2	0.778	0.021	37.117	0.000
NGTEA2	0.876	0.012	75.606	0.000
NCURU2	0.923	0.007	128.529	0.000
NASSESS2	0.882	0.011	81.282	0.000
NPEER2	0.839	0.015	55.621	0.000
NSURF2	0.741	0.024	31.177	0.000
NCHJUST2	0.719	0.026	27.779	0.000
NCHEAT2	0.778	0.021	37.336	0.000

Latent Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SUB1	0.084	0.037	2.257	0.024
HON1	0.005	0.010	0.471	0.638
PERF1	0.027	0.026	1.055	0.292
GTEACH1	0.747	0.061	12.237	0.000
ASSESS1	0.846	0.063	13.386	0.000
CURUSE1	0.275	0.058	4.777	0.000
PEER1	0.371	0.076	4.906	0.000
SURF1	0.562	0.083	6.770	0.000
CHJUST1	0.622	0.069	8.967	0.000

CHEAT1	0.592	0.065	9.168	0.000
TEACHER1	0.261	0.061	4.289	0.000
SUB2	0.648	0.048	13.485	0.000
HON2	0.548	0.066	8.319	0.000
PERF2	0.150	0.058	2.601	0.009
GTEACH2	0.610	0.062	9.805	0.000
ASSESS2	0.988	0.064	15.407	0.000
CURUSE2	0.386	0.056	6.901	0.000
PEER2	0.463	0.064	7.207	0.000
SURF2	0.608	0.070	8.727	0.000
CHJUST2	0.489	0.070	7.009	0.000
CHEAT2	0.607	0.074	8.185	0.000
TEACHER2	0.435	0.061	7.183	0.000

QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix 0.178E-05
(ratio of smallest to largest eigenvalue)

ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	SUB1	HON1	PERF1	GTEACH1	ASSESS1
SUB1	0.921				
HON1	0.220	0.962			
PERF1	0.110	0.012	0.942		
GTEACH1	0.378	0.235	-0.100	0.970	
ASSESS1	0.399	0.248	-0.106	0.765	0.954
CURUSE1	0.470	0.128	0.099	0.574	0.606
PEER1	-0.159	-0.345	0.204	-0.423	-0.446
SURF1	-0.337	-0.292	0.252	-0.335	-0.353
CHJUST1	-0.131	-0.335	0.358	-0.368	-0.388
CHEAT1	-0.252	-0.424	0.189	-0.309	-0.327
TEACHER1	0.399	0.248	-0.106	0.765	0.807
SUB2	0.722	0.169	0.082	0.282	0.297
HON2	0.176	0.708	0.014	0.174	0.184
PERF2	0.089	-0.003	0.311	-0.029	-0.031
GTEACH2	0.330	0.223	-0.006	0.302	0.318
ASSESS2	0.412	0.279	-0.008	0.377	0.398
CURUSE2	0.438	0.106	0.058	0.268	0.283
PEER2	-0.213	-0.320	0.056	-0.329	-0.348
SURF2	-0.262	-0.236	0.218	-0.284	-0.299
CHJUST2	-0.200	-0.247	0.202	-0.271	-0.286
CHEAT2	-0.242	-0.234	0.104	-0.233	-0.245
TEACHER2	0.412	0.279	-0.008	0.377	0.398
GENDER	0.130	-0.031	0.059	0.018	0.019
GRADE2	0.016	-0.007	0.033	0.046	0.048

ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES, CONTINUED

	CURUSE1	PEER1	SURF1	CHJUST1	CHEAT1
CURUSE1	0.929				

PEER1	-0.222	0.955			
SURF1	-0.309	0.443	0.949		
CHJUST1	-0.264	0.571	0.591	0.754	
CHEAT1	-0.190	0.443	0.484	0.483	0.668
TEACHER1	0.606	-0.446	-0.353	-0.388	-0.327
SUB2	0.356	-0.102	-0.259	-0.089	-0.184
HON2	0.095	-0.247	-0.216	-0.239	-0.307
PERF2	0.036	0.095	0.076	0.143	0.080
GTEACH2	0.264	-0.163	-0.193	-0.148	-0.168
ASSESS2	0.329	-0.204	-0.242	-0.185	-0.210
CURUSE2	0.403	-0.086	-0.193	-0.101	-0.120
PEER2	-0.239	0.493	0.280	0.333	0.299
SURF2	-0.260	0.321	0.447	0.389	0.311
CHJUST2	-0.225	0.344	0.323	0.381	0.290
CHEAT2	-0.212	0.286	0.274	0.290	0.258
TEACHER2	0.329	-0.204	-0.242	-0.185	-0.210
GENDER	0.013	0.065	-0.001	0.084	0.056
GRADE2	0.029	-0.048	-0.008	-0.012	-0.018

	TEACHER1	SUB2	HON2	PERF2	GTEACH2
TEACHER1	0.807				
SUB2	0.297	0.907			
HON2	0.184	0.197	0.955		
PERF2	-0.031	0.094	0.010	0.929	
GTEACH2	0.318	0.366	0.261	0.101	0.961
ASSESS2	0.398	0.457	0.326	0.126	0.733
CURUSE2	0.283	0.503	0.112	0.146	0.482
PEER2	-0.348	-0.197	-0.314	-0.070	-0.338
SURF2	-0.299	-0.284	-0.210	0.299	-0.286
CHJUST2	-0.286	-0.226	-0.227	0.276	-0.244
CHEAT2	-0.245	-0.282	-0.223	0.073	-0.227
TEACHER2	0.398	0.457	0.326	0.126	0.733
GENDER	0.019	0.120	-0.001	0.095	0.076
GRADE2	0.048	-0.057	-0.012	-0.033	-0.071

	ASSESS2	CURUSE2	PEER2	SURF2	CHJUST2
ASSESS2	0.927				
CURUSE2	0.602	0.909			
PEER2	-0.422	-0.274	0.916		
SURF2	-0.357	-0.336	0.312	0.905	
CHJUST2	-0.305	-0.263	0.374	0.589	0.765
CHEAT2	-0.283	-0.277	0.371	0.466	0.523
TEACHER2	0.916	0.602	-0.422	-0.357	-0.305
GENDER	0.095	0.066	0.056	0.048	0.074
GRADE2	-0.089	-0.083	-0.039	0.069	0.036

	CHEAT2	TEACHER2	GENDER	GRADE2
CHEAT2	0.666			
TEACHER2	-0.283	0.916		
GENDER	0.039	0.095	0.218	
GRADE2	0.051	-0.089	0.016	0.248

ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	SUB1	HON1	PERF1	GTEACH1	ASSESS1
SUB1	1.000				
HON1	0.234	1.000			
PERF1	0.118	0.013	1.000		
GTEACH1	0.400	0.243	-0.105	1.000	
ASSESS1	0.426	0.259	-0.112	0.795	1.000
CURUSE1	0.508	0.136	0.106	0.605	0.644

PEER1	-0.170	-0.360	0.215	-0.440	-0.468
SURF1	-0.361	-0.306	0.266	-0.349	-0.371
CHJUST1	-0.157	-0.394	0.425	-0.430	-0.457
CHEAT1	-0.322	-0.529	0.238	-0.384	-0.409
TEACHER1	0.463	0.281	-0.121	0.865	0.920
SUB2	0.790	0.181	0.089	0.300	0.319
HON2	0.188	0.739	0.015	0.181	0.192
PERF2	0.097	-0.003	0.333	-0.031	-0.033
GTEACH2	0.351	0.232	-0.006	0.312	0.332
ASSESS2	0.446	0.295	-0.008	0.397	0.423
CURUSE2	0.479	0.113	0.062	0.285	0.303
PEER2	-0.232	-0.341	0.061	-0.349	-0.372
SURF2	-0.287	-0.253	0.236	-0.303	-0.322
CHJUST2	-0.239	-0.287	0.238	-0.314	-0.335
CHEAT2	-0.309	-0.292	0.131	-0.289	-0.308
TEACHER2	0.449	0.297	-0.008	0.400	0.425
GENDER	0.289	-0.068	0.130	0.040	0.043
GRADE2	0.034	-0.015	0.068	0.093	0.099

	CURUSE1	PEER1	SURF1	CHJUST1	CHEAT1
CURUSE1	1.000				
PEER1	-0.235	1.000			
SURF1	-0.329	0.466	1.000		
CHJUST1	-0.315	0.673	0.698	1.000	
CHEAT1	-0.241	0.555	0.608	0.680	1.000
TEACHER1	0.700	-0.509	-0.404	-0.497	-0.445
SUB2	0.388	-0.110	-0.279	-0.107	-0.236
HON2	0.101	-0.258	-0.227	-0.281	-0.384
PERF2	0.039	0.101	0.081	0.171	0.101
GTEACH2	0.279	-0.170	-0.202	-0.174	-0.210
ASSESS2	0.355	-0.217	-0.258	-0.222	-0.267
CURUSE2	0.439	-0.093	-0.207	-0.122	-0.155
PEER2	-0.259	0.527	0.301	0.401	0.383
SURF2	-0.284	0.345	0.482	0.471	0.400
CHJUST2	-0.267	0.402	0.379	0.501	0.406
CHEAT2	-0.269	0.359	0.345	0.410	0.387
TEACHER2	0.357	-0.218	-0.259	-0.223	-0.269
GENDER	0.029	0.142	-0.003	0.208	0.146
GRADE2	0.060	-0.099	-0.016	-0.027	-0.045

ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES, CONTINUED

	TEACHER1	SUB2	HON2	PERF2	GTEACH2
TEACHER1	1.000				
SUB2	0.347	1.000			
HON2	0.209	0.212	1.000		
PERF2	-0.036	0.103	0.010	1.000	
GTEACH2	0.361	0.392	0.272	0.107	1.000
ASSESS2	0.460	0.498	0.346	0.136	0.776
CURUSE2	0.330	0.554	0.120	0.159	0.516
PEER2	-0.404	-0.216	-0.336	-0.076	-0.360
SURF2	-0.350	-0.313	-0.226	0.326	-0.306
CHJUST2	-0.364	-0.271	-0.265	0.328	-0.285
CHEAT2	-0.335	-0.363	-0.280	0.092	-0.283
TEACHER2	0.463	0.501	0.348	0.137	0.781
GENDER	0.046	0.269	-0.002	0.211	0.167
GRADE2	0.107	-0.120	-0.025	-0.069	-0.146

	ASSESS2	CURUSE2	PEER2	SURF2	CHJUST2
ASSESS2	1.000				
CURUSE2	0.656	1.000			
PEER2	-0.458	-0.300	1.000		
SURF2	-0.390	-0.371	0.342	1.000	
CHJUST2	-0.362	-0.315	0.446	0.708	1.000
CHEAT2	-0.361	-0.356	0.475	0.600	0.733
TEACHER2	0.994	0.660	-0.461	-0.392	-0.364
GENDER	0.212	0.149	0.126	0.108	0.180
GRADE2	-0.186	-0.176	-0.082	0.145	0.082

	CHEAT2	TEACHER2	GENDER	GRADE2
CHEAT2	1.000			
TEACHER2	-0.363	1.000		
GENDER	0.104	0.213	1.000	
GRADE2	0.126	-0.187	0.070	1.000

Appendix AF:

