FACULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES

The Measurement of Implicit Motives in Applied Settings

Jan Malte Runge

Supervisor: Prof. Dr. Jonas W. B. Lang

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Doctor of Psychology

Academic year 2018–2019

GHENT
UNIVERSITY

WORDS OF THANKS

I am happy and satisfied that I have finished my dissertation. I owe the fact that my PhD went smoothly overall to all those who have supported me over the past few years.

First of all, and in particular, I would like to express my deepest appreciation to *Jonas* as a promoter for the unique opportunity he has given and the faith he always had in me. Uncountable mails and messages have guided me through the academic jungle in the last four years. Thank you for showing me the formal rules of how to be a good researcher, the informal tips and tricks, and the inside knowledge of how to behave in this world unknown to me. You were always there with guidance and support in lightning speed, at office hours, evenings, weekends and in the night (not in the mornings though) to share successes, discuss specific problems, and to tell me why my new idea is really not as great as I first thought. With you, the more I gave, the more I received—for me it developed into a successful team play that made me very satisfied. I can only hope that you are also satisfied with me. If nothing else, you guided me to be an optimistic researcher enjoying this exiting profession anew every day.

I also wish to thank my guidance committee for listening to my research talks, reading my papers, and giving high quality feedback that improved my work. I thank Jan De Houwer, for probing questions that challenged me to scrutinize my work and helped me to improve it. I thank Filip Lievens for his genuine interest, intelligent remarks, and letting me tap into his invaluable knowledge. I thank Frederik Anseel for asking the seemingly easy questions that I tend to overlook, which helped me tremendously to learn how to frame my studies, and for being the most supportive audience I can imagine. I thank Nicola Baumann for showing me a first look into the academic world, sparking my interest in implicit motives, and an encouraging positive attitude. I further want to thank all coauthors on their valuable feedback in all stages of the projects we did together.

I also want to give special thanks to the old and new colleagues that helped me on my way. Special thanks goes to Stefan Engeser, who guided my first research project when I was a student, and showed me how rigorous and detailed one needs to work in academia, he sparked the passion for research in me; special thanks goes also to Marjolein Feys and Bart Wille who helped me settle in Belgium in many different ways—thank you for everything at work and outside of work; special thanks goes also to Christoph Herde who was always encouraging and interested in my research—thank you for countless talks about research and soccer; special thanks go also Gudrun Reindl, for going through this together and supporting me in so many ways—thank you for becoming a real friend; another words of thanks goes to my office colleagues for the fun we had during the last years: Celine, Sam, Ruben, and Sirarpi—with me gone, you have a chance to win the darts week too.

I also want to thank my parents, who have always supported me in my education in many ways. Without their support, I would not be where I am now, and I am very grateful for it.

Dear Nathalie, I cannot thank you enough. For deciding to come to Gent with me without even having seen the place—because I thought it is a good idea (it was). For sharing my excitement when I had a new idea and being happy when something went well. For listening to my confused thoughts after a day of thinking. For distraction when things were difficult and encouraging me when things felt unsolvable. For support, tea, and food during long evenings and weekends. For giving up much of your own and covering my back for everything else so I could focus on my dissertation. For giving me and this dissertation an unreasonable amount of time and place in our family. For letting me chase my dreams.

Malte Runge

May 2019

TABLE OF CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

"Measuring things better advances everything that we do."
(Cecil Reynolds, April 2018)

In the introduction of this dissertation, we describe how two personality theories that were proposed at the same time—motives and traits—took a very different development. One developed into the most widely applied personality theory (traits) and the other plays a much less popular role (motives). We argue that a major reason for the success of one and the lagging behind of the other lays in differences closely tied to their measurement. While trait theory profited from advancements in psychometrics and a straightforward measurement procedure, the measurement of implicit motives suffered from an inconvenient expert coding procedure and a response process that is difficult to model. We then describe how this dissertation contributes to measuring implicit motives better *to advance research and applications in the field of Organizational Behavior (OB). The goal of this dissertation is, thus, to innovate the psychometric foundation as well as the measurement procedure of implicit motives to contribute to closing the gap between motive and trait research in OB.*

Definition of Implicit Motives

We start with defining implicit motives and continue with a short history. The basic concept of implicit motives is that people differ in what motivates them not only on a conscious level but also on a nonconscious level. A comprehensive and concise definition describes implicit motives as "motivational dispositions that operate outside of a person's conscious awareness and are aimed at the attainment of specific classes of incentives and the avoidance of specific classes of disincentives" (Schultheiss & Brunstein, 2010, p. 603).

