# Hans Eysenck, Education and the Experimental Approach:

# A Meta-analysis of Academic Capabilities in University Students

Philip J. Corr

City University London

Rock Braithwaite

Humboldt State University

Author Note

Philip J. Corr, Department of Psychology, City University London, UK.

Correspondence concerning this article should be addressed to Philip J. Corr, Department of Psychology, City University London, Northampton Square, London, EC1V 0HB, UK. E-mail: philip.corr.1@city.ac.uk

**Abstract**

Hans Eysenck had a long-established interest in the influence of individual differences on educational attainment, noting that typically personality traits and cognitive abilities are ignored in debates concerning educational policy and practice. His general scientific approach emphasized the importance of applying an experimental approach to answering social questions. Inspired by this perspective, in this article, we conducted a meta-analysis of the literature on (largely quasi) experimental intervention studies ($N$ = 47, with 49 independent samples) aimed at enhancing self-concept (comprising various forms of self-efficacy and self-confidence) in order to influence a range of academic outcomes in university students ($N$ = 5,535). Results revealed small-to-moderate, but statistically significant, positive effects across all the outcome domains examined. There was only little evidence for moderation of these effects, with only the quality of the study intervention being statistically significant (lower quality studies showing the largest effect sizes). Although our analysis shows the paucity of purely experimental studies in higher education research, the results are sufficiently clear to suggest that the study of individual differences variables are important in educational design and instruction.

*Keywords*: Self-efficacy, student confidence, university, capabilities, economics, experimental intervention

Hans Eysenck, Education and the Experimental Approach:

A Meta-analysis of Academic Capabilities in University Students

Hans Eysenck always had strong interests in applying psychological principles and findings to social issues, and one of the most important of these is education (Corr, 2016). As Eysenck (1991) stated, education is important because, "it concerns the future of our children, and hence our whole culture and civilization" (p. 191). During the late 1960s, Eysenck involvement in the 'progressive education' debate resulted in the Black Papers that challenged the rush to 'comprehensive education' in the UK, the 'grammar school' system that was based on intellectual merit and intelligence testing at age eleven (see Corr, 2016, pp. 181-3). Even the indomitable Margaret Thatcher, during her time as Education Secretary (1970-74), could do little to stem this political tide – indeed, under her ministership she closed more grammar schools than any other Education Secretary. When Eysenck asked her why there was no research into the effectiveness and success of the new 'progressive' system, she told him that she had asked her Civil Servants to undertake this but they had done nothing. It seemed then, and now, that education is a research-free zone, at least of the Eysenck type which places data over dogma. The same is true also of higher education, where we might expect the influence of individual differences in appetitive and aptitude for learning and scholastic attainment to be, at least, as marked if not more so than in earlier years. As Eysenck (1998) observed, 'Children, as they grow up, increasingly *choose* their environment; this choice itself is driven by genetic factors. And they *interpret* their environment in terms of their genetic contributions. Our environment is *structured* by ourselves, on the basis of genetic drives' (p. 42; Eysenck's italics).

Education at all levels would seem ripe for the type of experimental approach favoured by Eysenck to answer social questions.

In this article, we examine systematically experimental interventions at university designed to enhance self-esteem and self-efficacy, which are known to be *associated* with educational success. The direction of causation is not known, hence the need for the type of experimental methodology advocated by Eysenck. Specifically, we examine all extant experimental inventions, most of which are quasi-experimental, to determine: (a) the extent and quality of such research; and (b) the patterns of findings and their implications for psychological factors in university success. The focus is on those interventions designed to enhance university students' core self-evaluations, specifically self-efficacy and self-confidence, on a range of academic outcomes (affect, cognition, knowledge, and assessment grades).

University education is an integral part of a progressive and productive society. A major challenge for the higher education sector is to identify the potential for the development of students' capabilities. Effective instructional design features can be identified which allow successful investment in educational processes to produce desired outcomes (e.g., employability, social competence, intellectual curiosity, and more generally good citizenship). The importance of university education is underscored by the finding that over half of life-time human capital is acquired after post-compulsory school investment (Heckman, Lochner & Taber, 1988).

The importance of using evidence-based knowledge to inform educational processes and practices at university is highlighted by (a) the psychological challenges facing students at university, and (b) the wide diversity of outcomes (with some students excelling, while others fail to do so, or fail outright). This is important because the university sector s charged by society with the task of developing the human capital of *all* students to the best of their abilities. Given that university aims to develop students' capabilities, it would be sensible to assume that some 'interventions' (e.g., methods of teaching) are more effective than others. To get at causal influences, experimental interventions are required – but these are few and far between and most that fall under this rubric do not entail randomization of participants and are, thus, are not purely experimental in design.  The main of this article is to summarize these interventions.  Our focus is not on instructional design, but on the wider psychological environment of students' core self-evaluations, centred on self-efficacy and self-confidence, which we assume permeate most learning processes at university (e.g., initiation of study, persistence, reaction to feedback, tolerance of frustration, and so on).

## Psychological Factors in Education

This section provides a summary of psychological constructs that have been widely applied to education in general and which has obvious relevance for higher education. It supports the view that these psychological constructs are important in academic experience and outcomes; and, further, this raises the possibility that they may be subject to influence by experimental means.

**Self-efficacy theory**. The psychological construct of self-efficacy can be defined as a person's perception of his/her ability to perform successfully a behaviour (Sitzmann & Yeo, 2013). It is purported to influence decisions about which behaviour/s to engage in, and persistence in response to difficulty, as well as actual task performance (Multon, Brown & Lent, 1991). For this reason, self-efficacy theory has been highly influential in the educational field.  It is aligned with social learning and social cognitive theory (Bandura & Schunk, 1981), and as such gives consideration to the impact of social factors on self-efficacy beliefs.

Self-efficacy is composed of four factors: mastery experience, verbal persuasion, vicarious feedback, and psychological feedback, with mastery experience seen as the most important factor (Bandura & Adams, 1977). In the educational context, it lends itself well to targeted interventions (e.g., structuring effective feedback), and accounts for how an individual's conceptualisation of ability moderates any potential self-efficacy effect (Wood & Bandura, 1989).

Work on self-efficacy (e.g., Bandura & Schunk, 1981) has provided valuable insight into the successful motivation of individuals performing tasks (particularly via an emphasis on proximal sub-goals). Bandura (1993) provides a detailed account of the influence which self-efficacy beliefs have on the cognitive development of an individual through cognitive, affective, motivational and selection processes; he also gives consideration to how the psychosocial network of influences affects performance (Bandura, Barbaranelli, Caprara & Pastorelli, 1996). Self-efficacy theorists emphasize specific self-efficacy (as opposed to an overall global measure) to predict domain specific behaviours (Pajares, 1996); for example, computer literacy (Compeau & Higgins, 1995), capabilities of arthritis sufferers (Lorig, Chastain, Ung, Shoor & Holman, 1989), and mathematical ability (Pajares & Miller, 1995). Rodgers, Conner and Murray (2008) found behavioural-specific self-efficacy beliefs to be superior predictors of selected academic (reading 1, 30, or 100 pages) and health-related (e.g., tooth flossing everyday, eating 5-10 servings of fruit and vegetables everyday) behaviours when compared to other prominent types of control measures, namely perceived control and perceived difficulty. Specific empirical work has further demonstrated that self-efficacy has predictive utility for academic outcomes across different age-groups (e.g. Bandura *et al*., 1996; Caprara *et al*., 2008) and for work-related performance in a range of different study settings from simulated/laboratory based work to actual/field work (Stajkovic & Luthans, 1998).

**Self-confidence.** As a concept, self-confidence is similar to, but not isomorphic with, Bandura's concept of self-efficacy; it is recognized as crucial in psychology, education, and employability literatures, and is itself composed of specific features, (e.g., academic, inter-personal and occupational). It is commonly seen as a key determinant of how students respond to academic and employment-related opportunities and challenges. Broadly speaking, 'confidence' reflects those cognitive and affective processes that relate to the perceived capacity to use current capabilities to achieve some, not yet attained, desired outcome. It can be viewed as a general attitude to action-outcomes relations; and. In contrast, self-efficacy can be seen to comprise the psychological processes that enable these relations.

In contrast to individuals low in confidence, those high in this concept believe that they can reach their desired future states with the necessary personal investment (e.g., time, effort, and commitment). This self-belief has important emotional consequences which motivate behaviour towards sources of potential reward – important in this regard, too, is the tolerance of negative emotions, such as frustration and anxiety (Corr, 2013). University life is as much about emotional experiences as it is intellectual ones.

***Academic self-confidence*.** Linked to the above literature, there have been attempts to define the notion of specific academic 'confidence' **-** usually defined as the belief in one's capability to achieve some specific outcome (e.g., give a successful tutorial presentation). For example, a series of studies have developed and used the Academic Behavioural Confidence (ABC) scale (Sander & Sanders, 2003, 2006, 2009). This research applies the ABC model to self-efficacy (with the same four determining variables; see above). This scale has been refined to a four-factor (Sander & Sanders, 2009), leading to a 17-item scale that measures: *confidence in grades, confidence in verbalising, confidence in study, and confidence in attendance*. The authors obtained a significant correlation between ABC score and final year degree grade. This work is highly consistent with Multon and colleagues' (1991) meta-analytical work which found a statistically significant relationship between self-efficacy and academic performance based upon a review of 39 studies.

**Other constructs related to self-confidence**. There are additional approaches which, whilst not directly relating to confidence, are worth noting for completeness. The learned optimism work of Seligman (1998), and the considerable body of work within social psychology on locus of control, are two key examples. There have also been notable attempts to create specific instruments for related concepts that may be useful for any study of confidence. For example, Vallerand *et al*. (1992) provided a measure of academic motivation, which was proposed as a key intermediate variable between specific self-efficacy and task performance. The Rosenberg Self-Esteem Scale (Rosenberg, 1965) is of potential use in accounting for the role of what is potentially a linked construct. Stark, Bentley, Lowther and Shaw (1991) provide a Student Goals Exploration Test which is also of relevance. Carroll and Garavalia (2004) and Klomegah (2007) apply a Motivated Strategies for Learning Questionnaire alongside the Self-Efficacy in Self-Directed Learning Questionnaire.

**Perceived behavioural control and optimism.** Various other theories supplement insights from the self-concept literature. For example, Orbell (2003) uses the concept of perceived behavioural control, which integrates material from a variety of psychological sources, particularly the theory of planned behaviour, behavioural control research and personality systems interaction theory. This literature reveals that the addition of perceived behavioural control to attitudes and norms substantially increases the explanation of variance for academic behaviours.