Indirect effects (standardized) in the longitudinal model

```
STANDARDIZED TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS

STDYX Standardization
                                                      Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Effects from CURUSE2 to CHEAT2

  Total               -0.169      0.108     -1.566      0.117
  Total indirect      -0.109      0.074     -1.458      0.145

  Specific indirect

    CHEAT2
    PEER2
    CURUSE2           -0.008      0.015     -0.534      0.594

    CHEAT2
    CHJUST2
    CURUSE2           -0.093      0.070     -1.322      0.186

    CHEAT2
    CHJUST2
    PEER2
    CURUSE2           -0.008      0.014     -0.546      0.585

  Direct
    CHEAT2
    CURUSE2           -0.060      0.103     -0.584      0.559


Effects from PERF2 to CHEAT2

  Total                0.083      0.081      1.025      0.305
  Total indirect       0.179      0.070      2.546      0.011

  Specific indirect

    CHEAT2
    PEER2
    PERF2             -0.017      0.015     -1.141      0.254

    CHEAT2
    CHJUST2
    PERF2              0.212      0.067      3.166      0.002

    CHEAT2
    CHJUST2
    PEER2
    PERF2             -0.016      0.013     -1.217      0.224

  Direct
    CHEAT2
    PERF2             -0.096      0.090     -1.067      0.286
```

```
Effects from PEER2 to CHEAT2

  Total                 0.310      0.091      3.415      0.001
  Total indirect        0.153      0.066      2.334      0.020

  Specific indirect

    CHEAT2
    CHJUST2
    PEER2               0.153      0.066      2.334      0.020

  Direct
    CHEAT2
    PEER2               0.157      0.094      1.664      0.096


Effects from SUB2 to CHEAT2

  Total                -0.291      0.077     -3.782      0.000
  Total indirect       -0.156      0.062     -2.496      0.013

  Specific indirect

    CHEAT2
    PERF2
    SUB2               -0.002      0.008     -0.199      0.842

    CHEAT2
    CURUSE2
    SUB2               -0.025      0.043     -0.582      0.560

    CHEAT2
    CHJUST2
    SUB2               -0.071      0.058     -1.219      0.223

    CHEAT2
    TEACHER2
    SUB2                0.029      0.033      0.880      0.379

    CHEAT2
    PEER2
    PERF2
    SUB2                0.000      0.001     -0.200      0.842

    CHEAT2
    PEER2
    CURUSE2
    SUB2               -0.003      0.006     -0.532      0.595

    CHEAT2
    PEER2
    TEACHER2
    SUB2               -0.016      0.011     -1.384      0.166

    CHEAT2
    CHJUST2
    PERF2
    SUB2                0.003      0.017      0.202      0.840

    CHEAT2
    CHJUST2
    CURUSE2
```

```
    SUB2              -0.038      0.030     -1.294      0.196

    CHEAT2
    CHJUST2
    TEACHER2
    SUB2              -0.014      0.022     -0.627      0.531

    CHEAT2
    CHJUST2
    PEER2
    PERF2
    SUB2               0.000      0.001     -0.200      0.842

    CHEAT2
    CHJUST2
    PEER2
    CURUSE2
    SUB2              -0.003      0.006     -0.544      0.586

    CHEAT2
    CHJUST2
    PEER2
    TEACHER2
    SUB2              -0.015      0.009     -1.747      0.081

  Direct
    CHEAT2
    SUB2              -0.135      0.087     -1.561      0.119


Effects from HON2 to CHEAT2

  Total               -0.106      0.081     -1.308      0.191
  Total indirect      -0.072      0.057     -1.264      0.206

  Specific indirect

    CHEAT2
    PEER2
    HON2              -0.017      0.015     -1.105      0.269

    CHEAT2
    CHJUST2
    HON2              -0.026      0.052     -0.505      0.613

    CHEAT2
    TEACHER2
    HON2               0.022      0.025      0.870      0.384

    CHEAT2
    PEER2
    TEACHER2
    HON2              -0.012      0.009     -1.359      0.174

    CHEAT2
    CHJUST2
    PEER2
    HON2              -0.016      0.013     -1.229      0.219

    CHEAT2
    CHJUST2
    TEACHER2
    HON2              -0.011      0.017     -0.623      0.533
```

```
    CHEAT2
    CHJUST2
    PEER2
    TEACHER2
    HON2              -0.012      0.007     -1.703      0.089

  Direct
    CHEAT2
    HON2              -0.034      0.080     -0.425      0.671


Effects from TEACHER2 to CHEAT2

  Total               -0.057      0.113     -0.504      0.614
  Total indirect      -0.160      0.085     -1.890      0.059

  Specific indirect

    CHEAT2
    PEER2
    TEACHER2          -0.056      0.038     -1.476      0.140

    CHEAT2
    CHJUST2
    TEACHER2          -0.049      0.077     -0.635      0.525

    CHEAT2
    CHJUST2
    PEER2
    TEACHER2          -0.055      0.028     -1.940      0.052

  Direct
    CHEAT2
    TEACHER2           0.103      0.114      0.903      0.366


Effects from CHJUST1 to CHEAT2

  Total                0.200      0.060      3.334      0.001
  Total indirect       0.200      0.060      3.334      0.001

  Specific indirect

    CHEAT2
    CHEAT1
    CHJUST1            0.025      0.041      0.614      0.539

    CHEAT2
    CHJUST2
    CHJUST1            0.175      0.058      3.022      0.003


Effects from PEER1 to CHEAT2

  Total                0.215      0.047      4.590      0.000
  Total indirect       0.215      0.047      4.590      0.000

  Specific indirect

    CHEAT2
    CHEAT1
    PEER1              0.005      0.010      0.462      0.644
```

```
    CHEAT2
    PEER2
    PEER1                0.061       0.037       1.617       0.106

    CHEAT2
    CHEAT1
    CHJUST1
    PEER1                0.011       0.019       0.609       0.543

    CHEAT2
    CHJUST2
    CHJUST1
    PEER1                0.079       0.031       2.559       0.010

    CHEAT2
    CHJUST2
    PEER2
    PEER1                0.059       0.028       2.143       0.032


Effects from PERF1 to CHEAT2

  Total                  0.120       0.037       3.195       0.001
  Total indirect         0.120       0.037       3.195       0.001

  Specific indirect

    CHEAT2
    CHEAT1
    PERF1                0.001       0.006       0.203       0.839

    CHEAT2
    PERF2
    PERF1               -0.030       0.029      -1.028       0.304

    CHEAT2
    CHEAT1
    PEER1
    PERF1                0.001       0.001       0.436       0.663

    CHEAT2
    CHEAT1
    CHJUST1
    PERF1                0.008       0.014       0.606       0.545

    CHEAT2
    PEER2
    PEER1
    PERF1                0.007       0.007       1.033       0.302

    CHEAT2
    PEER2
    PERF2
    PERF1               -0.005       0.005      -1.092       0.275

    CHEAT2
    CHJUST2
    CHJUST1
    PERF1                0.058       0.024       2.460       0.014
```

```
    CHEAT2
    CHJUST2
    PERF2
    PERF1              0.067      0.027      2.440      0.015

    CHEAT2
    CHEAT1
    CHJUST1
    PEER1
    PERF1              0.001      0.002      0.559      0.576

    CHEAT2
    CHJUST2
    CHJUST1
    PEER1
    PERF1              0.010      0.008      1.234      0.217

    CHEAT2
    CHJUST2
    PEER2
    PEER1
    PERF1              0.007      0.006      1.141      0.254

    CHEAT2
    CHJUST2
    PEER2
    PERF2
    PERF1             -0.005      0.004     -1.158      0.247


Effects from TEACHER1 to CHEAT2

  Total               -0.151      0.056     -2.709      0.007
  Total indirect      -0.151      0.056     -2.709      0.007

  Specific indirect

    CHEAT2
    CHEAT1
    TEACHER1          -0.005      0.010     -0.463      0.643

    CHEAT2
    TEACHER2
    TEACHER1           0.035      0.039      0.884      0.377

    CHEAT2
    CHEAT1
    PEER1
    TEACHER1          -0.002      0.005     -0.459      0.646

    CHEAT2
    CHEAT1
    CHJUST1
    TEACHER1          -0.001      0.004     -0.305      0.760

    CHEAT2
    PEER2
    PEER1
    TEACHER1          -0.033      0.022     -1.505      0.132
```

```
    CHEAT2
    PEER2
    TEACHER2
    TEACHER1             -0.019      0.013     -1.423      0.155

    CHEAT2
    CHJUST2
    CHJUST1
    TEACHER1             -0.009      0.025     -0.350      0.726

    CHEAT2
    CHJUST2
    TEACHER2
    TEACHER1             -0.016      0.026     -0.631      0.528

    CHEAT2
    CHEAT1
    CHJUST1
    PEER1
    TEACHER1             -0.006      0.010     -0.600      0.549

    CHEAT2
    CHJUST2
    CHJUST1
    PEER1
    TEACHER1             -0.043      0.021     -2.084      0.037

    CHEAT2
    CHJUST2
    PEER2
    PEER1
    TEACHER1             -0.032      0.017     -1.903      0.057

    CHEAT2
    CHJUST2
    PEER2
    TEACHER2
    TEACHER1             -0.018      0.010     -1.828      0.068


Effects from CURUSE1 to CHEAT2

  Total                -0.045      0.048     -0.946      0.344
  Total indirect       -0.045      0.048     -0.946      0.344

  Specific indirect

    CHEAT2
    CHEAT1
    CURUSE1               0.007      0.013      0.549      0.583

    CHEAT2
    CURUSE2
    CURUSE1              -0.017      0.029     -0.580      0.562

    CHEAT2
    CHEAT1
    PEER1
    CURUSE1               0.001      0.002      0.437      0.662
```

```
    CHEAT2
    CHEAT1
    CHJUST1
    CURUSE1            -0.005      0.009     -0.573      0.567

    CHEAT2
    PEER2
    PEER1
    CURUSE1             0.010      0.010      1.007      0.314

    CHEAT2
    PEER2
    CURUSE2
    CURUSE1            -0.002      0.004     -0.529      0.597

    CHEAT2
    CHJUST2
    CHJUST1
    CURUSE1            -0.034      0.024     -1.455      0.146

    CHEAT2
    CHJUST2
    CURUSE2
    CURUSE1            -0.026      0.021     -1.271      0.204

    CHEAT2
    CHEAT1
    CHJUST1
    PEER1
    CURUSE1             0.002      0.003      0.546      0.585

    CHEAT2
    CHJUST2
    CHJUST1
    PEER1
    CURUSE1             0.013      0.012      1.118      0.264

    CHEAT2
    CHJUST2
    PEER2
    PEER1
    CURUSE1             0.010      0.009      1.105      0.269

    CHEAT2
    CHJUST2
    PEER2
    CURUSE2
    CURUSE1            -0.002      0.004     -0.541      0.588


Effects from SUB1 to CHEAT2

  Total                -0.315      0.059     -5.302      0.000
  Total indirect       -0.315      0.059     -5.302      0.000

  Specific indirect

    CHEAT2
    CHEAT1
    SUB1               -0.013      0.021     -0.602      0.547
```

```
CHEAT2
   SUB2
   SUB1                 -0.105       0.068      -1.557       0.120

   CHEAT2
   CHEAT1
   PERF1
   SUB1                  0.000       0.000       0.199       0.842

   CHEAT2
   CHEAT1
   CURUSE1
   SUB1                  0.004       0.007       0.548       0.584

   CHEAT2
   CHEAT1
   CHJUST1
   SUB1                  0.001       0.002       0.254       0.800

   CHEAT2
   CHEAT1
   TEACHER1
   SUB1                 -0.002       0.005      -0.462       0.644

   CHEAT2
   PERF2
   PERF1
   SUB1                 -0.003       0.004      -0.725       0.469

   CHEAT2
   PERF2
   SUB2
   SUB1                 -0.001       0.006      -0.199       0.842

   CHEAT2
   CURUSE2
   CURUSE1
   SUB1                 -0.009       0.016      -0.579       0.563

   CHEAT2
   CURUSE2
   SUB2
   SUB1                 -0.019       0.033      -0.582       0.561

   CHEAT2
   CHJUST2
   CHJUST1
   SUB1                  0.004       0.015       0.278       0.781

   CHEAT2
   CHJUST2
   SUB2
   SUB1                 -0.055       0.046      -1.217       0.224

   CHEAT2
   TEACHER2
   TEACHER1
   SUB1                  0.015       0.017       0.875       0.381

   CHEAT2
   TEACHER2
   SUB2
```

```
SUB1                 0.022      0.026      0.880      0.379

CHEAT2
CHEAT1
PEER1
PERF1
SUB1                 0.000      0.000      0.401      0.689

CHEAT2
CHEAT1
PEER1
CURUSE1
SUB1                 0.000      0.001      0.436      0.663

CHEAT2
CHEAT1
PEER1
TEACHER1
SUB1                -0.001      0.002     -0.458      0.647

CHEAT2
CHEAT1
CHJUST1
PERF1
SUB1                 0.001      0.001      0.520      0.603

CHEAT2
CHEAT1
CHJUST1
CURUSE1
SUB1                -0.003      0.005     -0.572      0.568

CHEAT2
CHEAT1
CHJUST1
TEACHER1
SUB1                -0.001      0.002     -0.305      0.761

CHEAT2
PEER2
PEER1
PERF1
SUB1                 0.001      0.001      0.726      0.468

CHEAT2
PEER2
PEER1
CURUSE1
SUB1                 0.005      0.005      1.000      0.317

CHEAT2
PEER2
PEER1
TEACHER1
SUB1                -0.015      0.010     -1.459      0.145

CHEAT2
PEER2
PERF2
PERF1
SUB1                 0.000      0.001     -0.746      0.455
```

```
CHEAT2
PEER2
PERF2
SUB2
SUB1                 0.000      0.001     -0.200      0.842

CHEAT2
PEER2
CURUSE2
CURUSE1
SUB1                -0.001      0.002     -0.528      0.597

CHEAT2
PEER2
CURUSE2
SUB2
SUB1                -0.003      0.005     -0.532      0.595

CHEAT2
PEER2
TEACHER2
TEACHER1
SUB1                -0.008      0.006     -1.386      0.166

CHEAT2
PEER2
TEACHER2
SUB2
SUB1                -0.012      0.009     -1.381      0.167

CHEAT2
CHJUST2
CHJUST1
PERF1
SUB1                 0.005      0.005      0.937      0.349

CHEAT2
CHJUST2
CHJUST1
CURUSE1
SUB1                -0.019      0.013     -1.433      0.152

CHEAT2
CHJUST2
CHJUST1
TEACHER1
SUB1                -0.004      0.011     -0.349      0.727

CHEAT2
CHJUST2
PERF2
PERF1
SUB1                 0.006      0.006      0.943      0.346

CHEAT2
CHJUST2
PERF2
SUB2
SUB1                 0.003      0.013      0.202      0.840
```