This definition addresses three specific properties of implicit motives. First, implicit motives operate outside of a person's conscious awareness which is reflected in the term *implicit*. Second, different classes of incentives are mentioned. Researchers typically focus on the Big Three implicit motives: the *affiliation*, the *achievement,* and the *power* motive. An affiliation motivated person has a higher incentive to build and maintain positive emotional relationships with others. An achievement motivated person has a high incentive to do well in challenging tasks and improve their performance. Power motivated people have a high incentive to exert influence on others. The individual differences in incentives for certain classes of behavior are tied to individual differences in being sensible for (emotional) rewards that come with these classes of behavior. The third aspect of this definition refers to two different directions in which implicit motives operate. The disposition to attain an implicit motive—or *approach* motivation—refers to the motive to reach an incentive that is relevant for the motive. Avoidance motivation refers to the disposition to prevent not reaching a motive relevant incentive. For instance, a person who is approach motivated for achievement wants to do well in an exam and a person who is avoidance motivated for achievement wants to not fail in an exam.

We use this definition in the present dissertation as a foundation and contribute to two of the properties. The first property entails that implicit motives operate outside of a person's awareness which makes researchers measure them indirectly. In implicit motive measures, respondents see ambiguous pictures and write stories in response to those pictures, which are then evaluated by expert coders for motivational themes. In this dissertation, we challenge the view that respondents cannot categorize their implicit motives themselves with the development of a new measurement procedure. The second

property of the definition refers to the existence of three implicit motives. Previous psychometric work focused on modeling only a single motive (Kuhl, 1978; Tuerlinckx, De Boeck, & Lens, 2002). We continue a recently introduced psychometric approach (Lang, 2014) that models a response process including all three motives. The third property referring to approach and avoidance motivation plays a less important role in this dissertation—we mainly focus on modeling approach motivation.

The Roots of Implicit Motive Research

What researchers study today under the term " implicit motives" has its roots in ancient Greek philosophy (Winter, John, Stewart, Klohnen, & Duncan, 1998). Philosophers have been interested in understanding, studying and explaining what motivates people and how people differ in that. For instance, Empedocles (5th century B.C.) described two fundamental *motive forces*: love, the need to unite and fuse, and strife, the need to dissolve fusions. Building on Empedocles dualism of life and death instincts, Freud later theorized that people differ in motivational forces (Freud, 1937/1964, pp. 245-247). Researchers were interested in interindividual differences in motivational forces from the early days of modern psychology.

The founding father of contemporary research on motives was Henry Murray, who developed a broad framework to describe and analyze personality in the 1930ies at Harvard University (Murray, 1938). The fundament of his theory on personality—that he named personology—is based on differences between people's motivational *needs*. The core is a set of universal needs where people differ in need strength. The basic principle behind those needs is that a high need creates a disequilibrium and individuals strive for behaviors satisfying the need to reduce tension. To measure individual differences in motivational needs, Murray developed a test procedure based on coding imaginative stories written in response to ambiguous pictures, the Thematic Apperception Test (TAT; Morgan & Murray, 1935). His idea was that people project their needs on characters depicted in the pictures which then reflects in the written stories. This procedure introduced a standardized measure of Freud's principle of projection.

The Roots of Personality Traits – the Other Big Personality Theory of that Time

At the same time and in the same place, Gordon Allport built a different theory of personality based on personality traits (Allport, 1931). Traits are more similar to Jung's idea of personality types using the concepts of extraversion and introversion. Allport described a grand vision for the field of personality and put characteristic behaviors and thoughts of an individual person in the focus of his studies. For him, traits were basic and explained a considerable consistency in a person's behavior. An advantage of personality traits as envisioned by Allport was that one could ask people directly for their characteristic behavior and thoughts using questionnaires.