Perceived behavioural control shares much in common with notions of locus of control, and the two concepts have been conceptually linked by Ajzen (2002). This variable is of importance as it may be seen as the capability to monitor goal-directed behaviour under conditions of uncertainty and, also, in the absence of immediate positive reinforcement; and it entails restraint of action to control behaviour to match prevailing environment conditions.

Ruthig, Haynes, Stupnisky and Perry (2009) conducted work in a related area, establishing a link between perceived academic control and optimism (thus linking the concept to the influential model of learned optimism; Seligman, 1998). This study built on the foundations established in an earlier study by Ruthig, Perry, Hall and Hladkyj (2004), which detailed a successful intervention that targeted optimism and attributions via attribution retraining. Results showed that optimism alone was a *risk* factor in this particular study, the implication being that only when optimism is channeled by appropriate attributions can it manifest effectively in performance. Relevant literature relating to attributional retraining has focused on academic performance, with has produced mixed results. Examples of evidence for and against include Mitchell and Hirom (2002) and Bridges (2001), respectively (see Gibb, Zhu, Alloy & Abramson, 2002).

**Aims of Study**

Given the importance of core self-evaluation, largely focused as they are on self-efficacy and self-confidence, in the development of university-related capabilities and academic outcomes, it is necessary to know whether they can, by targeted experimental intervention, be enhanced. However, the education literature is mixed. For example, Kahn and Nauta (2001) tested a social learning theory model of first-year college persistence to test precollege and first-semester college performance predictors. Contrary to their hypotheses, they did not find a significant role of first-semester self-efficacy beliefs, outcome expectations, or performance goals. In contrast, a similar finding was demonstrated within an experimental study that determined a *negative* relationship between self-efficacy and performance due to the likelihood of committing logic errors because of overconfidence (Vancouver, Thompson, Tischner & Putka, 2002). Clarification of this literature seems warranted.

An alternative viewpoint is that, as core-evaluations may reflect dispositional aspects of personality and cognitive abilities, these psychological constructs are not amenable to change and are, so to speak, set in stone. This viewpoint is not consistent with the investment model of personality which emphasizes the malleability of personality traits (see Ferguson, Heckman & Corr, 2011), and nor with the *raison d'être* of the role of the university in society. However, empirical evidence is needed to reveal whether, or not, this is the case.

To address the above issue, we conducted a meta-analysis of the existing literature on (largely) non-experimental interventions that focused on enhancing, in general terms, core self-evaluations, largely comprising the related constructs of self-efficacy and self-confidence. Although the number of such experimentally-controlled studies is relatively small, a sufficient number now exist to render a meta-analysis viable and conclusions potentially of relevance for instructional design and for fostering a productive psychological environment.

## Method

The literature search and inclusion criteria are detailed in Supplementary Material, as are the coding and data extraction forms.

### Literature Search and Inclusion Criteria

A comprehensive literature search was employed using electronic databases, review articles, and manual searches of article reference lists matching inclusion criteria. Computer database searches were completed in *British Educational Index (BEI), Cambridge Journals Online, Educational Resources Information Center (ERIC), Proquest Dissertations & Theses, PsycINFO, PsycARTICLES, OmniFile Full Text, Oxford Journals Online*, and *Web of Science/Web of Knowledge*, using variations of the key words: *Self-Confidence, Self-Efficacy, Social Cognitive Theory, Intervention, Experiment, Control*, *Student*, *College*, and *University*.

The search process was conducted by entering combinations of keywords and then recording dates, search limits, total results of each database searched, number of duplicates for each search, and number of articles included for each search. Since the intent of the current study was to determine the effectiveness of self-concept interventions on affective, behavioural, and cognitive variables related to academic performance, operational definitions and inclusion criteria were used to determine research evidence that was relevant to the study purpose.

The definition of student academic management was the use of affective, behavioural, or cognitive tools or skills that facilitated the completion of academic tasks that contributed to current and/or future success. *A priori* inclusion criteria for screening decisions included: (a) articles or papers conducting general self-concept, but specifically self-efficacy and self-confidence, interventions using a control group or comparison measure; (b) studies reporting pre-test and post-test measures; (c) studies conducted on college/university students (18 years and older); (d) the study purpose was to enhance student academic achievement and/or performance though a self-concept intervention; (e) papers that included quantitative measures that would provide a calculation of effect size; and (f) articles available in the English language from January 1977 to September 2013. Study titles and abstracts were reviewed to make preliminary inclusion decisions; and then were exported into Endnote version X5 (Thompson Reuters, 2011) to organize and manage reference lists.

Coding and data extraction forms were developed using established protocols (Brown, Upchurch, & Acton, 2003; Hattie, Biggs, & Purdie, 1996; Wilson & Lipsey, 2001). Study information was separated into three categories: *Intervention*, *Participant*, and *Study* features.

***Intervention.*** *Design* included using an experimental approach employing randomized allocation and the use of a control group, or quasi experimental using nonrandomized same group or different group comparisons (most of the identified studies fell into the second group). *Duration* was the length of time students were exposed to the experimental intervention: Studies were coded as *semester* (equivalent to one academic term), *academic year* (equivalent to two academic terms), or *other* (shorter or longer than a semester or year). *Characteristic* related to the use of a *theory-based* (explicit connections identified and linked the theoretical framework when designing and implementing the intervention/s) or *non-theoretical* (lack of or minimal connection to theoretical framework). *Intent* provided classification of studies being the *primary* or *secondary* focus of the study: Studies using self-concept as a *primary* focus intended to improve student outcomes through the use of a specified self-concept treatment while studies with a *secondary* focus employed several treatment methods. *Follow-up* was coded as having (*yes*) or not (*no*) delayed posttest measures (after post-test) of student based outcomes to determine the treatment or program effectiveness after the study had been completed. *Classification* was based on Biggs and Collins (1982) structure of the observed learning outcomes (SOLO) and subsequent review by Hattie et al. (1996) used to classify intervention effectiveness of studies designed to improve learning: Studies classified as *unistructural* focused on enhancing a singular feature or characteristic compared to *multistructural* interventions that attempted to improve a range of strategies or procedures, while *relational* interventions served to facilitate self-assessment through self-monitoring or self-regulated strategies. *Nature* referred to how the intervention transferred to the academic achievement or outcome: *Reproductive* outcomes focused on content, while *transformational* outcomes were used to develop additional strategies or skills beyond the current context. *Quality* was determined using the criteria established to evaluate interventions used in education (Reed et al., 2005) and medicine (Higgins & Green, 2011) – it was decided to include studies with apparently lower quality in order to empirically assess whether, indeed, this variable serves as a moderator.

All studies were coded according to explanation and rationale of study design, specifically: randomization of group processes, sample baseline characteristics were reported, incomplete or selective reporting of outcome measurements, statistics were reported on reliability and/or validity of outcome measures, study procedures provided details permitting replication. Studies were assigned values of '0' or '1' based on the coding process with a total possible score of 6 points for each study. Studies scoring between 0 to 2 points were considered of low quality, medium quality were scores of 3 to 4 points, and high quality interventions were scores above 5 points.

***Participant.*** S*ample size*, *gender*, *country* in which the intervention occurred, and *learner level* (low ability, mixed ability, and high ability).

***Study***. *Publication type* (published or unpublished), *outcome measure* (affective, performance-based, study skills, or combinations of study outcomes), and study *reporting method* (student report, instructor report, or combined reporting method).

Two coders independently reviewed and reported codes for each of the studies meeting inclusion criteria; these codes were then examined by a third coder, who also looked at any discrepancies between the first two coders. Coding results were compared and analyzed using agreement rates and an inter-rater reliability coefficient. Prior to the statistics being calculated, discrepancies between study codes were reviewed and classified as factual or interpretative. (Factual errors were considered transcription errors where the correct answer was present in the study and either missed by the coder or inaccurately reported. Interpretative errors were considered to be errors where study information was inferred or not clear and required the coder to make an interpretation on the classification). All factual errors were corrected; interpretative errors were reviewed by a third author and a simple majority decision determined the appropriate code.

**Outcome Measures**

Outcome measures for each study were coded to provide a summary treatment effect. This process involved listing corresponding measurement tools, questionnaires, subscales, and individual items that were involved in measuring student performance. Due to the variability for measuring student achievement and to provide consistency in reporting outcome measures three separate authors reviewed and grouped dependent variables according to the purpose of the measurement tool.

Categories were individually established and then discussed to select the best method to group and report student achievement. The emergent categories included: Self-efficacy/confidence (specific and general measures); knowledge/learning (e.g., IQ, content knowledge questionnaires, etc.); learning strategies (e.g., metacognition, and feedback loops); anxiety; self-regulatory processes (e.g., self-monitoring, and organization, planning); motivation (e.g., goal profiles, and internal attributions); attitude/interest; self-perceptions (e.g., self-appraisal, and self-evaluation); social skills (e.g., collaboration, cohesion, and social integration);

professional aspirations (e.g., career decisions); and academic attainment (e.g., tests, assignments, and GPA).

As shown in Table 3, outcome measures were grouped in to three main broad domains: *Process* (i.e., self-efficacy, knowledge intellect, and learning strategies); *Orientation* (i.e., attitude/interest, self-regulation, motivation, self-perceptions, social skills, professional aspirations, and anxiety); and *Performance* (i.e., academic attainment) – these were reported in one instance (no pretest just posttest) and this contrasted with *Process* based knowledge (as represented by a 'gain score', or improvement e.g. reported improvements in confidence, organizational skills, etc.). Irrespective of the validity of the construction of three domains (and, of course, they are different ways to categorize the variables), statistics are reported separately for all the outcomes measures and, thus, their interpretation is not significantly affected by this classification. However, it seemed sensible to attempt some thematic organization, to reflect what might be different psychological levels. *Performance* is an outcome achievement measure and, thus readily stands apart from *Process* and *Orientation*; and in relation to the last two categories, *Process* is seen to reflect more formal mechanisms enabling the *Orientation* variables (e.g., *Process* 'knowledge/intellect' should be expected to influence the expression of *Orientation* 'attitude/interest', and the same may be said of 'self-efficacy' and 'anxiety', respectively).

Statistical consideration, outlier analysis, and publication bias are detailed in Supplementary Material.

## Statistical Considerations

**Effect size calculations.** Comprehensive Meta-Analysis version-2 software was used to calculate all effect sizes (Borenstein, Hedges, Higgins & Rothstein, 2005). A random effects model using Hedges *g* as the effect size index was selected to measure differences between experimental and comparison groups (Hedges & Olkin, 1985; Hedges & Vevea, 1998). Since there were several separate analyses (outcome and subgroup), the statistical assumption supporting a random effects model suggests that there will be within-study error (sampling error) and between-study variance. Standardized mean differences were adjusted by the inverse weight of the variance to prevent inflation of study weights and to provide more accurate estimates of effect size. Meta-analytic literature has found that Hedges *g* prevents overestimation of an effect size value when sample sizes are fewer than 20 studies (Field, 2001; Hedges & Olkin, 1985).