CHEAT2				
CHJUST2				
CURUSE2				
CURUSE1				
SUB1	-0.014	0.011	-1.258	0.209
CHEAT2				
CHJUST2				
CURUSE2				
SUB2				
SUB1	-0.030	0.023	-1.291	0.197
CHEAT2				
CHJUST2				
TEACHER2				
TEACHER1				
SUB1	-0.007	0.012	-0.628	0.530
CHEAT2				
CHJUST2				
TEACHER2				
SUB2				
SUB1	-0.011	0.017	-0.627	0.531
CHEAT2				
CHEAT1				
CHJUST1				
PEER1				
PERF1				
SUB1	0.000	0.000	0.490	0.624
CHEAT2				
CHEAT1				
CHJUST1				
PEER1				
CURUSE1				
SUB1	0.001	0.002	0.545	0.586
CHEAT2				
CHEAT1				
CHJUST1				
PEER1				
TEACHER1				
SUB1	-0.003	0.005	-0.597	0.551
CHEAT2				
CHJUST2				
CHJUST1				
PEER1				
PERF1				
SUB1	0.001	0.001	0.786	0.432
CHEAT2				
CHJUST2				
CHJUST1				
PEER1				
CURUSE1				
SUB1	0.007	0.006	1.108	0.268

```
CHEAT2
CHJUST2
CHJUST1
PEER1
TEACHER1
SUB1                -0.019      0.010      -1.966      0.049

CHEAT2
CHJUST2
PEER2
PEER1
PERF1
SUB1                 0.001      0.001       0.760      0.447

CHEAT2
CHJUST2
PEER2
PEER1
CURUSE1
SUB1                 0.005      0.005       1.096      0.273

CHEAT2
CHJUST2
PEER2
PEER1
TEACHER1
SUB1                -0.014      0.008      -1.810      0.070

CHEAT2
CHJUST2
PEER2
PERF2
PERF1
SUB1                 0.000      0.001      -0.767      0.443

CHEAT2
CHJUST2
PEER2
PERF2
SUB2
SUB1                 0.000      0.001      -0.200      0.841

CHEAT2
CHJUST2
PEER2
CURUSE2
CURUSE1
SUB1                -0.001      0.002      -0.540      0.589

CHEAT2
CHJUST2
PEER2
CURUSE2
SUB2
SUB1                -0.003      0.005      -0.544      0.586

CHEAT2
CHJUST2
PEER2
TEACHER2
TEACHER1
SUB1                -0.008      0.005      -1.750      0.080
```

```
    CHEAT2
    CHJUST2
    PEER2
    TEACHER2
    SUB2
    SUB1                -0.012      0.007     -1.740      0.082


Effects from HON1 to CHEAT2

  Total                -0.205      0.059     -3.477      0.001
  Total indirect       -0.205      0.059     -3.477      0.001

  Specific indirect

    CHEAT2
    CHEAT1
    HON1                -0.014      0.023     -0.606      0.545

    CHEAT2
    HON2
    HON1                -0.025      0.059     -0.425      0.671

    CHEAT2
    CHEAT1
    PEER1
    HON1                -0.001      0.002     -0.457      0.648

    CHEAT2
    CHEAT1
    CHJUST1
    HON1                -0.005      0.008     -0.598      0.550

    CHEAT2
    CHEAT1
    TEACHER1
    HON1                -0.001      0.002     -0.459      0.646

    CHEAT2
    PEER2
    PEER1
    HON1                -0.013      0.010     -1.392      0.164

    CHEAT2
    PEER2
    HON2
    HON1                -0.012      0.011     -1.103      0.270

    CHEAT2
    CHJUST2
    CHJUST1
    HON1                -0.034      0.017     -1.965      0.049

    CHEAT2
    CHJUST2
    HON2
    HON1                -0.020      0.039     -0.505      0.614

    CHEAT2
    TEACHER2
    TEACHER1
    HON1                 0.006      0.007      0.843      0.399
```

CHEAT2 TEACHER2 HON2 HON1	0.016	0.019	0.869	0.385
CHEAT2 CHEAT1 PEER1 TEACHER1 HON1	0.000	0.001	-0.452	0.651
CHEAT2 CHEAT1 CHJUST1 PEER1 HON1	-0.003	0.004	-0.594	0.552
CHEAT2 CHEAT1 CHJUST1 TEACHER1 HON1	0.000	0.001	-0.305	0.761
CHEAT2 PEER2 PEER1 TEACHER1 HON1	-0.006	0.004	-1.320	0.187
CHEAT2 PEER2 TEACHER2 TEACHER1 HON1	-0.003	0.003	-1.267	0.205
CHEAT2 PEER2 TEACHER2 HON2 HON1	-0.009	0.007	-1.353	0.176
CHEAT2 CHJUST2 CHJUST1 PEER1 HON1	-0.018	0.009	-1.880	0.060
CHEAT2 CHJUST2 CHJUST1 TEACHER1 HON1	-0.002	0.004	-0.349	0.727
CHEAT2 CHJUST2 PEER2 PEER1 HON1	-0.013	0.008	-1.689	0.091