The development of psychometric methods—correlational techniques and factor analysis (Cattell, 1952; Spearman, 1904; Thurstone, 1931)—facilitated the development of trait theories. Using factor analysis, Cattell (1946, 1950) developed a comprehensive system of personality with the goal to predict people's behavior. For Cattell, traits were a mental structure that could be inferred from observable behavior. The underlying mental structure would explain consistency in behavior over time. Another popular factor theory influencing the progress of research on personality was suggested by Hans Eysenck (1952) and has a strong influence on personality research up until today. He described human personality with three basic traits: introversion-extraversion, neuroticism, and psychoticism. Although trait psychology experienced a major crisis in the 1970ies and 1980ies, it recovered relatively well with the rise of the Five Factor Model of personality (McCrae & John, 1992) and is today the major framework to describe personality.

An important factor in the success of trait psychology was psychometric advancements using factor analysis and the relatively easy to obtain measurement using self-rating questionnaires. The principles of the classical test theory (Lord, Novick, & Birnbaum, 1968) lend themselves very well to study personality traits using questionnaires and Likert scales. For a comprehensive overview of the history of personality traits and personality in general, we refer the reader to the excellent summary by McAdams (McAdams, 1997).

The Early Years of Implicit Motives Research

The history of implicit motives as a personality construct took a different route. A major reason were difficulties and challenges in measurement and the absence of an appropriate psychometric framework. In the following section, we describe how research on implicit motives first popularized and then how psychometric criticism lead to decreasing interest. McClelland and Atkinson started a research program in the 1940ies to study human motivation. Their first interest was the *need for hunger*. McClelland had some initial doubts that people would respond honestly when researchers would ask them to rate their motives. As a young man, he observed differences between the values people expressed on Sunday in church and the behavior they showed during the week (McClelland, 1984). He was searching for a way to measure the implicit values that guided behavior. They decided to use the TAT and their initial results were promising (McClelland & Atkinson, 1948).

With the use of the TAT, McClelland also adopted terminology from Murrays personology. The construct—today studied under the term "implicit motives"—was termed *need for* hunger or *n* hunger and the term *projective* was adopted to describe the response process. After studying the need for hunger, McClelland and Atkinson quickly shifted their attention to studying the *need for achievement*, because they believed it to be more relevant for predicting job outcomes. The initial work on the achievement motive was summarized in the book *The Achievement Motivation* (McClelland, Atkinson, Clark, & Lowell, 1953). The book sparked interest in research on implicit motives and led to a rapid proliferation of research. Based on the TAT version from McClelland, researchers developed and validated different measures to study motivational needs such as the *need for affiliation, sex, autonomy,* or *power*. Today, researchers typically study the "big three" implicit motives: affiliation, achievement, and power (e.g. Kehr, 2004b). The effort to understand picture story coding methods in the initial boom was documented in *Motives in Fantasy, Action, and Society* (Atkinson, 1958). During this time, researchers showed in various studies the utility and predictive power of implicit motive measures.

Psychometric Criticism and Responses

In the next 30 years, research on implicit motives continued to diversify, but also stagnated and suffered from harsh criticism on psychometric properties of the TAT. The field of psychology started to divide into researchers who thought that implicit motives were a valuable source that adds unique information to predict human behavior and those who were convinced that it is not justified to continue using them. Critical voices towards implicit motives focused mainly on the low reliability of picture story coding measures. Klinger (1966) reported low test-retest reliabilities and concluded that the temporal stability of implicit motive measures is low. Therefore, he concluded that the TAT is not useful to measure stable motivational dispositions. Entwisle (1972) found that internal consistencies of implicit motive measures were too low—typically in the range from .30 to .40—which seriously limits the validity of implicit motive measures. Fineman (1977) reiterated the criticisms of Klinger and Entwisle and added some additional studies in a review. Although the criticisms also had some shortcomings—such as the omission of studies not in line with their message (Smith, 1992)—they received a lot of attention and led to a decline in research on implicit motives.

Proponents of implicit motive measures responded in various ways. Mitchell (1961) suggested that some studies included by Entwisle did not show sufficient agreement between raters and therefore did not hold the high standard of other studies. Other responses included omissions of important findings such as Heckhausen (1963) who found a retest reliability of .53 and studies that used different ways to control for story length (Reuman, Alwin, & Veroff, 1984).