The total sample ($N$ =49) suggests Cohen's *d* would be the appropriate statistic to report; however, to ensure consistency of reporting methods, Hedges' *g* was selected due to the smaller sample sizes in outcome and subgroup analyses. The standard formula for Hedges' *g* used to correct for bias in small samples was:

$$g = d\left[1 - \frac{3}{4N - 9}\right]$$

Descriptive measures such as means, standard deviations, and sample sizes were used to calculate estimates of effect size. When descriptive data were not available effect sizes were calculated from *F*, *t*, r, and/or *p*-values. The standard procedure for reporting results was that each study contributed one effect size calculation to the overall analysis. When a study contained more than one measurement, outcomes were averaged to provide an overall summary effect.

**Heterogeneity of Variance.** The three statistics used to assess homogeneity of variance included the $Q_{Total}$ ($Q_T$) value which is based on a χ-square ($\chi^2$) distribution, tau-square ($\tau^2$) value, and I-square ($I^2$) value. All three statistics ($Q_T$, $\tau^2$, *and* $I^2$) were used to interpret heterogeneity of variance. When the $Q_T$ statistic is significant then a procedure is used to conduct subgroup (moderator) analyses by compartmentalizing variance into $Q_{Between}$ ($Q_B$) and $Q_{Within}$ ($Q_W$) values with significant $Q_B$ values (p < .05) needing a statistical technique (*t-test* or *ANOVA*) to determine group differences (Borenstein, Hedges, Higgins & Rothstein, 2009; Hedges & Olkin, 1985). The $\tau^2$ statistic provides an estimate of total variance between studies with larger values reflecting the proportion of variance that can be attributed to real differences between studies in a random effects model. When the number of studies per subgroup is small ($k \leq 5$) $\tau^2$ can be imprecise, therefore, a pooled estimate of variance was used for all calculations (Borenstein *et al*., 2009). The $I^2$ statistic is the ratio of excess dispersion to total dispersion and can be interpreted as the overlap of confidence intervals explaining low (25%), moderate (50%), and high (75%) values of the total variance attributed to covariates (Higgins, Thompson, Deeks & Altman, 2003). Larger values $I^2$ values require techniques (i.e., moderator/subgroup analysis or meta-regression) to provide explanations (Borenstein *et al*., 2009). Research suggests that smaller samples sizes increase the likelihood that assumptions will be violated when using a random effects model as error can be overestimated (Field, 2001; 2005; Overton, 1998). A conservative alpha level ($\alpha < .01$) was established to prevent type I errors from being committed when interpreting results from the subgroup analyses.

**Outlier Analysis and Publication Bias**

Outliers were identified by analyzing relative residual values ($Z$ < or > $\pm$ 1.96) and if present were analyzed by using a ‘one study removed’ technique that is available with the CMA version-2 software (Borenstein et al., 2005). The criteron for outlier inclusion was a large residual value that did not influence significant ($p \leq .01$) effect sizes (Hedges *g*) and remained within the 95% confidence interval.

Publication bias was analyzed through visual inspection of a funnel plot, a Fail Safe *N* calculation (Rosenthal, 1979), and a ‘Trim and Fill’ procedure (Duval & Tweedie, 2000a, b).

Funnel plots provide a visual representation of studies according to standard error (y-axis) and effect size (x-axis) with symmetrical distributions being indicative of a lack of publication bias. The 'Trim and Fill' procedure is iterative statistical process that adds/removes studies to balance an asymmetrical funnel plot and provide an unbiased estimate of effect size (Duval & Tweedie, 2000a, b). Fail Safe N calculations are based on the number of studies needed to nullify significant effects (Rosenthal, 1979).

### Final Studies

Combinations of search terms generated a possible list of 18,489 articles that were reviewed by title and abstract and reduced to a total of 356 studies. This was reduced to 47 studies with 49 independent samples meeting inclusion criteria, involving 5,535 participants from 11 different countries. Results from the coding process produced an inter-rater agreement of 96.7% ($r$ = .939) that ranged from 89.7% to 100% across the three characteristics (intervention, sample, and study features; see Supplementary Material). Based on the types of coding disagreements, there were 21 total disagreements including 9 factual errors that were corrected and 12 interpretation errors and independently reviewed and coded by a third author to determine final codes.

## Results

We examined the effects of experimental interventions on university students' academic-related outcomes. Tables 1 and 2 provide the descriptive information for the studies that met our inclusion criteria. Cohen's (1988) criteria has established that effect sizes are small ($\leq$ 0.20), medium (0.50), or large ($\geq$ 0.80) with positive effect sizes interpreted as treatment groups having stronger results than control or comparison conditions. Negative effect sizes indicated that control groups or comparison measures yielded larger outcomes.

**Random Effects Model**

Overall, we found a small positive effect (*g* = 0.273, *SE* = 0.041, *C.I.* = 0.194, 0.393, $p <$ .001) for university students exposed to experimental treatments. This small effect represented approximately one quarter of a standard deviation improvement on learning outcomes and achievement. The differences between the individual outcome measures (Tables 3 & 4) qualify this omnibus statistic.

Analysis of homogeneity statistics determined there was a significant heterogeneous distribution ($Q_{Total}$ = 208.6, $p < .001$) of studies requiring subgroup analyses to explain a large portion of variance ($I^2$ = 76.99) between study covariates. Review of standard residuals produced six outliers (Chyung, Winiecki & Fenner, 1998, *z* = 4.34; Duijnhower, Prins & Stokking, 2010, *z* = -2.50; Gaudine & Saks, 2004, *z* = -2.08; Latham, 2006, *z* = -2.48; Papinczak, Young, Groves & Haynes, 2008, *z* = -2.17; Rampp & Guffey, 1999, *z* = 1.98), therefore, a sensitivity analysis was performed. The CMA version 2 software (Borenstein *et al*., 2005) provides a 'one study removed program feature that completed the sensitivity analysis finding only a small change in effect size (*g* = 0.262) would result from removing any single study and remain close to or within the 95% confidence interval. The Fail Safe *N* calculation determined that an additional 1810 studies were needed to produce results that would exceed the predetermined alpha value ($\alpha$ = .05). Publication bias was deemed marginal, therefore, the Trim and Fill procedure was not needed to provide an unbiased estimate of overall treatment effect.

**Outcome Analyses**

We found that experimental interventions influenced a diverse range of outcomes measures. Outcome variables not reported by more than four studies were removed from the analysis as estimates of effect size can be imprecise (Borenstein *et al*., 2009). Overall, there were small-to-moderate treatment effects across all outcomes, ranging from -0.134 to 0.598 effects sizes.

The largest treatment effects in the Orientation domain were perceptions of self ($k$ = 8, $g$ = 0.368), attitude/interest ($k$ = 6, $g$ = 0.319), and professional aspirations ($k$ = 7, $g$ = 0.269). Homogeneity and publication bias statistics suggest that study distributions ($Q_T$ <.05) had a high degree of variability, and that the number of studies needed to increase significant p-values beyond the threshold were suspect to publication bias.

Process variables produced the most variable treatment effects that were small-to-moderate. Objective measures of knowledge or intellect (IQ tests, content knowledge assessment, etc.) showed the largest effect size ($k$ = 7, $g$ = 0.598). Significant heterogeneity statistics ($Q_T$ = 59.13, $p$ < .05) were indicative of a diverse distribution of study results and publication bias was improbable (Fail safe $N$ = 124).

Performance (i.e., exam scores, assignments, or grade point averages) showed only a small overall treatment effect ($k$ = 14, $g$ = 0.259). Review of heterogeneity and publication bias statistics indicated consistent findings ($Q_T$= 39.30, $p$ > .05) for improved performance across studies with marginal publication bias (Fail Safe N = 96).

**Subgroup Analyses**

Interpretation of the homogeneity statistics for the random effects model determined there was a heterogeneous distribution ($Q_{Total}$ = 208.6, $p$ < .001) and that a larger portion of variance ($I^2$ = 76.99) could be explained by conducting subgroup analyses. Tables 4 and 5 summarize the moderator statistics for the coded intervention characteristics (Table 4), sample characteristics (Table 5), and study characteristics (Table 5).

In summary, there were overall trends ($p < .05$) indicating improved learning and performance outcomes for students experiencing experimental treatments; however, study quality was the only moderating variable within intervention characteristics to produce significant differences ($Q_B = 6.601$, $p < .05$) between categories. Borenstein **et al**. (2009) have recommended that when interpreting moderating variable differences conservative approaches should be employed when subgroups are minimal ($k < 5$) as estimates of treatment effect may be imprecise. We have selected to report moderator statistics and provide a conservative interpretation in order to highlight trends and recommend future directions.

**Intervention characteristics**. Most of the categories within intervention characteristics produced positive trends including larger effect sizes for: (a) Experimental designs ($g = 0.429$, $Z = 3.429$, $p < .05$); (b) studies employing academic year interventions ($g = 0.327$, $Z = 2.276$, $p < .05$); (c) multistructural interventions attempting to improve student performance using a range of strategies or procedures ($g = 0.298$, $Z = 4.142$, $p < .05$); (d) interventions as the primary basis for improving student learning and performance ($g = 0.332$, $Z = 5.195$, $p < .05$); and (e) interventions that focused on improving student content knowledge and/or skills ($g = 0.299$, $Z = 4.586$, $p < .05$). Unexpected positive trends were found for interventions not conducting follow-up measures after post-tests ($g = 0.278$, $Z = 5.944$, $p < .05$), and studies utilizing atheoretical interventions ($g = 0.438$, $Z = 3.429$, $p < .05$). Study quality was the only category within intervention characteristics to produce significant differences between subgroups. Lower quality interventions ($g = 0.640$, $Z = 3.875$, $p < .05$) produced significantly greater effects for treatment groups or conditions. Overall, there were small positive treatment effects for intervention subgroup variables.

**Sample and Study characteristics.** No significant differences were found within sample variables; however, there were several trends including larger treatment effects for interventions conducted on low ability students ($g$ = 0.339, $Z$ = 4.113, $p < .05$), and treatments completed at Universities within the United States ($g$ = 0.353, $Z$ = 6.163, $p < .05$). Analyses of study characteristics produced no significant subgroup differences but, similar to previous findings, trends were present in the data. Larger treatment effects were found for unpublished studies ($g$ = 0.305, $Z$ = 3.064, $p < .05$), affective study outcomes ($g$ = 0.359, $Z$ = 5.937, $p < .05$), and measures that involved student self-reporting ($g$ = 0.300, $Z$ = 5.543, $p < .05$). In summary, sample and study subgroup variables produced small positive treatment effects.