CHEAT2 CHJUST2 PEER2 HON2 HON1	-0.012	0.010	-1.228	0.219
CHEAT2 CHJUST2 TEACHER2 TEACHER1 HON1	-0.003	0.005	-0.615	0.538
CHEAT2 CHJUST2 TEACHER2 HON2 HON1	-0.008	0.013	-0.622	0.534
CHEAT2 CHEAT1 CHJUST1 PEER1 TEACHER1 HON1	-0.001	0.002	-0.586	0.558
CHEAT2 CHJUST2 CHJUST1 PEER1 TEACHER1 HON1	-0.007	0.005	-1.658	0.097
CHEAT2 CHJUST2 PEER2 PEER1 TEACHER1 HON1	-0.006	0.004	-1.564	0.118
CHEAT2 CHJUST2 PEER2 TEACHER2 TEACHER1 HON1	-0.003	0.002	-1.528	0.126
CHEAT2 CHJUST2 PEER2 TEACHER2 HON2 HON1	-0.009	0.005	-1.692	0.091

```
Effects from CURUSE2 to SURF2

  Total                -0.201      0.104     -1.925      0.054
  Total indirect       -0.074      0.053     -1.388      0.165

  Specific indirect

    SURF2
    CHJUST2
    CURUSE2            -0.068      0.053     -1.302      0.193

    SURF2
    CHJUST2
    PEER2
    CURUSE2            -0.006      0.011     -0.545      0.586

  Direct
    SURF2
    CURUSE2            -0.127      0.101     -1.259      0.208

Effects from PERF2 to SURF2

  Total                 0.331      0.078      4.246      0.000
  Total indirect        0.144      0.049      2.939      0.003

  Specific indirect

    SURF2
    CHJUST2
    PERF2               0.156      0.051      3.088      0.002

    SURF2
    CHJUST2
    PEER2
    PERF2              -0.012      0.010     -1.216      0.224

  Direct
    SURF2
    PERF2               0.187      0.084      2.212      0.027

Effects from SUB2 to SURF2

  Total                -0.199      0.082     -2.440      0.015
  Total indirect       -0.174      0.057     -3.021      0.003

  Specific indirect

    SURF2
    PERF2
    SUB2                0.003      0.015      0.201      0.841

    SURF2
    CURUSE2
    SUB2               -0.053      0.043     -1.232      0.218

    SURF2
    CHJUST2
    SUB2               -0.052      0.044     -1.197      0.231

    SURF2
    TEACHER2
    SUB2               -0.022      0.029     -0.751      0.453
```

```
    SURF2
    CHJUST2
    PERF2
    SUB2                 0.003      0.012      0.202      0.840

    SURF2
    CHJUST2
    CURUSE2
    SUB2                -0.028      0.022     -1.275      0.202

    SURF2
    CHJUST2
    TEACHER2
    SUB2                -0.010      0.016     -0.625      0.532

    SURF2
    CHJUST2
    PEER2
    PERF2
    SUB2                 0.000      0.001     -0.200      0.842

    SURF2
    CHJUST2
    PEER2
    CURUSE2
    SUB2                -0.002      0.004     -0.544      0.587

    SURF2
    CHJUST2
    PEER2
    TEACHER2
    SUB2                -0.011      0.006     -1.747      0.081

  Direct
    SURF2
    SUB2                -0.025      0.088     -0.290      0.771


Effects from TEACHER2 to SURF2

  Total                 -0.154      0.104     -1.478      0.139
  Total indirect        -0.077      0.056     -1.361      0.174

  Specific indirect

    SURF2
    CHJUST2
    TEACHER2            -0.036      0.057     -0.633      0.527

    SURF2
    CHJUST2
    PEER2
    TEACHER2            -0.040      0.021     -1.939      0.052

  Direct
    SURF2
    TEACHER2            -0.077      0.101     -0.766      0.444


Effects from CHJUST1 to SURF2

  Total                  0.295      0.068      4.358      0.000
  Total indirect         0.295      0.068      4.358      0.000
```

```
  Specific indirect

    SURF2
    SURF1
    CHJUST1              0.165      0.066      2.519      0.012

    SURF2
    CHJUST2
    CHJUST1              0.129      0.046      2.800      0.005


Effects from PERF1 to SURF2

  Total                  0.230      0.047      4.846      0.000
  Total indirect         0.230      0.047      4.846      0.000

  Specific indirect

    SURF2
    SURF1
    PERF1                0.007      0.024      0.273      0.785

    SURF2
    PERF2
    PERF1                0.059      0.030      1.946      0.052

    SURF2
    SURF1
    CHJUST1
    PERF1                0.055      0.026      2.122      0.034

    SURF2
    CHJUST2
    CHJUST1
    PERF1                0.043      0.018      2.345      0.019

    SURF2
    CHJUST2
    PERF2
    PERF1                0.049      0.021      2.402      0.016

    SURF2
    SURF1
    CHJUST1
    PEER1
    PERF1                0.009      0.008      1.199      0.230

    SURF2
    CHJUST2
    CHJUST1
    PEER1
    PERF1                0.007      0.006      1.218      0.223

    SURF2
    CHJUST2
    PEER2
    PEER1
    PERF1                0.005      0.005      1.142      0.253
```

```
    SURF2
    CHJUST2
    PEER2
    PERF2
    PERF1              -0.004      0.003     -1.158      0.247


Effects from TEACHER1 to SURF2

  Total              -0.137      0.057     -2.423      0.015
  Total indirect     -0.137      0.057     -2.423      0.015

  Specific indirect

    SURF2
    SURF1
    TEACHER1            0.025      0.036      0.710      0.477

    SURF2
    TEACHER2
    TEACHER1           -0.026      0.034     -0.764      0.445

    SURF2
    SURF1
    CHJUST1
    TEACHER1           -0.008      0.024     -0.346      0.729

    SURF2
    CHJUST2
    CHJUST1
    TEACHER1           -0.006      0.019     -0.350      0.727

    SURF2
    CHJUST2
    TEACHER2
    TEACHER1           -0.012      0.019     -0.629      0.529

    SURF2
    SURF1
    CHJUST1
    PEER1
    TEACHER1           -0.041      0.021     -1.910      0.056

    SURF2
    CHJUST2
    CHJUST1
    PEER1
    TEACHER1           -0.032      0.016     -2.007      0.045

    SURF2
    CHJUST2
    PEER2
    PEER1
    TEACHER1           -0.024      0.013     -1.904      0.057

    SURF2
    CHJUST2
    PEER2
    TEACHER2
    TEACHER1           -0.014      0.007     -1.826      0.068
```

```
Effects from CURUSE1 to SURF2

  Total                -0.100       0.054      -1.857       0.063
  Total indirect       -0.100       0.054      -1.857       0.063

  Specific indirect

    SURF2
    SURF1
    CURUSE1            -0.014       0.030      -0.452       0.651

    SURF2
    CURUSE2
    CURUSE1            -0.036       0.030      -1.218       0.223

    SURF2
    SURF1
    CHJUST1
    CURUSE1            -0.033       0.023      -1.395       0.163

    SURF2
    CHJUST2
    CHJUST1
    CURUSE1            -0.025       0.018      -1.429       0.153

    SURF2
    CHJUST2
    CURUSE2
    CURUSE1            -0.020       0.016      -1.253       0.210

    SURF2
    SURF1
    CHJUST1
    PEER1
    CURUSE1             0.012       0.011       1.086       0.277

    SURF2
    CHJUST2
    CHJUST1
    PEER1
    CURUSE1             0.010       0.009       1.106       0.269

    SURF2
    CHJUST2
    PEER2
    PEER1
    CURUSE1             0.007       0.006       1.105       0.269

    SURF2
    CHJUST2
    PEER2
    CURUSE2
    CURUSE1            -0.002       0.003      -0.541       0.588
```

```
Effects from SUB1 to SURF2

  Total                -0.302      0.063     -4.827      0.000
  Total indirect       -0.302      0.063     -4.827      0.000

  Specific indirect

    SURF2
    SURF1
    SUB1               -0.059      0.031     -1.920      0.055

    SURF2
    SUB2
    SUB1               -0.020      0.068     -0.291      0.771

    SURF2
    SURF1
    PERF1
    SUB1                0.001      0.002      0.262      0.793

    SURF2
    SURF1
    CURUSE1
    SUB1               -0.007      0.016     -0.452      0.651

    SURF2
    SURF1
    CHJUST1
    SUB1                0.004      0.014      0.277      0.782

    SURF2
    SURF1
    TEACHER1
    SUB1                0.011      0.016      0.704      0.481

    SURF2
    PERF2
    PERF1
    SUB1                0.005      0.006      0.902      0.367

    SURF2
    PERF2
    SUB2
    SUB1                0.002      0.012      0.201      0.841

    SURF2
    CURUSE2
    CURUSE1
    SUB1               -0.020      0.016     -1.206      0.228

    SURF2
    CURUSE2
    SUB2
    SUB1               -0.041      0.033     -1.229      0.219

    SURF2
    CHJUST2
    CHJUST1
    SUB1                0.003      0.011      0.278      0.781
```