A key response was the development of the *Dynamics of Action* theory by Atkinson and Birch (Atkinson & Birch, 1970; Atkinson, Bongort, & Price, 1977) challenging the assumption that low internal consistency limits validity for the TAT. They argued that classical test theory assumes that a response is not affected by previous responses to another item, which may not be true for implicit motives. The dynamics of actions theory proposes that a need is dynamic and need strength may change over time. For instance, a need can be satisfied and temporally lowered (e.g. when someone ate a full meal, their need for hunger is temporarily satisfied). Applied to the measurement context of the TAT, the

dynamic of actions theory suggests that writing a story with achievement content may temporarily lower the motive so that respondents are less likely to write an achievement story in the following picture. Another aspect of the theory is that different motives (such as achievement and affiliation motives) compete for enactment in a given situation and the strongest motive determines the following behavior. Atkinson and colleagues (1977) could show in simulation studies that these principles may account for the low internal consistencies reported in the criticisms of the TAT. They argued that over the test, the TAT sum score averages out dynamic changes of motives. Therefore, it is not surprising that the TAT predicts behavior even if internal consistency is low. An excellent review of the criticism and the responses is written by Smith (1992).

Although implicit motive researchers responded to the criticism, it took them a long while to do so (Smith, 1992). Meanwhile, researchers that were interested in the construct of motives started to use self-report measures based on questionnaires as an alternative to the TAT. However, that approach led to confusion in the field, because self-report measures did not correlate with picture story coding (Köllner & Schultheiss, 2014; McClelland, Koestner, & Weinberger, 1989; Spangler, 1992). If TAT and self-report would measure the same construct, they should correlate. McClelland, consistent with his beliefs that people may not be able to accurately report on their personality, initially discounted self-report measures (Schultheiss & Brunstein, 2010b). However, research continuously showed that self-reported motives—although they do not correlate with TAT measures—also predicted various behaviors. In what is now a hallmark paper of the field, McClelland finally suggested that self-report and picture-story based measures describe different constructs (McClelland, Koestner, & Weinberger, 1989).

He proposed a two-system model to differentiate between *self-attributed needs* (for self-ratings) and *implicit motives* (for picture-story based measures). Self-attributed—or explicit motives—refer to motives on a conscious level and implicit motives refer to motives on an nonconscious level. He proposed that both constructs predict different types of behavior. Implicit motives predict what he called *operant* behavior referring to behavior in unstructured situations (note that the term "operant" is used differently in learning

psychology). Explicit motives predict *respondent* behavior referring to behavior in response to specific demands and expectations.

McClelland replaced the term "need for" with *implicit motives* and as a contrast *explicit motives* for self-attributed motives, in analogy to implicit memory research that distinguished between implicit and explicit memory (Schacter, 1987). The paper also introduced a new term for the measurement procedure for implicit motives, the Picture Story Exercise (PSE). Although the way implicit motives were measured was initially strongly influenced by the measurement procedure of the TAT, many of the procedures have changed over time.

The conceptual clarification of this work led to an increased interest in motivational constructs and reinvigorated research on implicit motives. Subsequent meta-analyses supported the two-system theory (Köllner & Schultheiss, 2014; Spangler, 1992). A comprehensive overview of the history of implicit motives, that goes into more detail, is provided in the introduction of the book *Implicit Motives*, edited by Schultheiss and Brunstein (2010b).

Implicit Motives in Organizational Behavior

Personality traits—mainly (but not only) in the form of the Five-Factor Model—are now a pillar of OB research. Many theoretical papers, empirical studies, and several important meta-analyses have provided evidence that individual differences, conceptualized as personality traits, play an important role in explaining and predicting organizational behavior such as performance (Barrick & Mount, 1991), counterproductive work behavior (Hershcovis et al., 2007), organizational citizenship behavior (Chiaburu, Oh, Berry, Li, & Gardner, 2011), and many more. Effects of personality on OB are studied in long term perspectives (Lievens, Ones, & Dilchert, 2009; Woods, Lievens, De Fruyt, & Wille, 2013), on a daily level (Ilies, Scott, & Judge, 2006), or in teams (Bell, 2007). The basis for this outstanding progress are easy to use, reliable, and validated measures for personality traits like the NEO-PI (Costa & McCrae, 1992; McCrae, Costa, & Martin, 2005), BFI (Soto & John, 2017), or the HEXACO (Lee & Ashton, 2018) that provide researchers with tools to study personality traits.