## Discussion

Taking our inspiration from Hans Eysenck's emphasis on the need for empirical, and preferably experimental, studies of the effectiveness of education design and instruction, we examined the effects of (largely) non-experimental interventions aimed at enhancing core self-evaluation, mainly comprising self-efficacy and self-confidence on university educational outcomes. Results revealed a wide variety of effects for the various experimental interventions. Overall there were consistently small-to-moderate treatment effects across all outcomes measures. There were notable effects on the *process* variable of knowledge/intellection, and on the *orientation* variables of self-perceptions and professional aspirations. The overall effect on the *performance* variable of academic attainment was modest. In terms of subgroup effects, although all subgroups produced overall trends indicating improved learning and achievement outcomes for the experimental treatments, only study quality significantly moderated these effects: Lower quality interventions produced significantly greater effects for treatment groups or conditions.

In addition, most of the categories within intervention characteristics produced positive trends including larger effect sizes for: (a) experimental design studies employing academic year interventions; (b) multistructural interventions attempting to improve student achievement using a range of strategies or procedures; (c) interventions that used self-efficacy/self-confidence as the primary basis for improving student learning and achievement; and (d) reproductive interventions that focused on improving student content knowledge and/or skills.

No significant differences were found within sample subgroup variables; however, there were several trends including larger treatment effects for interventions conducted on low scoring ability students and experimental treatments completed at Universities within the United States. Analyses of study characteristics produced no significant subgroup differences but larger treatment effects were found for unpublished studies, affective study outcomes, and measures that involved student self-reporting. The fact that lower quality studies, unpublished studies, and self-report yielded significant differences suggests that these variables need to be considered when interpreting the results of any one study.

**Theoretical and Practical Implications**

Although the data set was relatively small, it did yield results that hold relevant theoretical and practical implications. The first one is that it may be too easy to over-interpret the results of non-experimental studies. As shown here, even with experimentally-controlled studies, effect sizes are small and there is a negative relationship between study quality and effect size. This is a rather unfortunate outcome for evidence-based design of university education because it seems that poorly designed and conducted studies are the ones which may have the largest influence by virtue of their larger effect sizes.

Results from several outcome and subgroup analyses connect several conceptual elements of the influence that self-concepts have on university-level capabilities. Our findings may have both immediate and lasting implications as they show that increases in time (length of intervention subgroup variable) produces improved student outcomes. Furthermore, studies that performed follow-up analyses (retention measures) demonstrated that the development of capabilities remained consistent after interventions. This finding is perhaps unsurprising, but it is important: interventions need to be targeted and sustained in order for enhanced outcomes to be sustained.

One obvious implication is that programme design to foster university capabilities needs to be implemented early, and also often. Although we did not examine the possibility, it is likely that there is considerable synergy between the development of cognitive, affective and behavioural capabilities. Indeed, it is likely that this synergy takes a statistical interaction form, with each component multiplied by all others and, as such, theses bundles of capabilities need to be jointly considered.

Learning is a process that has the potential to alter a student's current level of cognition and affect, and self-efficacy and self-confidence are identified as significant predictors of several outcome variables. Results from the subgroup analyses on the intervention characteristics classification, nature, and intent provide suggestions on the implementation of targeted interventions. When considering the number of skills (intervention classification subgroup analyses) to be implemented during an intervention, our results indicate that singular or multi-component skill sets can be developed with success. Also, apparent from the results was the nature of interventions that facilitate students' (a) specific (content) skills that produce immediate effects, or (b) more general (transformational) skills that may be expected to have continuing effects on academic outcomes. Finally, when attempting to develop students' general core self-evaluations (intervention intent subgroup analysis) careful consideration should be given to ensuring self-efficacy is the primary focus of strategies to improve student outcomes.

One longer term outcome of such enhanced academic capabilities may be seen in employability, the success of which requires bundles of cognitive, affective and interpersonal capabilities that develop over life. Indeed, the relevance of concepts such as self-efficacy and self-confidence has been shown in a number of studies. For example, Wanberg, Zhang and Diehn (2010) position job search confidence as one of the seven key factors affecting employment outcomes; and Wanberg, Zhu and Van Hooft (2010) use the concept of re-employment efficacy for the unemployed, suggesting that this operates in a direct feedback loop with achievements relating to re-employment (see also, Knight & Yorke, 2004). The positive influence of the development of key academic capabilities should be expected to persist beyond the confines of the university campus.

**Limitations**

When conducting meta-analyses appropriate methods need to be employed to prevent inflated estimates of effect size. Two such concerns that have the potential to influence effect size estimates include publication bias and studies not reporting sufficient data that would permit accurate calculations of effect sizes. To address publication bias the authors established and followed *a priori* inclusion criteria when conducting the literature search, reported inter-rater reliability statistics for relevant information extracted from studies, and used several statistical procedures (i.e., funnel plot review, Trim and Fill procedure, and Fail Safe-N calculation) to control for publication bias. Statistical results indicated that the influence of publication bias was negligible; however, the authors recognize the possibility that studies (either published or unpublished) could have been missed during the literature search process.

Insufficient data can also influence the estimate of effect size and there were several studies that failed to report baseline information or only reported data that was significant, were ambiguous when reporting validity and reliability of the measurement tools used to collect data, and/or did not provide enough information concerning moderating variables that influence self-efficacy such as gender or ethnicity. We have attempted to control for these issues by conducting statistical analyses (i.e., subgroup and outcome analyses) and by providing interpretative precautions for an accurate perspective of the self-efficacy treatment effects on a variety of outcomes in university contexts.

Several other limitations of this literature are noteworthy. First, the majority of studies do not use pure intervention designs as they do not employ full randomization of participants – this is in the very nature of the types of studies conducted in higher education. This limits the scope of interpretation of results, which did suggest that more pure experimental designs yielded larger effect sizes. Second, studies differ in the nature of their interventions and the types of outcome measures used. This makes the task of comparison all the more difficult, requiring grouping of different measures into rather broad categories (e.g., performance attainment). Despite these problems, theoretically relevant results are still observed.

**Conclusions**

Results of our meta-analysis reveal that (albeit largely) non-experimental interventions aimed at enhancing core self-evaluations have statistically significant impacts on a range of university-related capabilities and outcomes. However, as the effect sizes were small-to-moderate, our results suggest that, as researchers and instructors, we should moderate our enthusiasm for results from studies that are not purely experimental in design and where quality of intervention is in question. This conclusion points to the need for far more pure, and higher quality, experimental studies, including ones applying interventions for longer duration, and employing follow-up measures to determine the extent of change overtime.

Our conclusions are consistent with Hans Eysenck's emphasis on the need for rigorous empirical studies to decide the psychological dynamics of educational attainment, even at university level. As we have shown, the fragmented and inadequate nature of his literature reflects the failure to apply purely experimental approaches. As in many other areas of social concern, Eysenck's scientific principles and practices are badly needed.

## References

*References marked with an asterisk indicate studies included in the meta-analysis.*

Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, *32*, 665-683.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*, 117-148.

Bandura, A. & Adams, N. E. (1977). Analysis of self-efficacy theory of behavioral change. *Cognitive Therapy and Research, 1,* 287-310.

Bandura, A., Barbaranelli, C., Caprara, G. V. & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development*, *67*, 1206-1222.

Bandura, A. & Schunk, D. S. (1981). Cultivating competence, self-efficacy and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*, 586-598.

Becker, G. S. & Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of Labor Economics*, *4*, S1-S39. Chicago, IL: University of Chicago Press.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *The Journal of Political Economy*, *75*, 352-365.

*Berzoff, J., Dane, B. & Cait C-A. (2005). Innovative models for developing post-masters curriculum in end-of-life care. *Journal of Teaching in Social Work*, *25*, 63-88.

*Betz N. E. & Borgen F. H. (2009). Comparative effectiveness of CAPA and FOCUS online: career assessment systems with undecided college students. *Journal of Career Assessment, 17,* 351-366.

*Betz, N. E. & Schifano R. S. (2000). Evaluation of an intervention to increase realistic self-efficacy and interests in college women. *Journal of Vocational Behavior, 56*, 35-52.

Biggs, J. B. & Collins, K. F. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy*. New York, NY: Academic Press.

Borenstein, M., Hedges, L., Higgins, J. & Rothstein, H. (2005). *Comprehensive Meta-analysis: A Computer Program for Research Synthesis* (Version 2.0) [Computer software]. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L., Higgins, J. & Rothstein, H. (2009). *Introduction to Meta-Analysis.* Hoboken, NJ: Jon Wiley Publications.

Borghans, L., Golsteyn, B. H. H., Heckman. J. & Humphries, J. E. (2011). Identification problems in personality psychology. *Personality and Individual Differences*, *51*, 315–320.

*Bresó, E., Schaufeli, W. & Salanova, M. (2011). Can a self-efficacy-based intervention decrease burnout, increase engagement, and enhance performance? A quasi-experimental study. *Higher Education: The International Journal of Higher Education and Educational Planning*, *61*, 339-355.

Bridges, K. R. (2001). Using attributional style to predict academic performance: how does it compare to traditional methods? *Personality and Individual Differences, 31*, 723-730.

Brown, S. A., Upchurch, S. L. & Acton, G. J. (2003). A framework for developing a coding scheme for meta-analysis. *Western Journal of Nursing Research*, *25*, 205-222.

*Brown, T. & Morrissey, L. (2004). The effectiveness of verbal self-guidance as a transfer of training intervention: its impact on presentation performance, self efficacy and anxiety. *Innovations in Education and Teaching International, 41*, 255-271.

*Butler D. L. (1997). The roles of goal setting and self-monitoring in students' self-regulated engagement in tasks. *Paper presented at the Annual Meeting of the American Educational Research Association (Chicago*, 20).

Caprara, V.C., Fida, R., Vecchione, M., Del Bove, G., Vecchio, G.M., Barbaranelli, C. & Bandura, A. (2008). Longitudinal analysis of the role of perceived self-efficacy for self-regulated learning in academic continuance and achievement. *Journal of Educational Psychology, 100*, 525-534.

Carneiro, P. & Heckman, J. J. (2003). Human capital policy. In J. J. Heckman, A. B. Krueger, & B. M. Friedman (Eds.). *Inequality in America: What Role for Human Capital Policies?* (pp. 77-239). Cambridge, MA: MIT Press.