SURF2 CHJUST2 SUB2 SUB1	-0.041	0.034	-1.195	0.232
SURF2 TEACHER2 TEACHER1 SUB1	-0.011	0.015	-0.758	0.448
SURF2 TEACHER2 SUB2 SUB1	-0.017	0.022	-0.750	0.453
SURF2 SURF1 CHJUST1 PERF1 SUB1	0.005	0.005	0.914	0.361
SURF2 SURF1 CHJUST1 CURUSE1 SUB1	-0.018	0.013	-1.375	0.169
SURF2 SURF1 CHJUST1 TEACHER1 SUB1	-0.004	0.011	-0.345	0.730
SURF2 CHJUST2 CHJUST1 PERF1 SUB1	0.004	0.004	0.929	0.353
SURF2 CHJUST2 CHJUST1 CURUSE1 SUB1	-0.014	0.010	-1.408	0.159
SURF2 CHJUST2 CHJUST1 TEACHER1 SUB1	-0.003	0.008	-0.349	0.727
SURF2 CHJUST2 PERF2 PERF1 SUB1	0.004	0.005	0.939	0.348
SURF2 CHJUST2 PERF2 SUB2 SUB1	0.002	0.010	0.202	0.840

```
SURF2
CHJUST2
CURUSE2
CURUSE1
SUB1                -0.011      0.009     -1.240      0.215

SURF2
CHJUST2
CURUSE2
SUB2
SUB1                -0.022      0.017     -1.272      0.204

SURF2
CHJUST2
TEACHER2
TEACHER1
SUB1                -0.005      0.009     -0.626      0.532

SURF2
CHJUST2
TEACHER2
SUB2
SUB1                -0.008      0.013     -0.624      0.532

SURF2
SURF1
CHJUST1
PEER1
PERF1
SUB1                 0.001      0.001      0.776      0.438

SURF2
SURF1
CHJUST1
PEER1
CURUSE1
SUB1                 0.007      0.006      1.078      0.281

SURF2
SURF1
CHJUST1
PEER1
TEACHER1
SUB1                -0.018      0.010     -1.818      0.069

SURF2
CHJUST2
CHJUST1
PEER1
PERF1
SUB1                 0.001      0.001      0.781      0.435

SURF2
CHJUST2
CHJUST1
PEER1
CURUSE1
SUB1                 0.005      0.005      1.096      0.273
```

```
SURF2
CHJUST2
CHJUST1
PEER1
TEACHER1
SUB1                -0.014      0.007     -1.900      0.057

SURF2
CHJUST2
PEER2
PEER1
PERF1
SUB1                 0.000      0.001      0.760      0.447

SURF2
CHJUST2
PEER2
PEER1
CURUSE1
SUB1                 0.004      0.004      1.096      0.273

SURF2
CHJUST2
PEER2
PEER1
TEACHER1
SUB1                -0.010      0.006     -1.812      0.070

SURF2
CHJUST2
PEER2
PERF2
PERF1
SUB1                 0.000      0.000     -0.766      0.444

SURF2
CHJUST2
PEER2
PERF2
SUB2
SUB1                 0.000      0.001     -0.200      0.842

SURF2
CHJUST2
PEER2
CURUSE2
CURUSE1
SUB1                -0.001      0.002     -0.540      0.589

SURF2
CHJUST2
PEER2
CURUSE2
SUB2
SUB1                -0.002      0.003     -0.544      0.587

SURF2
CHJUST2
PEER2
TEACHER2
TEACHER1
SUB1                -0.006      0.003     -1.749      0.080
```

```
    SURF2
    CHJUST2
    PEER2
    TEACHER2
    SUB2
    SUB1              -0.009      0.005     -1.741      0.082


Effects from SUB2 to CHJUST2

  Total               -0.217      0.079     -2.761      0.006
  Total indirect      -0.106      0.046     -2.291      0.022

  Specific indirect

    CHJUST2
    PERF2
    SUB2               0.005      0.026      0.202      0.840

    CHJUST2
    CURUSE2
    SUB2              -0.060      0.045     -1.328      0.184

    CHJUST2
    TEACHER2
    SUB2              -0.021      0.034     -0.635      0.525

    CHJUST2
    PEER2
    PERF2
    SUB2               0.000      0.002     -0.200      0.842

    CHJUST2
    PEER2
    CURUSE2
    SUB2              -0.005      0.009     -0.546      0.585

    CHJUST2
    PEER2
    TEACHER2
    SUB2              -0.024      0.013     -1.838      0.066

  Direct
    CHJUST2
    SUB2              -0.111      0.089     -1.247      0.212


Effects from CURUSE2 to CHJUST2

  Total               -0.157      0.108     -1.456      0.146
  Total indirect      -0.012      0.022     -0.548      0.584

  Specific indirect

    CHJUST2
    PEER2
    CURUSE2           -0.012      0.022     -0.548      0.584

  Direct
    CHJUST2
    CURUSE2           -0.145      0.107     -1.359      0.174
```

```
Effects from PERF2 to CHJUST2

  Total                 0.305      0.080      3.837      0.000
  Total indirect       -0.025      0.020     -1.247      0.213

  Specific indirect

    CHJUST2
    PEER2
    PERF2              -0.025      0.020     -1.247      0.213

  Direct
    CHJUST2
    PERF2               0.331      0.080      4.142      0.000


Effects from TEACHER2 to CHJUST2

  Total                -0.163      0.113     -1.438      0.150
  Total indirect       -0.086      0.041     -2.068      0.039

  Specific indirect

    CHJUST2
    PEER2
    TEACHER2           -0.086      0.041     -2.068      0.039

  Direct
    CHJUST2
    TEACHER2           -0.077      0.119     -0.644      0.520


Effects from HON2 to CHJUST2

  Total                -0.102      0.077     -1.330      0.183
  Total indirect       -0.061      0.031     -1.947      0.052

  Specific indirect

    CHJUST2
    PEER2
    HON2               -0.026      0.020     -1.259      0.208

    CHJUST2
    TEACHER2
    HON2               -0.016      0.026     -0.631      0.528

    CHJUST2
    PEER2
    TEACHER2
    HON2               -0.018      0.010     -1.789      0.074

  Direct
    CHJUST2
    HON2               -0.041      0.081     -0.508      0.612
```

```
Effects from PEER1 to CHJUST2

  Total                0.216      0.051      4.228      0.000
  Total indirect       0.216      0.051      4.228      0.000

  Specific indirect

    CHJUST2
    CHJUST1
    PEER1              0.124      0.046      2.716      0.007

    CHJUST2
    PEER2
    PEER1              0.093      0.040      2.325      0.020


Effects from PERF1 to CHJUST2

  Total                0.213      0.047      4.501      0.000
  Total indirect       0.213      0.047      4.501      0.000

  Specific indirect

    CHJUST2
    CHJUST1
    PERF1              0.091      0.035      2.612      0.009

    CHJUST2
    PERF2
    PERF1              0.104      0.037      2.829      0.005

    CHJUST2
    CHJUST1
    PEER1
    PERF1              0.015      0.012      1.252      0.211

    CHJUST2
    PEER2
    PEER1
    PERF1              0.011      0.010      1.167      0.243

    CHJUST2
    PEER2
    PERF2
    PERF1             -0.008      0.007     -1.185      0.236


Effects from TEACHER1 to CHJUST2

  Total               -0.186      0.058     -3.187      0.001
  Total indirect      -0.186      0.058     -3.187      0.001

  Specific indirect

    CHJUST2
    CHJUST1
    TEACHER1          -0.014      0.039     -0.350      0.726

    CHJUST2
    TEACHER2
    TEACHER1          -0.026      0.040     -0.640      0.522
```

```
    CHJUST2
    CHJUST1
    PEER1
    TEACHER1           -0.067      0.031     -2.167      0.030

    CHJUST2
    PEER2
    PEER1
    TEACHER1           -0.050      0.025     -2.027      0.043

    CHJUST2
    PEER2
    TEACHER2
    TEACHER1           -0.029      0.015     -1.938      0.053


Effects from CURUSE1 to CHJUST2

  Total                -0.063      0.051     -1.233      0.217
  Total indirect       -0.063      0.051     -1.233      0.217

  Specific indirect

    CHJUST2
    CHJUST1
    CURUSE1            -0.054      0.036     -1.483      0.138

    CHJUST2
    CURUSE2
    CURUSE1            -0.041      0.032     -1.304      0.192

    CHJUST2
    CHJUST1
    PEER1
    CURUSE1             0.020      0.018      1.130      0.258

    CHJUST2
    PEER2
    PEER1
    CURUSE1             0.015      0.013      1.127      0.260

    CHJUST2
    PEER2
    CURUSE2
    CURUSE1            -0.003      0.006     -0.544      0.587


Effects from SUB1 to CHJUST2

  Total                -0.261      0.064     -4.046      0.000
  Total indirect       -0.261      0.064     -4.046      0.000

  Specific indirect

    CHJUST2
    CHJUST1
    SUB1                0.006      0.023      0.278      0.781

    CHJUST2
    SUB2
    SUB1               -0.087      0.070     -1.244      0.213
```

CHJUST2				
CHJUST1				
PERF1				
SUB1	0.008	0.008	0.944	0.345
CHJUST2				
CHJUST1				
CURUSE1				
SUB1	-0.029	0.020	-1.460	0.144
CHJUST2				
CHJUST1				
TEACHER1				
SUB1	-0.006	0.017	-0.349	0.727
CHJUST2				
PERF2				
PERF1				
SUB1	0.009	0.009	0.960	0.337
CHJUST2				
PERF2				
SUB2				
SUB1	0.004	0.020	0.202	0.840
CHJUST2				
CURUSE2				
CURUSE1				
SUB1	-0.022	0.017	-1.290	0.197
CHJUST2				
CURUSE2				
SUB2				
SUB1	-0.047	0.035	-1.324	0.185
CHJUST2				
TEACHER2				
TEACHER1				
SUB1	-0.011	0.018	-0.636	0.525
CHJUST2				
TEACHER2				
SUB2				
SUB1	-0.017	0.026	-0.635	0.526
CHJUST2				
CHJUST1				
PEER1				
PERF1				
SUB1	0.001	0.002	0.789	0.430
CHJUST2				
CHJUST1				
PEER1				
CURUSE1				
SUB1	0.011	0.010	1.120	0.263
CHJUST2				
CHJUST1				
PEER1				
TEACHER1				
SUB1	-0.030	0.015	-2.034	0.042