Studying the role of implicit motives in predicting organizational behavior is, however, much more difficult. Picture story coding-based procedures are challenging to administer, scoring may not always be objective, and self-report measures are absent. Many researchers criticize implicit motive measures for low reliabilities and therefore discount them (Entwisle, 1972; Fineman, 1977; Lilienfeld, Wood, & Garb, 2000). Nevertheless, researchers studied the role of implicit motives in OB (Kanfer, 1992). For instance, McClelland was interested in the effect of achievement motivation on the economic development of societies. He focused on societal and organizational applications of the achievement motive and found several links that periods of economic growth were preceded by periods of increased collective achievement motivation (McClelland, 1961). He could also show that a two-week training program in an Indian city aimed at improving achievement motivation led participants to work harder, invest more, and create more jobs than in an untrained control group from a different city in the same region (McClelland & Winter, 1969; McClelland, 1965). The success of the training also showed two years later, when the city that had received the training had obtained one-third more jobs compared to the other city.

After the initial focus on a macro level, researchers started to study implicit motive also on an organizational level. One study showed that an achievement motive training also led small businesses in the United States to successfully expand and improve (Miron & McClelland, 1979). A prominent focus in the following years was on leadership. McClelland and Boyatzis (1982) predicted long-term management success over eight and 16 years using a configuration of motives—the Leader Motive Pattern—with moderate to high power motivation, low affiliation motivation and high activity inhibition (a self-control measure, conceptualized as the frequency of the word *not* in a story). This work was extended by Winter (1991) with the concept of *responsibility* that interacted with high power motive to predict leader success. Researchers also found that successful female and male leaders differ in their power motivation (Chusmir & Parker, 1984; Jacobs & McClelland, 1994). While female leaders were more likely to use resourceful power themes (protecting, supporting, and helping others), male leaders were more likely to use reactive power themes (assertive and aggressive actions). As it is not always possible to access

leaders, especially highly visible political and corporate leaders, Winter developed a content coding system to assess implicit motives from a distance using verbal material such as speeches (Winter, 1977, 1994). Measuring motives from a distance, researchers studied leader effectiveness in US and South African presidents and predicted important political dispositions and styles (Donley & Winter, 1970; Winter, 1980).

Researchers have also studied the role of implicit motives in entrepreneurship. Business owners had a higher achievement motivation than managers and the achievement motive was positively related to business success (McClelland, 1961). Roberts (1989) found that another less frequently studied motive, the autonomy motive, was high in a sample of technical entrepreneurs. Researchers have also studied relationships for performance measures and found a link between the achievement motive and learning as well as speed of performance (Lowell, 1952), and also that executive performance is predicted by the achievement motive (Cummin, 1967). One of the few meta-analyses for implicit motives found that the achievement motive predicts career success and that the effect for implicit motives is on average higher when compared to explicit motives (Spangler, 1992).

More recent research has found that implicit motives interact with personality traits to predict performance (Lang, Zettler, Ewen, & Hülsheger, 2012), a person-job fit between implicit motives and the possibility to satisfy them at work predicts burnout and physical symptoms (Brandstätter, Job, & Schulze, 2016), the discrepancy between implicit and explicit motives is associated with volitional depletion among managers (Kehr, 2004a), and that the affiliation motive predicts income over several years (Apers, Lang, & Derous, 2018).

Despite these recent efforts, the number of studies investigating implicit motives in the context of OB is limited. Kanfer, Frese, and Johnson (2017) recently reviewed 100 years of research on work motivation in the article "*Motivation related to work: A century of progress*" in the centennial special issue of the Journal of Applied Psychology. They acknowledged the early work on implicit motives in OB and emphasized that most studies focused on the achievement motive, but also concluded that "research on motives waned, however, during the 1970s owing to criticisms about personality research in general

(Mischel, 1968) and in I/O psychology specifically (Mitchell, 1979).” (Kanfer et al., 2017; p. 341). Criticism about personality research, in general, was an impactful reason for the decline in implicit motive research in OB. Personality trait research, however, recovered well after the success of the Five-Factor Model, while research on implicit motives waned. We believe that a crucial reason for implicit motive research to lag behind personality trait research in OB lies in the challenging measurement and absence of appropriate psychometric models. In this dissertation, we aim to close this gap with fundamental research on how to measure implicit motives better.