Carneiro, P., Cunha, F. & Heckman, J. J. (2003). Interpreting the evidence of family influence on child development. Paper presented at the conference, “The economics of early childhood development: Lessons for economic policy" Federal Reserve Bank of Minneapolis, Minneapolis, MN, October, 17, 2003.

Carroll, C. A. & Garavalia, L. S. (2004). Factors contributing to the academic achievement of pharmacy students: Use of the goal efficacy framework. *American Journal of Pharmacy Education, 68*, 1–8.

*Chyung, Y., Winiecki, D. J. & Fenner, J. A. (1998). A case study: Increase enrollment by reducing dropout rates in adult distance education. In *Distance Learning ‘98: Proceedings of the 14th Annual Conference on Distance Teaching and Learning*, 97-102.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edition. Hillsdale, NJ: Lawrence Erlbaum.

Compeau, D. R. & Higgins, C. A. (1995). Computer self-efficacy: development of a measure and initial test. *MIS Quarterly, 19*, 189-211.

*Cordero, E. D., Porter, S., Israel, T. & Brown, M. T. (2010). Math and science pursuits: a self-efficacy intervention comparison study. *Journal of Career Assessment*, *18*, 362-375.

Corr, P. J. (2013). Approach and avoidance behavior: multiple systems and their interactions. *Emotion Review*, *5*, 286-291.

Corr, P. J. (2016). *Hans Eysenck*: *A Contradictory Psychology*. London: Palgrave.

Cunha, F. & Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, *43*, 738-782.

*Duijnhouwer, H., Prins F. J. & Stokking K. M. (2010). Progress feedback effects on students' writing mastery goal, self-efficacy beliefs, and performance. *Educational Research and Evaluation, 16*, 53-74.

Duval, S. & Tweedie, R. (2000a). A nonparametric “trim and fill” method accounting for publication bias in meta-analysis. *Journal of American Statistical Association*, *9*5, 89-98

Duval, S. & Tweedie, R. (2000b). Trim and fill: a simple funnel-plot-based method of testing for publication bias in meta-analysis. *Biometrics*, *56*, 455-463.

*Elander, J., Pittam, G., Lusher, J., Fox, P. & Payne, N. (2010). Evaluation of an intervention to help students avoid unintentional plagiarism by improving their authorial identity. *Assessment and Evaluation in Higher Education*, *35*, 157-171.

Eysenck, H. J. (1991). *Rebel with a Cause: The Autobiography of Hans Eysenck*. London: W. H. Allen.

Eysenck, H. J. (1998). *Intelligence: The New Look.* London: Transaction Publishers.

Ferguson, E., Heckman J. J. & Corr, P. J. (2011). Personality and economics: overview and proposed framework. *Personality and Individual Differences, 51,* 201–209.

Field, A. (2001). Meta-analysis of correlation coefficients: a Monte Carlo comparison of fixed-and random-effects methods. *Psychological Methods, 6*, 161-180.

Field, A. (2005). Is the meta-analysis of correlation coefficients accurate when population correlations vary. *Psychological Bulletin, 10,* 444-467.

*Fletcher, K. M. (2005). Self-efficacy as an evaluation measure for programs in support of online learning literacies for undergraduates. *Internet and Higher Education, 8,* 307-322.

*Ford-Gilboe, M., Laschinger, H. , Laforet-Fleisser, Y., Foran, S. & Ward-Griffin, C. (1997). The effect of a clinical practicum on undergraduate nursing students' self-efficacy for community-based family nursing practice. *Journal of Nursing Education, 36*, 212-219.

*Freedman, A. M. (1996). Effects of a multicomponent group intervention on the self-efficacy and academic achievement of at-risk undergraduates. [Dissertation]. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 57*(6-A), 2380.

*Gaudine A. P. & Saks, A. M. (2004). A longitudinal quasi-experiment on the effects of posttraining transfer interventions. *Human Resource Development Quarterly, 15*, 57-76.

Gibb, B. E., Zhu, L., Alloy, L. B. & Abramson, L. Y. (2002). Attributional styles and academic achievement in university students: a longitudinal investigation. *Cognitive Therapy and Research, 26*, 309-315.

*Graham, G. M. (2008). Achievement motivation, internal locus of control, goal orientation, and academic self-efficacy as outcome measures for a course designed to positively affect student academic performance. [Dissertation]. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 69* (6-A), 2142.

*Grier R. T. L. & Skaar N. R. (2010). An outcome study of career decision self-efficacy and indecision in an undergraduate constructivist career course. *Career Development Quarterly, 59*, 42-53.

*Griffin, B. W. & Griffin, M. M. (1997). The effects of reciprocal peer tutoring on graduate students' achievement, test anxiety, and academic self-efficacy. *Journal of Experimental Education, 65*, 197-209.

*Hanlon E. H. & Schneider, Y. (1999). Improving math proficiency through self efficacy training. *Paper presented at the Annual Meeting of the American Educational Research Association (Montreal*, 12).

Hattie, J., Biggs, J. & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66*, 99-136.

Heckman, J. J., Hsee, J. & Rubinstein, Y. (2000). The GED is a mixed signal: the effect of cognitive and noncognitive skills on human capital and labor market outcomes. University of Chicago. Unpublished working paper.

Heckman, J. J., Lochner, L. & Taber, C. (1988). Explaining rise wage inequality: explorations with a dynamic general equilibrium model of earnings with heterogeneous agents. *Review of Economic Dynamics, 1*, 1-58.

Heckman, J. J. & Rubinstein, Y. (2001). The importance of noncognitive skills: lessons from the GED testing program. *American Economic Review, 91*, 145-149.

Hedges, L. V. & Olkin, I. (1985). *Statistical Methods for Meta-analysis*. New York, NJ: Academic Press.

Hedges, L. V. & Vevea, J. L. (1998). *Fixed and Random Effects Models in Meta-analysis*. New York, NJ: Russell Sage Foundation.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J. & Altman, D. G. (2003). Measuring inconsistency in meta-analysis. *British Medical Journal*, *327,* 557e560.

Higgins J. P. T. & Green S. (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org. Hofer B. K., & Yu, S. L. (2003). Teaching self-regulated learning through a "learning to learn" course. *Teaching of Psychology, 30*, 30-33.

*Jungert, T. & Rosander, M. (2010). Self-efficacy and strategies to influence the study environment. *Teaching in Higher Education, 15*, 647-659.

Kahn, J. H. & Nauta, M. M. (2001). Social-cognitive predictors of first-year college persistence: the importance of proximal assessment. *Research in Higher Education*, *42*, 633-652.

*Klobas, J. (2001). Student psychological response to computer-supported collaborative learning. *ASIST 2001: Proceedings of the 64th ASIST Annual Meeting*, *38*, 190-198.

Klomegah, R. Y. (2007). Predictors of academic performance of university students: an application of the goal efficacy model. *College Student Journal*, *41,* 407-415.

Knight, P. & Yorke, M. (2004). A new view of employability. In P. Knight & M. Yorke (Eds.), *Learning, Curriculum and Employability in Higher Education* (pp. 36-49). London: Routledge Farmer.

*Latham, G. P. (2006). The effect of learning vs. outcome goals on self-efficacy, satisfaction and performance in an MBA program. *Applied Psychology – An International Review*, *55*, 606-623.

*Lee S. Y., Hoerr S. L., Weatherspoon, L. & Schiffman R. F. (2008). Nutrition students improve attitudes after a guided experiential assignment with older adults. *Journal of Nutrition Education and Behavior, 40*, 279-287.

*Lent, R. W., Brown, S. D. & Larkin, K. C. (1986). Self-efficacy in the prediction of academic performance and perceived career options. *Journal of Counseling Psychology, 33,* 265-269.

*Lent R. W., Schmidt, J. & Schmidt, L. (2006). Collective efficacy beliefs in student work teams: relation to self-efficacy, cohesion, and performance. *Journal of Vocational Behavior, 68*, 73-84.

Lorig, K., Chastain, R. L., Ung, E., Shoor, S. & Holman, H. R. (1989). Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. *Arthritis and Rheumatism, 32*, 37-44.

*Luzzo D. A., Funk, D. P. & Strang, J. (1996). Attributional retraining increases career decision-making self- efficacy. *Career Development Quarterly, 44,* 378-386.

*Luzzo, D. A., Hasper, P., Albert, K. A., Bibby, M. A. & Martinelli, E. A. (1999). Effects of self-efficacy-enhancing interventions on the math/science self-efficacy and career interests, goals, and actions of career undecided college students. *Journal of Counselling Psychology, 46*, 233-243.

*Maples, M. R. & Luzzo, D. A. (2005). Evaluating DISCOVER's effectiveness in enhancing college students' social cognitive career development. *Career Development Quarterly, 53*, 274-285.

*Mathisen, G. E. & Bronnick, K. S. (2009). Creative self-efficacy: an intervention study. *International Journal of Educational Research, 48*, 21-29.

McArdle, S., Waters, L., Briscoe, J.  P. & Hall D. T. (2007). Employability during unemployment: adaptability, career identity and human and social capital. *Journal of Vocational Behavior, 71*, 247-264.

*McCabe. P. P., Kraemer L. A., Miller, P. M., Parmar, R. S. & Ruscica, M. B. (2006). The effect of text format upon underachieving first year college students' self-efficacy for reading and subsequent reading comprehension. *Journal of College Reading and Learning*, *37*, 19-44.

Mitchell, G. & Hirom, K. (2002, September). *The role of explanatory style in the academic underperformance of boys*. Paper presented at the European Conference on Educational Research, Lisbon.

Multon, K. D., Brown, S. D. & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: a Meta-analytic investigation. *Journal of Counselling Psychology*, *38*, 30-38.

Orbell, S. (2003). Personality systems interactions theory and the theory of planned behaviour: evidence that self-regulatory volitional components enhance enactment of studying behaviour. *British Journal of Social Psychology*, *42*, 95–112.

Overton, R. C. (1998). A comparison of fixed-effects and mixed (random-effects) models for meta-analysis tests of moderator variable effects. *Psychological Methods*, *3*, 354-379.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, *66*, 543-578.

Pajares, F. & Miller, M. D. (1995). Mathematics self-efficacy and mathematics performances: the need for specificity of assessment. *Journal of Counseling Psychology*, *42*, 190-198.

*Papinczak, T., Young, L., Groves, M. & Haynes, M. (2008). Effects of a metacognitive intervention on students' approaches to learning and self-efficacy in a first year medical course. *Advances in Health Sciences Education, 13,* 213-232.

*Rampp L. C. & Guffey, J. S. (1999). The impact of metacognition training on academic self-efficacy of selected underachieving college students. (ERIC Document Reproduction Services No. ED432 607.)