CHJUST2 PEER2 PEER1 PERF1 SUB1	0.001	0.001	0.767	0.443
CHJUST2 PEER2 PEER1 CURUSE1 SUB1	0.008	0.007	1.118	0.264
CHJUST2 PEER2 PEER1 TEACHER1 SUB1	-0.022	0.012	-1.916	0.055
CHJUST2 PEER2 PERF2 PERF1 SUB1	-0.001	0.001	-0.774	0.439
CHJUST2 PEER2 PERF2 SUB2 SUB1	0.000	0.002	-0.200	0.841
CHJUST2 PEER2 CURUSE2 CURUSE1 SUB1	-0.002	0.004	-0.543	0.587
CHJUST2 PEER2 CURUSE2 SUB2 SUB1	-0.004	0.007	-0.546	0.585
CHJUST2 PEER2 TEACHER2 TEACHER1 SUB1	-0.013	0.007	-1.845	0.065
CHJUST2 PEER2 TEACHER2 SUB2 SUB1	-0.019	0.010	-1.830	0.067

```
Effects from HON1 to CHJUST2

  Total               -0.209      0.060     -3.501      0.000
  Total indirect      -0.209      0.060     -3.501      0.000

  Specific indirect

    CHJUST2
    CHJUST1
    HON1              -0.053      0.026     -2.032      0.042

    CHJUST2
    HON2
    HON1              -0.031      0.060     -0.507      0.612

    CHJUST2
    CHJUST1
    PEER1
    HON1              -0.028      0.014     -1.938      0.053

    CHJUST2
    CHJUST1
    TEACHER1
    HON1              -0.002      0.007     -0.349      0.727

    CHJUST2
    PEER2
    PEER1
    HON1              -0.021      0.012     -1.773      0.076

    CHJUST2
    PEER2
    HON2
    HON1              -0.019      0.015     -1.257      0.209

    CHJUST2
    TEACHER2
    TEACHER1
    HON1              -0.004      0.007     -0.623      0.533

    CHJUST2
    TEACHER2
    HON2
    HON1              -0.012      0.019     -0.630      0.528

    CHJUST2
    CHJUST1
    PEER1
    TEACHER1
    HON1              -0.012      0.007     -1.699      0.089

    CHJUST2
    PEER2
    PEER1
    TEACHER1
    HON1              -0.009      0.005     -1.631      0.103

    CHJUST2
    PEER2
    TEACHER2
    TEACHER1
    HON1              -0.005      0.003     -1.590      0.112
```

```
    CHJUST2
    PEER2
    TEACHER2
    HON2
    HON1               -0.014      0.008     -1.776      0.076


Effects from CURUSE2 to PEER2

  Total                -0.051      0.091     -0.559      0.576
  Total indirect        0.000      0.000      0.000      1.000

  Direct
    PEER2
    CURUSE2            -0.051      0.091     -0.559      0.576


Effects from HON2 to PEER2

  Total                -0.184      0.070     -2.644      0.008
  Total indirect       -0.077      0.031     -2.504      0.012

  Specific indirect

    PEER2
    TEACHER2
    HON2               -0.077      0.031     -2.504      0.012

  Direct
    PEER2
    HON2               -0.107      0.073     -1.462      0.144


Effects from PERF1 to PEER2

  Total                 0.014      0.042      0.324      0.746
  Total indirect        0.014      0.042      0.324      0.746

  Specific indirect

    PEER2
    PEER1
    PERF1               0.047      0.036      1.307      0.191

    PEER2
    PERF2
    PERF1              -0.034      0.024     -1.406      0.160


Effects from TEACHER1 to PEER2

  Total                -0.330      0.068     -4.855      0.000
  Total indirect       -0.330      0.068     -4.855      0.000

  Specific indirect

    PEER2
    PEER1
    TEACHER1           -0.210      0.063     -3.319      0.001

    PEER2
    TEACHER2
    TEACHER1           -0.120      0.040     -2.996      0.003
```

```
Effects from CURUSE1 to PEER2

  Total                  0.048      0.055      0.874      0.382
  Total indirect         0.048      0.055      0.874      0.382

  Specific indirect

    PEER2
    PEER1
    CURUSE1              0.063      0.050      1.256      0.209

    PEER2
    CURUSE2
    CURUSE1             -0.015      0.026     -0.554      0.579


Effects from HON1 to PEER2

  Total                 -0.280      0.056     -4.990      0.000
  Total indirect        -0.280      0.056     -4.990      0.000

  Specific indirect

    PEER2
    PEER1
    HON1                -0.086      0.035     -2.470      0.014

    PEER2
    HON2
    HON1                -0.080      0.055     -1.459      0.144

    PEER2
    PEER1
    TEACHER1
    HON1                -0.037      0.017     -2.122      0.034

    PEER2
    TEACHER2
    TEACHER1
    HON1                -0.021      0.010     -2.045      0.041

    PEER2
    TEACHER2
    HON2
    HON1                -0.057      0.023     -2.470      0.014


Effects from SUB1 to PERF2

  Total                  0.040      0.065      0.615      0.539
  Total indirect         0.040      0.065      0.615      0.539

  Specific indirect

    PERF2
    PERF1
    SUB1                 0.027      0.028      0.992      0.321

    PERF2
    SUB2
    SUB1                 0.012      0.062      0.202      0.840
```

```
Effects from SUB1 to TEACHER2

  Total                0.366      0.052      7.056      0.000
  Total indirect       0.366      0.052      7.056      0.000

  Specific indirect

    TEACHER2
    TEACHER1
    SUB1               0.148      0.037      4.028      0.000

    TEACHER2
    SUB2
    SUB1               0.218      0.053      4.095      0.000


Effects from HON1 to TEACHER2

  Total                0.217      0.047      4.635      0.000
  Total indirect       0.217      0.047      4.635      0.000

  Specific indirect

    TEACHER2
    TEACHER1
    HON1               0.058      0.024      2.466      0.014

    TEACHER2
    HON2
    HON1               0.159      0.045      3.537      0.000


Effects from SUB1 to CURUSE2

  Total                0.479      0.047     10.142      0.000
  Total indirect       0.479      0.047     10.142      0.000

  Specific indirect

    CURUSE2
    CURUSE1
    SUB1               0.155      0.038      4.124      0.000

    CURUSE2
    SUB2
    SUB1               0.324      0.053      6.109      0.000


Effects from CURUSE1 to CHEAT1

  Total                0.090      0.116      0.778      0.437
  Total indirect      -0.043      0.070     -0.619      0.536

  Specific indirect

    CHEAT1
    PEER1
    CURUSE1            0.014      0.022      0.627      0.530

    CHEAT1
    CHJUST1
    CURUSE1           -0.090      0.062     -1.455      0.146
```

```
    CHEAT1
    CHJUST1
    PEER1
    CURUSE1             0.034       0.030       1.130       0.259

  Direct
    CHEAT1
    CURUSE1             0.133       0.116       1.146       0.252


Effects from PERF1 to CHEAT1

  Total                 0.208       0.080       2.592       0.010
  Total indirect        0.187       0.062       3.009       0.003

  Specific indirect

    CHEAT1
    PEER1
    PERF1               0.010       0.016       0.624       0.533

    CHEAT1
    CHJUST1
    PERF1               0.152       0.060       2.551       0.011

    CHEAT1
    CHJUST1
    PEER1
    PERF1               0.025       0.020       1.253       0.210

  Direct
    CHEAT1
    PERF1               0.021       0.094       0.220       0.826


Effects from PEER1 to CHEAT1

  Total                 0.290       0.095       3.046       0.002
  Total indirect        0.207       0.076       2.735       0.006

  Specific indirect

    CHEAT1
    CHJUST1
    PEER1               0.207       0.076       2.735       0.006

  Direct
    CHEAT1
    PEER1               0.083       0.117       0.707       0.479


Effects from SUB1 to CHEAT1

  Total                -0.271       0.070      -3.851       0.000
  Total indirect       -0.041       0.058      -0.704       0.481

  Specific indirect

    CHEAT1
    PERF1
    SUB1                0.002       0.008       0.214       0.831
```

```
  CHEAT1
  CURUSE1
  SUB1                0.072       0.064       1.133       0.257

  CHEAT1
  CHJUST1
  SUB1                0.011       0.039       0.278       0.781

  CHEAT1
  TEACHER1
  SUB1               -0.039       0.059      -0.665       0.506

  CHEAT1
  PEER1
  PERF1
  SUB1                0.001       0.002       0.532       0.595

  CHEAT1
  PEER1
  CURUSE1
  SUB1                0.007       0.012       0.626       0.531

  CHEAT1
  PEER1
  TEACHER1
  SUB1               -0.020       0.029      -0.694       0.488

  CHEAT1
  CHJUST1
  PERF1
  SUB1                0.013       0.014       0.941       0.347

  CHEAT1
  CHJUST1
  CURUSE1
  SUB1               -0.049       0.034      -1.432       0.152

  CHEAT1
  CHJUST1
  TEACHER1
  SUB1               -0.010       0.029      -0.350       0.727

  CHEAT1
  CHJUST1
  PEER1
  PERF1
  SUB1                0.002       0.003       0.790       0.430

  CHEAT1
  CHJUST1
  PEER1
  CURUSE1
  SUB1                0.018       0.016       1.120       0.263

  CHEAT1
  CHJUST1
  PEER1
  TEACHER1
  SUB1               -0.050       0.024      -2.039       0.041

Direct
  CHEAT1
  SUB1               -0.230       0.079      -2.927       0.003
```

```
Effects from HON1 to CHEAT1

  Total                -0.453      0.066     -6.881      0.000
  Total indirect       -0.201      0.050     -4.028      0.000

  Specific indirect

    CHEAT1
    PEER1
    HON1               -0.018      0.027     -0.690      0.490

    CHEAT1
    CHJUST1
    HON1               -0.089      0.042     -2.104      0.035

    CHEAT1
    TEACHER1
    HON1               -0.015      0.024     -0.656      0.512

    CHEAT1
    PEER1
    TEACHER1
    HON1               -0.008      0.012     -0.677      0.499

    CHEAT1
    CHJUST1
    PEER1
    HON1               -0.046      0.024     -1.952      0.051

    CHEAT1
    CHJUST1
    TEACHER1
    HON1               -0.004      0.012     -0.350      0.727

    CHEAT1
    CHJUST1
    PEER1
    TEACHER1
    HON1               -0.020      0.012     -1.708      0.088

  Direct
    CHEAT1
    HON1               -0.252      0.074     -3.383      0.001


Effects from TEACHER1 to CHEAT1

  Total                -0.270      0.123     -2.184      0.029
  Total indirect       -0.181      0.081     -2.227      0.026

  Specific indirect

    CHEAT1
    PEER1
    TEACHER1           -0.045      0.065     -0.699      0.485

    CHEAT1
    CHJUST1
    TEACHER1           -0.023      0.066     -0.351      0.726
```

```
    CHEAT1
    CHJUST1
    PEER1
    TEACHER1          -0.113      0.052     -2.172      0.030

  Direct
    CHEAT1
    TEACHER1          -0.089      0.132     -0.669      0.503


Effects from CURUSE1 to SURF1

  Total               -0.144      0.133     -1.084      0.279
  Total indirect      -0.086      0.086     -1.004      0.315

  Specific indirect

    SURF1
    CHJUST1
    CURUSE1           -0.138      0.085     -1.624      0.104

    SURF1
    CHJUST1
    PEER1
    CURUSE1            0.051      0.044      1.184      0.237

  Direct
    SURF1
    CURUSE1           -0.057      0.125     -0.459      0.646


Effects from PERF1 to SURF1

  Total                0.298      0.091      3.262      0.001
  Total indirect       0.270      0.072      3.725      0.000

  Specific indirect

    SURF1
    CHJUST1
    PERF1              0.231      0.068      3.394      0.001

    SURF1
    CHJUST1
    PEER1
    PERF1              0.039      0.029      1.335      0.182

  Direct
    SURF1
    PERF1              0.028      0.102      0.273      0.785


Effects from SUB1 to SURF1

  Total               -0.330      0.079     -4.190      0.000
  Total indirect      -0.080      0.073     -1.096      0.273

  Specific indirect

    SURF1
    PERF1
    SUB1               0.002      0.009      0.263      0.793
```

```
    SURF1
    CURUSE1
    SUB1                -0.031      0.068     -0.459      0.646

    SURF1
    CHJUST1
    SUB1                 0.016      0.059      0.279      0.781

    SURF1
    TEACHER1
    SUB1                 0.047      0.064      0.736      0.462

    SURF1
    CHJUST1
    PERF1
    SUB1                 0.020      0.021      0.969      0.333

    SURF1
    CHJUST1
    CURUSE1
    SUB1                -0.075      0.047     -1.593      0.111

    SURF1
    CHJUST1
    TEACHER1
    SUB1                -0.015      0.045     -0.348      0.728

    SURF1
    CHJUST1
    PEER1
    PERF1
    SUB1                 0.003      0.004      0.809      0.419

    SURF1
    CHJUST1
    PEER1
    CURUSE1
    SUB1                 0.028      0.024      1.172      0.241

    SURF1
    CHJUST1
    PEER1
    TEACHER1
    SUB1                -0.076      0.031     -2.441      0.015

  Direct
    SURF1
    SUB1                -0.250      0.091     -2.741      0.006


Effects from TEACHER1 to SURF1

  Total                 -0.100      0.140     -0.714      0.475
  Total indirect        -0.207      0.100     -2.080      0.038

  Specific indirect

    SURF1
    CHJUST1
    TEACHER1            -0.035      0.101     -0.349      0.727
```

```
    SURF1
    CHJUST1
    PEER1
    TEACHER1          -0.172      0.064     -2.676      0.007

  Direct
    SURF1
    TEACHER1           0.107      0.144      0.743      0.458


Effects from SUB1 to CHJUST1

  Total               -0.141      0.079     -1.788      0.074
  Total indirect      -0.164      0.063     -2.613      0.009

  Specific indirect

    CHJUST1
    PERF1
    SUB1               0.029      0.029      0.982      0.326

    CHJUST1
    CURUSE1
    SUB1              -0.107      0.066     -1.622      0.105

    CHJUST1
    TEACHER1
    SUB1              -0.022      0.063     -0.351      0.726

    CHJUST1
    PEER1
    PERF1
    SUB1               0.005      0.006      0.814      0.416

    CHJUST1
    PEER1
    CURUSE1
    SUB1               0.040      0.034      1.191      0.234

    CHJUST1
    PEER1
    TEACHER1
    SUB1              -0.108      0.042     -2.589      0.010

  Direct
    CHJUST1
    SUB1               0.023      0.084      0.279      0.780


Effects from CURUSE1 to CHJUST1

  Total               -0.123      0.122     -1.009      0.313
  Total indirect       0.074      0.061      1.203      0.229

  Specific indirect

    CHJUST1
    PEER1
    CURUSE1            0.074      0.061      1.203      0.229

  Direct
    CHJUST1
    CURUSE1           -0.197      0.119     -1.655      0.098
```

```
Effects from PERF1 to CHJUST1

  Total                 0.386      0.082      4.716      0.000
  Total indirect        0.055      0.041      1.358      0.175

  Specific indirect

    CHJUST1
    PEER1
    PERF1               0.055      0.041      1.358      0.175

  Direct
    CHJUST1
    PERF1               0.331      0.080      4.143      0.000


Effects from TEACHER1 to CHJUST1

  Total                -0.296      0.130     -2.279      0.023
  Total indirect       -0.246      0.085     -2.888      0.004

  Specific indirect

    CHJUST1
    PEER1
    TEACHER1           -0.246      0.085     -2.888      0.004

  Direct
    CHJUST1
    TEACHER1           -0.050      0.143     -0.352      0.725


Effects from HON1 to CHJUST1

  Total                -0.346      0.071     -4.874      0.000
  Total indirect       -0.152      0.045     -3.362      0.001

  Specific indirect

    CHJUST1
    PEER1
    HON1               -0.101      0.042     -2.408      0.016

    CHJUST1
    TEACHER1
    HON1               -0.009      0.025     -0.351      0.726

    CHJUST1
    PEER1
    TEACHER1
    HON1               -0.043      0.021     -1.995      0.046

  Direct
    CHJUST1
    HON1               -0.194      0.074     -2.610      0.009


Effects from CURUSE1 to PEER1

  Total                 0.163      0.127      1.279      0.201
  Total indirect        0.000      0.000      0.000      1.000
```

```
  Direct
    PEER1
    CURUSE1             0.163      0.127      1.279      0.201


Effects from HON1 to PEER1

  Total                -0.318      0.077     -4.107      0.000
  Total indirect       -0.095      0.042     -2.284      0.022

  Specific indirect

    PEER1
    TEACHER1
    HON1               -0.095      0.042     -2.284      0.022

  Direct
    PEER1
    HON1               -0.223      0.081     -2.738      0.006
```