Research Objectives and Overview of the Chapters

In this dissertation, we built on recent psychometric developments for modeling implicit motives (Lang, 2014) to advance the measurement of implicit motives and study test criteria. Researchers typically evaluate the quality of a measure using test criteria. The three major test criteria are objectivity, reliability, and validity. A test is objective when different persons evaluating the test come to the same result; a test is reliable when the test score is accurate, reproducible, and consistent. A test is valid when it measures what it is supposed to measure. An additional test standard that is especially relevant in OB is test fairness. A test is fair when respondents with the same underlying trait who differ in gender, age, or cultural background score the same result. Another relevant criterion for researchers as well as practitioners is test economy. An economical test takes a reasonable amount of time and resources to administer and evaluate it. Below, we provide an overview of the four empirical chapters studying the overall aim to improve the measurement of implicit motives for research in OB. Based on these test criteria, we aim to improve the measurement of implicit motives for research in OB. In the following section, we provide an overview of the four empirical chapters which can be read as independent papers, each contributing to the overall aim of improving the measurement of implicit motives in an applied context.

Chapter 2: A Psychometric Analysis of the Operant Motive Test

Atkinson and Birch (1970) suggested that (1) implicit motives compete for enactment and that (2) motive strength may dynamically change during assessment due to a consumption effect and temporarily satisfied motives. Lang (2014) developed the dynamic

Thurstonian item response model (DTM) that models both response processes. He demonstrated that IRT reliability—accounting for the response processes—of the PSE scores in an archival dataset is much higher than internal consistency. The main goal of Chapter 2 is to apply the DTM to the Operant Motive Test (OMT; Kuhl & Scheffer, 2001), a modern implicit motive measure that asks for shorter responses and uses more pictures as the traditional PSE. Using the DTM, we study IRT reliability for the OMT and evaluate whether the findings in Lang (2014) also apply to modern implicit motive data. In addition, we provide an R tutorial explaining how to model the DTM for the OMT using the lme4 package and how to program a simulation study that estimates IRT reliabilities using a parameter recovery approach.

Chapter 3: Improving the Assessment of Implicit Motives Using IRT: Cultural Differences and Differential Item Functioning

In Chapter 3, we extended the DTM developed for the OMT in Chapter 2 to evaluate differential item functioning (DIF). DIF occurs when two respondents from different groups with the same score in the underlying trait have a different probability to respond to the item in question. For instance, assume that an OMT picture shows two people, one working on a task (achievement) and the other one supervising the first person (power). If some attributes in the picture indicate one person as a woman and the other as a man, it is likely that these attributes influence which story is written based on group membership (gender) and not the strength of an implicit motive. An item that shows DIF is potentially threatening research conclusions and personnel selections. If an item in an implicit motive test shows DIF for a particular group, motive comparisons for this group may be biased. In this chapter, we extend the DTM to test for DIF based on a procedure introduced by Thissen, Steinberg, and Gerrard (1986). In addition, we suggest an expansion of the model to account for DIF (when it is present) to get meaningful group comparisons. The suggested DIF analysis allows to identify and remove items that may lead to biased test results. For selection contexts, it is important to not systematically favor one group over another. We apply our DIF analysis approach to a dataset with respondents from three different cultural backgrounds.

Chapter 4: Predicting Counterproductive Work Behavior: Do Implicit Motives Have Incremental Validity Beyond Explicit Traits?

In Chapter 4, we study whether implicit motives contribute to the prediction of counterproductive work behavior above personality traits. Past research, especially in the early years of implicit motive research, has suffered from low sample sizes and has mainly focused on the implicit achievement motive to predict organizational behavior. However, research has shown that other aspects (e.g. social aspects) also play an important role in OB (e.g .Wrzesniewski, Dutton, & Debebe, 2003). In Chapter 4, we extend previous research by studying implicit motives as predictors of counterproductive work behavior. To demonstrate that implicit motives predict behavior in addition to personality traits, we investigate incremental validity in two studies.