*Reed M. J., Kennett D. J., Lewis, T., Lund L. E., Stallberg, C. & Newbold I. L. (2009). The relative effects of university success courses and individualized interventions for students with learning sisabilities. *Higher Education Research and Development, 28,* 385-400.

Reed D. A. , Price E. G., Windish D. M., Wright S. M., Gozu A., Hsu E., Beach M. C., Kern D. & Bass, E. B. (2005). Challenges in systematic reviews of educational intervention studies. *Annals of Internal Medicine*, *142*, 1080-1089

Reese R. J. & Miller C. D. (2010a, b). Using outcome to improve a career development course: closing the scientist-practitioner gap. *Journal of Career Assessment, 18,* 207-219.

Rodgers, W.M., Conner, M. & Murray, T.C. (2008). Distinguishing among perceived control, perceived difficulty, and self-efficacy as determinants of intentions and behaviours. *British Journal of Social Psychology, 47, 607-630.*

Rosenberg, M. (1965). *Society and the Adolescent Self-image*. Princeton, NJ: Princeton University Press.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*, 638-641.

Ruthig, J. C., Haynes. T. L., Stupnisky, R. H. & Perry, R. P. (2009). Perceived academic control: mediating the effects of optimism and social support on college students' psychological health. *Social Psychology of Education*, *12*, 233-249.

Ruthig, J. C., Perry, R. P., Hall N. C. & Hladkyj, S. (2004). Optimism and attributional retraining: longitudinal effects on academic achievement, test anxiety, and voluntary course withdrawal in college students. *Journal of Applied Social Psychology, 34*, 709-730.

Sander, P. & Sanders, L. (2003). Measuring confidence in academic study: a summary report. *Electronic Journal of Research in Educational Psychology and Psychopedagogy, 1*, 1. Retrieved from: http://www.investigacion psicopedagogica.org/revista/articulos/1/english/Art_1_1.pdf

Sander, P. & Sanders, L. (2006). Understanding academic confidence. *Psychology Teaching Review*, *12*, 29-39.

Sander, P., & Sanders, L. (2009). Measuring Academic Behavioural Confidence: the ABC Scale Revisited*. Studies in Higher Education, 34*, 19-35.

*Schunk, D. H. & Ertmer, P. A. (1999a, b). Self-regulatory processes during computer skill acquisition: goal and self-evaluative influences. *Journal of Educational Psychology, 91*, 251-260.

*Scott A. B. & Ciani K. D. (2008). Effects of an undergraduate career class on men's and women's career decision-making self-efficacy and vocational identity. *Journal of Career Development, 34*, 263-285.

Shonkoff, J. & Phillips, D. (2000) (Eds.). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, DC: National Academy Press.

Seligman, M. E. P. (1998). *Learned Optimism*. New York. Pocket Books.

Sitzmann, T. & Yeo, G. (2013). A meta-analytic investigation of the within-person self-efficacy domain: is self-efficacy a product of past performance or a driver of future performance? *Personnel Psychology, 66*, 531-668.

*Smith J. M. (1994). The effects of education on computer self-efficacy. *Journal of Industrial Teacher Education, 31,* 51-65.

*Sobral D. T. (1997). Improving learning skills: a self-help group approach. *Higher Education, 33*, 39-50.

*Sullivan K. R. & Mahalik J. R. (2000). Increasing career self-efficacy for women: evaluating a group intervention. *Journal of Counseling and Development, 78*, 54-62.

Stajkovic, A. D. & Luthans, F. (1998). Self-efficacy and work-related performance: a meta-anlaysis. *Psychological Bulletin, 124*, 240-261.

Stark, J. S., Bentley, R. J., Lowther, M. A. & Shaw, K. M. (1991). The student goals exploration: reliability and concurrent validity. *Educational and Psychological Measurement, 51,* 413-422.

Thompson Rueters. (2011). EndNote: A computer program for research citation (Version X5) [Computer software]. New York, NJ: Thompson Reuters.

*Trawick, L. (1992). Effects of a cognitive-behavioral intervention on the motivation, volition, and achievement of academically underprepared college students. *Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco*, 49).

*Uffelman, R. A., Subich, L. M., Diegelman, N. M., Wagner, K. S. & Bardash, R. J. (2004). Effect of mode of interest assessment on clients' career decision-making self-efficacy. *Journal of Career Assessment, 12*, 366-380.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N.M., Senecal, C. & Vallieres, E. F.

(1992). The Academic Motivation Scale: a measure of intrinsic, extrinsic and a motivation in education. *Educational and Psychological Measurement, 52*, 1003-1017.

*Vancouver, J. B., Thompson, C. M., Tischner, E. & Putka, D. J. (2002). Two studies examining the negative effect of self-efficacy on performance. *Journal of Applied Psychology*, *87*, 506-516.

*Vrugt A. J., Hoogstraten, J. & Langereis M. P. (1997). Academic self-efficacy and malleability of relevant capabilities as predictors of exam performance. *Journal of Experimental Education, 66*, 61-72.

Wanberg, C. R., Zhang, Z. & Diehn, E. W. (2010). Development of the "Getting Ready for your next job" Inventory for Unemployed Individuals. *Personnel Psychology, 63*, 439-478.

Wanberg, C. R., Zhu, J. & Van Hooft, E. A. J. (2010). The job search grind: perceived progress, self-reactions and self-regulation of search effort. *Academy of Management, 3*, 788-807.

*Wang S-L. & Wu P-Y. (2008). The role of feedback and self-efficacy on web-based learning: the social cognitive perspective. *Computers and Education, 51*, 1589-1598.

*Watters J. J. & Ginns, I. S. (1997). Peer assisted learning: impact on self-efficacy and achievement. *Paper presented at the Annual Meeting of the American Educational Research Association (Chicago*, 14).

Wilson, D. & Lipsey, M. (2001*). Practical Meta-analysis*. Thousand Oaks, CA: Sage Publications.

Wittekind, A., Raeder, S. & Grote, G. (2010). A longitudinal study of determinants of perceived employability. *Journal of Organizational Behavior*, *31*, 566-586.

Wood, R. E. & Bandura, A. (1989). Impact of conceptions of ability on self- regulatory mechanisms and complex decision making. *Journal of Personality and Social Psychology*, *56*, 407-415.

*Zorkina, Y. & Nalbone, D. P. (2003). Effect of induced level of confidence on college students performance on a cognitive test. *Current Research in Social Psychology, 8*, 148-162.

Table 1

*Summary Coding Intervention, Sample, and Study Characteristics*

| | | | Intervention Characteristics | | | | | Sample Characteristics | | | Study Characteristics | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Study** | **Design** | **Duration** | **Follow-up** | **Foundation** | **Class** | **Nature** | **Intent** | ***N*** | **Country** | **Level** | **Type** | **Outcome** | **Report** |
| Berzoff, Dane & Cait, 2005 | Q | Y | Y | N | M | T | S/P | 23/25 | US | H | P | C | C |
| Betz & Schifano., 2000 | Q | O | Y | T | M | T | P | 54 | US | M | P | O | S |
| Betz & Borgen, 2009 | Q | S/O | N | T | M | T | S | 867 | US | M | P | C | S |
| Bresó, Schaufeli & Salanova 2011 | Q | S | Y | T | M | T | P | 71 | Spain | M | P | O | S |
| Brown & Morrissey, 2004 | E | O | N | T | U | R | P | 65 | Canada | M | P | C | S |
| Butler, 1997 | Q | Y | N | T | M | R/T | S | 36 | Canada | L | U | P | S |
| Chyung et al., 1998 | Q | S | N/Y | N | U | R | S | 24 | US | H/M | U | C | S |
| Cordero, Porter, Israel & Brown, 2010 | E | O | Y | T | U | T | P | 99 | US | M | P | O | S |
| Duijnhower et al., 2010 | Q | S/O | N | T | U | R | S | 65 | Nether | M | P | P(A) | I |
| Elander, Pittam, Lusher, Fox & Payne, 2010 | Q | S | N | N | M/U | T/R | S | 279 | UK | M | P | C | S |
| Fletcher, 2005 | Q | S | N | T | U | R | P | 64 | Aust | M | P | P(A) | I |
| Ford-Gilboe, Laschinger, Laforet-Fleisser, Foran & Ward-Griffin, 1997 | Q | Y | N | T | M | R | P | 34 | Canada | M | P | P(A) | C |
| Freedman, 1996 | Q | O | N | T | U | R | P | 70 | US | L | U | C | C |
| Gaudine & Saks, 2004 | Q | O | Y | T/N | M | T | S | 147 | Canada | H/M | P | C | C |
| Graham, 2008 | Q | S | N | T | M | R | S | 144 | US | L | U | C | S |
| Grier & Skaar, 2010 | Q | S | N | T | U | T | P | 82 | US | L | P | O | S |
| Griffin & Griffin, 1997 | Q | S | N | T | M | R | P | 47 | US | M | U | C | C |
| Hanlon & Schneider, 1999 | Q | O | N | T | U | R | P | 17 | US | L | U | C | C |
| Hofer & Yu, 2003 | Q | S | N | T | R | T | S | 78 | US | M | P | O | S |
| Jungert & Rosander, 2010 | Q | O | N | T | U | T | P | 213 | Sweden | H | P | O | S |
| Klobas, 2001 | Q | S | Y | T | M | R/T | P | 69 | Italy | M | U | O | S |
| Latham, 2006 | Q | Y | N | T | U | T | P | 125 | Canada | H | P | C | S |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lee, Hoerr, Wetherspoon & Schiffman, 2008 | Q | S | N | N | M | T | S | 100 | US | M | P | P(A)/C | C |
| Lent, Brown & Larkin, 1986 | Q | S | Y | T | U | R | P | 42 | US | M | P | P(A) | C |
| Lent, Schmidt & Schmidt, 2006 | Q | S/O | N | T | U | R | P | 312 | US | M | P | C | C |
| Luzzo, Funk & Strang, 1996 | Q | O | N | T | U | T | P | 60 | US | M | P | O | S |
| Luzzo, Hasper, Albert, Bibby & Martinelli, 1999 | Q | O | Y | T | U | T | P | 94 | US | L/M | P | O | S |
| Maples & Luzzo, 2005 | E | O | N | T | M | T | P | 34 | US | L/M | P | O | S |
| Mathisen & Bronnick, 2009 | Q | O | Y | T | U | T | P | 195 | Norway | H | P | O | C |