Appendix AG: Human Research Ethics Committee Approval Letter (protocol 14193)

THE UNIVERSITY OF SYDNEY

RESEARCH INTEGRITY
Human Research Ethics Committee
Web: http://sydney.edu.au/ethics/
Email: ro.humanethics@sydney.edu.au

Address for all correspondence:
Level 6, Jane Foss Russell Building - G02
The University of Sydney
NSW 2006 AUSTRALIA

MF/PE

21 October 2011

Dr Paul Ginns
Faculty of Education and Social Work
Education Building - A35
The University of Sydney
paul.ginns@sydney.edu.au

Dear Dr Ginns

Thank you for your correspondence dated 20 October 2011 addressing comments made to you by the Human Research Ethics Committee (HREC).

I am pleased to inform you that with the matters now addressed your protocol entitled **"Unification of contextual perspectives on learner cognition: Achievement goal theory, student learning theory, and academic integrity"** has been approved.

Details of the approval are as follows:

Protocol No.: **14193**

Approval Date: **21 October 2011**

First Annual Report Due: **31 October 2012**

Authorised Personnel: **Dr Paul Ginns**
Associate Professor Richard Walker
Mr Bradford Barnhardt

Documents Approved:

Document	Version Number	Date
Student Survey	2	10/10/2011
School Consent Form	2	10/10/2011
Parental (or Caregiver) Information Statement	1	19/9/2011
Parental (or Caregiver) Information Statement (Questionnaire Trial)	1	19/9/2011
Participant Information Statement (Children Under 18 years)	1	19/9/2011
Participant Information Statement (Questionnaire Trial) (Children Under 18 years)	1	19/9/2011
Parental (or Caregiver) Consent Form	1	19/9/2011

HREC approval is valid for four (4) years from the approval date stated in this letter and is granted pending the following conditions being met:

Manager Human Ethics
Dr Margaret Faedo
T: +61 2 8627 8176
E: margaret.faedo@sydney.edu.au

Human Ethics Secretariat:
Ms Karen Greer T: +61 2 8627 8171 E: karen.greer@sydney.edu.au
Ms Patricia Engelmann T: +61 2 8627 8172 E: patricia.engelmann@sydney.edu.au
Ms Kala Retnam T: +61 2 8627 8173 E: kala.retnam@sydney.edu.au

ABN 15 211 513 464
CRICOS 00026A

THE UNIVERSITY OF SYDNEY

Special Condition of Approval

1. **Please provide a full list of the participating schools and consent forms from school principals when received.**

Condition/s of Approval

- Continuing compliance with the National Statement on Ethical Conduct in Research Involving Humans.
- Provision of an annual report on this research to the Human Research Ethics Committee from the approval date and at the completion of the study. Failure to submit reports will result in withdrawal of ethics approval for the project.
- All serious and unexpected adverse events should be reported to the HREC within 72 hours.
- All unforeseen events that might affect continued ethical acceptability of the project should be reported to the HREC as soon as possible.
- Any changes to the protocol including changes to research personnel must be approved by the HREC by submitting a Modification Form before the research project can proceed.

Chief Investigator / Supervisor's responsibilities:

1. You must retain copies of all signed Consent Forms and provide these to the HREC on request.
2. It is your responsibility to provide a copy of this letter to any internal/external granting agencies if requested.

Please do not hesitate to contact Research Integrity (Human Ethics) should you require further information or clarification.

Yours sincerely

Dr Margaret Faedo
Manager, Human Ethics
On behalf of the HREC

cc. Bradford Barnhardt bbar6232@uni.sydney.edu.au

This HREC is constituted and operates in accordance with the National Health and Medical Research Council's (NHMRC) National Statement on Ethical Conduct in Human Research (2007), NHMRC and Universities Australia Australian Code for the Responsible Conduct of Research (2007) and the CPMP/ICH Note for Guidance on Good Clinical Practice.