| | Intervention Characteristics | | | | | | | Sample Characteristics | | | Study Characteristics | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Study | Design | Duration | Follow-up | Foundation | Class | Nature | Intent | N | Country | Level | Type | Outcome | Report |
| McCabe, Kraemer, Miller, Parmar & Ruscica, 2006 | Q | O | N | T | U | R | P | 76 | US | L | P | P | C |
| Papinczak et al., 2008 | Q | O/Y | N | T | R | T | S | 213 | Aust | H | P | C | C |
| Rampp & Guffey 1999 | E | *S* | *N* | T | M | T | P | 25 | US | L | U | O | S |
| Reed et al., 2009 | Q | *S* | *N* | T/N | M | T | S | 27 | Canada | L | P | C | C |
| Reese & Miller, 2010a | Q | *S* | *N* | T | U | T | P | 133 | US | M | P | O | S |
| Reese & Miller, 2010b | Q | *S* | *Y* | T | U | T | P | 145 | US | M | P | O | S |
| Schunk & Ertmer, 1999a | E | *S* | *N* | T | R | T | S | 44 | US | M | P | O | S |
| Schunk & Ertmer, 1999b | E | S | N | T | R | T | S | 33 | US | M | P | O | S |
| Scott & Ciani, 2008 | Q | S | N | T | U | T | P | 88 | US | M | P | O | S |
| Smith, 1994 | Q | S | N | T | U | R | P | 147 | US | M | P | O | S |
| Sobral, 1997 | Q | S | N | T | R | T | S | 128 | Brazil | H | p | O | S |
| Sullivan & Mahalik, 2000 | Q | O | Y | T | M | T | P | 61 | US | L/M | P | O | S |
| Trawick, 1992 | E | O | Y | T | R | T | P | 79 | US | L | U | O | S |
| Uffelman, Subich, Diegelman, Wagner & Bardash, 2004 | Q | O | N | T | U | T | P | 81 | US | L | P | O | S |
| Vancouver, Thompson, Tischner & Putka, 2002 | E | O | N | T | U | R | P | 83 | US | M | P | P | I/C |
| Vrugt, Hoogstraten & Langereis, 1997 | Q | S/O | N | T | M | R | P | 438 | US | M | P | P | C |
| Wang & Wu, 2008 | Q | S | N | T | U | R | P | 76 | Taiwan | M | P | P | C |
| Watters & Ginns, 1997 | Q | S | Y | T | M | T | S | 124 | Aust | M | U | P | S |
| Zorkina & Nalbone, 2003 | E | O | N | T | U | R | P | 30 | US | M | P | C | S |

*Note*. Character = Intervention Design: E = Experimental; Q =Quasi Experimental. Intervention Duration: S = Semester; Y = Year Long; O= Other. Intervention Follow-up: Y = Yes; N= No. Intervention Characteristics (Character): N = Non-Theoretical; T = Theoretical. Intervention Classification (Class): U = Unistructural; M = Multistructural; or R = Relational. Intervention Intent (Intent): P = Primary Focus; S = Secondary Focus. Intervention Nature (Nature): R = Reproductive: T = Transformational. N = Sample Size. Country: UK= United Kingdom; US = United States. Student Level (Level): L = Low Ability; M = Mixed Ability; H = High Ability; U = Underachievers. Publication Type (Type): P = Published; U = Unpublished study. Outcome Measure (Outcome): O = Orientation; P(A) = Performance(attainment); P = Process; C = Combination. Study Reporting Method (Report): I = Instructor report; S = Student report; C = Combined instructor and student report.

Table 2.

*Study Quality Coding Scores for Studies Meeting Inclusion Criteria*

| Study | Study Design | Randomization | Descriptive Outcome | Outcome Reporting | Measurement Statistics | Replication | Total |
|---|---|---|---|---|---|---|---|
| Berzoff et al., 2005 | 1 | 0 | 0 | 1 | 0 | 1 | 3 |
| Betz & Schifano, 2000 | 1 | 1 | 1 | 0 | 1 | 1 | 5 |
| Betz & Borgen, 2009 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| Bresó et al., 2011 | 1 | 0 | 1 | 1 | 1 | 1 | 5 |
| Brown & Morrissey, 2004 | 1 | 1 | 1 | 0 | 0 | 1 | 4 |
| Butler, 1997 | 1 | 0 | 1 | 0 | 0 | 0 | 2 |
| Chyung et al., 1998 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| Cordero et al., 2010 | 1 | 1 | 1 | 0 | 1 | 1 | 5 |
| Duijnhower et al., 2010 | 1 | 0 | 1 | 1 | 1 | 1 | 5 |
| Elander et al., 2010 | 1 | 0 | 1 | 1 | 0 | 1 | 4 |
| Fletcher, 2005 | 1 | 0 | 0 | 0 | 1 | 0 | 2 |
| Ford-Gilboe et al., 1997 | 1 | 0 | 0 | 0 | 1 | 1 | 3 |
| Freedman, 1996 | 1 | 0 | 1 | 1 | 1 | 1 | 5 |
| Gaudine & Saks, 2004 | 1 | 1 | 1 | 0 | 1 | 1 | 5 |
| Graham, 2008 | 1 | 0 | 1 | 1 | 1 | 1 | 5 |
| Grier & Skaar., 2010 | 1 | 0 | 0 | 0 | 1 | 1 | 3 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Griffin & Griffin, 1997 | 1 | 0 | 1 | 1 | 1 | 1 | 5 |
| Hanlon & Schneider, 1999 | 1 | 0 | 1 | 1 | 0 | 1 | 4 |
| Hofer & Yu, 2003 | 1 | 0 | 1 | 1 | 1 | 0 | 4 |
| Jungert & Rosander, 2010 | 1 | 0 | 0 | 0 | 0 | 1 | 2 |
| Klobas et al., 2001 | 1 | 0 | 1 | 0 | 0 | 1 | 3 |
| Latham et al., 2006 | 1 | 0 | 1 | 0 | 0 | 1 | 3 |
| Lee et al., 2008 | 1 | 0 | 1 | 1 | 0 | 1 | 4 |

| Study | Study Design | Randomization | Sample Baseline | Outcome Reporting | Measurement Stats | Replication | Total |
|---|---|---|---|---|---|---|---|
| Lent et al., 1986 | 1 | 0 | 1 | 1 | 1 | 1 | 5 |
| Lent et al., 2006 | 1 | 0 | 0 | 1 | 1 | 1 | 4 |
| Luzzo et al., 1996 | 1 | 0 | 1 | 1 | 1 | 1 | 5 |
| Luzzo et al., 1999 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| Maples & Luzzo, 2005 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| Mathisen & Bronnick, 2009 | 1 | 0 | 1 | 1 | 1 | 1 | 5 |
| McCabe et al., 2006 | 1 | 0 | 1 | 0 | 0 | 1 | 3 |
| Papinczak et al., 2008 | 1 | 0 | 1 | 0 | 0 | 1 | 3 |
| Rampp & Guffey, 1999 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| Reed et al., 2009 | 1 | 0 | 1 | 1 | 1 | 1 | 5 |
| Reese & Miller, 2010a | 1 | 0 | 1 | 0 | 1 | 1 | 4 |
| Reese & Miller,2010b | 1 | 0 | 1 | 0 | 1 | 1 | 4 |
| Schunk & Ertmer, 1999a | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| Schunk & Ertmer, 1999b | 0 | 1 | 1 | 1 | 1 | 1 | 5 |
| Scott & Ciani, 2008 | 1 | 0 | 1 | 1 | 1 | 1 | 5 |
| Smith, 1994 | 1 | 0 | 0 | 1 | 1 | 1 | 4 |
| Sobral, 1997 | 1 | 0 | 1 | 1 | 1 | 1 | 5 |
| Sullivan & Mahalik, 2000 | 1 | 0 | 1 | 0 | 1 | 1 | 4 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Trawick, 1992 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| Uffelman et al., 2004 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| Vancouver et al., 2002 | 1 | 1 | 1 | 1 | 0 | 1 | 5 |
| Vrugt et al., 1997 | 1 | 0 | 0 | 1 | 1 | 1 | 4 |
| Wang & Wu, 2008 | 1 | 0 | 0 | 0 | 1 | 1 | 3 |
| Watters & Ginns, 1997 | 1 | 0 | 0 | 1 | 1 | 1 | 4 |
| Zorkina & Nalbone, 2003 | 1 | 1 | 0 | 0 | 1 | 1 | 4 |

*Note*. Study design = rationale for study design presented. Randomization = process used to assign subjects to experimental or control conditions. Descriptive Outcome = descriptive and outcome information reported on sample(s) at the beginning of study. Measurement statistics = provided information concerning validity and reliability of outcome measures. Replication = sufficient information provided to replicate study findings. Scores range from 0 to 6 with larger scores representing quality.

Table 3

*Outcome Analyses*

| | | Effect Size Statistics | | | | Null Test | Homogeneity Statistics | | | Publication Bias |
|---|---|---|---|---|---|---|---|---|---|---|
| **Variable** | ***k*** | ***g*** | ***SE*** | ***$s^2$*** | ***95% C.I.*** | ***Z*** | ***Q*** | ***$\tau^2$*** | ***$I^2$*** | **Fail Safe *N*** |
| Random Effects Model[A] | 49 | 0.273 | 0.041 | 0.002 | (0.194, 0.353) | 6.725* | 208.6* | 0.045 | 76.99 | 1080 |
| **Process** | | | | | | | | | | |
| Self-Efficacy | 40 | 0.289 | 0.047 | 0.002 | (0.197, 0.382) | 6.144* | 231.6* | 0.051 | 83.16 | 1304 |
| Knowledge/Intellect | 7 | 0.598 | 0.163 | 0.027 | (0.278, 0.917) | 3.662* | 59.13* | 0.120 | 89.85 | 124 |
| Learning Strategies | 15 | 0.255 | 0.076 | 0.006 | (0.107, 0.404) | 3.370* | 111.7* | 0.063 | 87.40 | 266 |
| **Orientation** | | | | | | | | | | |
| Attitude/Interest | 6 | 0.319 | 0.163 | 0.027 | (-0.001, 0.639) | 1.955 | 22.54* | 0.111 | 77.81 | 17 |
| Self-Regulation | 8 | 0.229 | 0.043 | 0.002 | (0.144, 0.314) | 5.290* | 3.119 | 0.000 | 0.000 | 54 |
| Motivation | 6 | 0.087 | 0.079 | 0.006 | (-0.069, 0.242) | 1.093 | 6.380 | 0.008 | 21.63 | 6 |
| Self-Perceptions | 6 | 0.368 | 0.190 | 0.036 | (-0.005, 0.741) | 1.935 | 42.67* | 0.156 | 88.28 | 46 |
| Social Skills | 4 | 0.241 | 0.151 | 0.023 | (-0.055, 0.538) | 1.595 | 23.73* | 0.072 | 87.37 | 13 |
| Professional Aspirations | 7 | 0.269 | 0.089 | 0.008 | (0.094, 0.443) | 3.019* | 12.18 | 0.024 | 50.75 | 41 |
| Anxiety | 4 | -0.134 | 0.181 | 0.033 | (-0.489, 0.222) | -0.736 | 6.899 | 0.072 | 56.52 | 0 |
| **Performance** | | | | | | | | | | |
| Academic Attainment | 14 | 0.259 | 0.088 | 0.008 | (0.085, 0.432) | 2.925* | 39.30 | 0.063 | 66.92 | 96 |

Note. *p $\leq$ .05. A=Total Q-value used to determine heterogeneity; B=Between Q-value used to determine significant differences between moderators. *k* = number of effect sizes. *g* = Effect size (Hedges *g*). *SE* = Standard Error. S*2* = variance. 95% *C.I.*= Confidence Intervals (lower limit, upper limit). Z = test of the null hypothesis. $\tau^2$ = Between study variance in Random Effects Model. $I^2$ = Total variance explained by moderators. A = Total *Q*-value used to determine heterogeneity. Fail Safe *N* = number of studies needed to increase p > .05.

Table 4

*Intervention Moderator Statistics*

| | Effect Size Descriptive Statistics | | | | | Null Test | Heterogeneity Statistics | | |
|---|---|---|---|---|---|---|---|---|---|
| | *k* | *g* | *SE* | $s^2$ | 95% *C.I.* | *Z* | *Q* | $\tau^2$ | $I^2$ |
| **Random Effects Model**[A] | 49 | 0.273 | 0.041 | 0.002 | (0.194, 0.353) | 6.725* | 208.6* | 0.045 | 76.99 |
| | | | | | | | | | |
| **Intervention Characteristics**[B] | | | | | | | | | |
| Design | | | | | | | 2.182 [B] | | |
| Experimental | 10 | 0.429 | 0.112 | 0.013 | (0.209, 0.648) | 3.429* | | 0.050 | 39.62 |
| Quasi Experimental | 39 | 0.251 | 0.044 | 0.002 | (0.166, 0.337) | 5.504* | | 0.044 | 79.57 |
| Duration | | | | | | | 0.457 [B] | | |
| Other | 19 | 0.301 | 0.070 | 0.005 | (0.164, 0.438) | 4.316* | | 0.020 | 67.54 |
| Semester | 26 | 0.251 | 0.055 | 0.003 | (0.143, 0.359) | 4.562* | | 0.472 | 94.68 |
| Year | 4 | 0.327 | 0.144 | 0.021 | (0.045 0.609) | 2.276* | | 0.093 | 73.55 |
| Follow-up | | | | | | | 0.920 [B] | | |
| No | 37 | 0.278 | 0.047 | 0.002 | (0.187 0.370) | 5.944* | | 0.046 | 79.97 |
| Yes | 12 | 0.268 | 0.087 | 0.008 | (0.098, 0.439) | 3.091* | | 0.051 | 58.11 |
| Foundation | | | | | | | 1.050 [B] | | |
| Atheoretical | 4 | 0.438 | 0.164 | 0.027 | (0.117, 0.759) | 2.676* | | 0.045 | 77.29 |
| Theoretical | 45 | 0.264 | 0.043 | 0.002 | (0.180, 0.349) | 6.139* | | 0.413 | 88.91 |
| Classification | | | | | | | 1.191 [B] | | |
| Multistructural | 18 | 0.298 | 0.072 | 0.005 | (0.157, 0.439) | 4.142* | | 0.071 | 79.14 |
| Relational | 6 | 0.163 | 0.117 | 0.014 | (-0.067, 0.393) | 1.391 | | 0.037 | 77.60 |
| Unistructural | 25 | 0.291 | 0.058 | 0.003 | (0.177, 0.405) | 5.006* | | 0.083 | 79.14 |
| Nature | | | | | | | 0.220 [B] | | |
| Reproductive | 20 | 0.299 | 0.065 | 0.004 | (0.171, 0.427) | 4.586* | | 0.039 | 72.96 |
| Transformational | 29 | 0.260 | 0.054 | 0.003 | (0.153, 0.366) | 4.785* | | 0.056 | 79.32 |
| Intent | | | | | | | 2.449 [B] | | |
| Primary | 33 | 0.332 | 0.051 | 0.003 | (0.222, 0.422) | 5.915* | | 0.057 | 75.58 |
| Secondary | 16 | 0.184 | 0.072 | 0.005 | (0.043, 0.325) | 2.514* | | 0.039 | 79.20 |
| Quality | | | | | | | 6.601* [B] | | |
| Low | 4 | 0.640 | 0.165 | 0.027 | (0.316, 0.963) | 3.875* | | 0.285 | 83.24 |
| Medium | 22 | 0.207 | 0.057 | 0.003 | (0.096, 0.318) | 3.633* | | 0.042 | 78.58 |
| High | 23 | 0.308 | 0.064 | 0.004 | (0.183, 0.433) | 4.836* | | 0.041 | 68.45 |

Note. A=Total Q-value used to determine heterogeneity; B=Between Q-value used to determine significant differences between moderators. *k* = number of effect sizes. *g* = Effect size (Hedges *g*). *SE* = Standard Error. S*2* = variance. 95% *C.I.*= Confidence Intervals (lower limit, upper limit). Z = test of the null hypothesis. $\tau^2$ = Between study variance in Random Effects Model. $I^2$ = Total variance explained by moderators. *p $\leq$ .05. A = Total *Q*-value used to determine heterogeneity. B = Between *Q*-value used to determine significant ($\alpha$ = .01) differences between moderators.

Table 5

*Sample and Study Moderator Statistics*

| | | Effect Size Descriptive Statistics | | | | Null Test | Heterogeneity Statistics | | |
|---|---|---|---|---|---|---|---|---|---|
| | ***k*** | ***g*** | ***SE*** | $s^2$ | **95% *C.I.*** | ***Z*** | ***Q*** | $\tau^2$ | $I^2$ |
| Random Effects Model[A] | 49 | 0.273 | 0.041 | 0.002 | (0.194, 0.353) | 6.725* | 208.6* | 0.045 | 76.99 |
| **Sample Characteristics**[B] | | | | | | | | | |
| Country | | | | | | | 14.66 [B] | | |
| Australia | 3 | 0.080 | 0.150 | 0.024 | (-0.215, 0.374) | 0.531 | | 0.093 | 83.58 |
| Brazil | 1 | 0.265 | 0.238 | 0.059 | (-0.200, 0.731) | 1.117 | | 0.000 | 0.000 |
| Canada | 6 | 0.157 | 0.118 | 0.015 | (-0.074, 0.388) | 1.331 | | 0.257 | 90.96 |
| Italy | 1 | 0.073 | 0.251 | 0.123 | (-0.419, 0.565) | 0.291 | | 0.000 | 0.000 |
| Netherlands | 1 | -0.542 | 0.333 | 0.115 | (-1.195, 0.111) | -1.628 | | 0.000 | 0.000 |
| Norway | 1 | 0.509 | 0.259 | 0.070 | (0.002, 1.017) | 1.969* | | 0.000 | 0.000 |
| Spain | 1 | 0.147 | 0.378 | 0.149 | (-0.593, 0.888) | 0.391 | | 0.000 | 0.000 |
| Sweden | 1 | 0.575 | 0.271 | 0.076 | (0.043, 1.107) | 2.118* | | 0.000 | 0.000 |
| Taiwan | 1 | 0.115 | 0.233 | 0.057 | (-0.333, 0.572) | 0.491 | | 0.000 | 0.000 |
| UK | 1 | 0.228 | 0.233 | 0.054 | (-0.337, 0.555) | 0.478 | | 0.000 | 0.000 |
| US | 32 | 0.353 | 0.053 | 0.003 | (0.249, 0.458) | 6.163* | | 0.027 | 65.24 |
| Level | | | | | | | 1.717 [B] | | |
| Low Ability | 13 | 0.339 | 0.082 | 0.007 | (0.177, 0.500) | 4.113* | | 0.020 | 44.63 |
| Mixed Ability | 28 | 0.276 | 0.053 | 0.003 | (0.172, 0.380) | 5.198* | | 0.022 | 63.99 |
| High Ability | 8 | 0.174 | 0.095 | 0.009 | (-0.013, 0.361) | 1.822 | | 0.172 | 92.20 |
| **Study Characteristics**[B] | | | | | | | | | |
| Type | | | | | | | 0.109 [B] | | |
| Published | 39 | 0.269 | 0.046 | 0.002 | (0.180, 0.358) | 5.898* | | 0.045 | 78.48 |
| Unpublished | 10 | 0.305 | 0.100 | 0.010 | (0.110 0.500) | 3.064* | | 0.074 | 71.33 |
| Outcome | | | | | | | 5.508 [B] | | |
| | 22 | 0.359 | 0.060 | 0.004 | (0.241, 0.477) | 5.937* | | 0.028 | 50.10 |
| Performance | 11 | 0.281 | 0.078 | 0.006 | (0.128 0.434) | 3.592* | | 0.047 | 74.29 |
| Combined | 16 | 0.164 | 0.068 | 0.005 | (0.012, 0.279) | 2.132* | | 0.043 | 83.47 |
| Reporting Method | | | | | | | 0.505 [B] | | |
| Student Report | 30 | 0.300 | 0.054 | 0.003 | (0.194, 0.406) | 5.543* | | 0.046 | 76.50 |
| Instructor Report | 4 | 0.241 | 0.158 | 0.025 | (-0.068, 0.550) | 1.527 | | 0.199 | 80.01 |
| Combined Method | 15 | 0.241 | 0.074 | 0.005 | (0.097, 0.385) | 3.275* | | 0.050 | 79.63 |

Note. A=Total Q-value used to determine heterogeneity; B=Between Q-value used to determine significant differences between moderators. *k* = number of effect sizes. *g* = Effect size (Hedges *g*). *SE* = Standard Error. S*2* = variance. 95% *C.I.* = Confidence Intervals (lower limit, upper limit). Z = test of the null hypothesis. $\tau^2$ = Between study variance in Random Effects Model. $I^2$ = Total variance explained by moderators. *p $\leq$ .05. A = Total *Q*-value used to determine heterogeneity. B = Between *Q*-value used to determine significant (α = .01) differences between moderators.

**Acknowledgement**

We wish to thank Dr Victoria Scaife (School of Psychology, University of East Anglia, UK) for her contribution to an earlier draft this manuscript.