Chapter 5: Can People Recognize their Implicit Thoughts? The Motive Self-Categorization Test

The dependence on expert coders in implicit motive measures leads to several restrictions. First, expert coders need to receive extensive coding training before they can code stories. Second, coding the stories itself is tedious and time-consuming. Third, because different raters may come to different conclusions about a story, motive scores may not always be objective in the sense that scores can depend on raters. Researchers have long argued that one cannot ask people directly for their implicit motives because they are, by definition, nonconscious and not accessible through introspection (e.g. McClelland et al., 1989). The common justification for the use of expert coders is that motivational content is not immediately apparent to the person producing the story. The main argument is that respondents actually do not have a direct route to self-knowledge of their implicit motives. This reasoning is also in line with evidence suggesting that implicit motive measures do not correlate with questionnaires directly asking respondents for their motives (McClelland et al., 1989; Spangler, 1992). In Chapter 5, we scrutinized the assumption that people cannot evaluate their implicit motives and developed a new measure of implicit motives that uses an innovative approach letting respondents self-categorize their motives. Instead of expert coders, respondents categorize their stories into motivational themes choosing contextualized statements that represent typical answers to this picture. The statements

were developed based on previous responses to the pictures and are directly linked to an implicit motive. We study the reliability of the new measure based on the DTM developed in Chapter 2, the convergence with expert coding using latent Multitrait-Multimethod models (Campbell & Fiske, 1959; Widaman, 1985), and overlap with self-ratings of personality.

Chapter 6: General Discussion

Finally, in Chapter 6 we discuss general conclusions, theoretical and practical implications resulting from our four empirical chapters and in addition, some limitations and recommendations for future research.

References

Allport, G. W. (1931). What is a trait of personality? *Journal of Abnormal and Social Psychology, 25,* 368–372.

Apers, C., Lang, J. W. B., & Derous, E. (2018). Who earns more? Explicit traits, implicit motives and income growth trajectories. *Journal of Vocational Behavior*. https://doi.org/10.1016/J.JVB.2018.12.004

Atkinson, J. W. (1958). *Motives in fantasy, action and society: A method of assessment and study*. Oxford, UK: Van Nostrand.

Atkinson, J. W., & Birch, D. (1970). *Dynamics of Action*. New York: Wiley.

Atkinson, J. W., Bongort, K., & Price, L. H. (1977). Explorations using computer simulation to comprehend thematic apperceptive measurement of motivation. *Motivation and Emotion*, *1*, 1–27. https://doi.org/10.1007/BF00997578

Barrick, M. R., & Mount, M. K. (1991). The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. *Personnel Pychology*, *44*, 1–26. http://dx.doi.org/10.1111/j.1744-6570.1991.tb00688.x

Bell, S. T. (2007). Deep-Level Composition Variables as Predictors of Team Performance: A Meta-Analysis. *Journal of Applied Psychology*, *92*, 595–615. https://doi.org/10.1037/0021-9010.92.3.595

Brandstätter, V., Job, V., & Schulze, B. (2016). Motivational Incongruence and Well-Being at the Workplace: Person-Job Fit, Job Burnout, and Physical Symptoms. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01153

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56,* 81–105. https://doi.org/10.1037/h0046016

Cattell, R. B. (1946). *Description and Measurement of Personality*. New York: World Books.

Cattell, R. B. (1950). *Personality: A systematic theoretical and factual study, 1st ed.* New York: McGraw-Hill. https://doi.org/10.1037/10773-000

Cattell, R. B. (1952). *Factor Analysis*. New York: Harper.

Chiaburu, D. S., Oh, I.-S., Berry, C. M., Li, N., & Gardner, R. G. (2011). The five-factor model of personality traits and organizational citizenship behaviors: A meta-analysis. *Journal of Applied Psychology*, *96*(6), 1140–1166. https://doi.org/10.1037/a0024004

Chusmir, L. H., & Parker, B. (1984). Dimensions of need for power: Personalized vs. socialized power in female and male managers. *Sex Roles*, *11*, 759–769. https://doi.org/10.1007/BF00287808

Costa, P. T. ., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PIR): Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Cummin, P. C. (1967). TAT correlates of executive performance. *Journal of Applied Psychology*, *51*, 78–81. https://doi.org/10.1037/h0024246

Donley, R. E., & Winter, D. G. (1970). Measuring the motives of public officials at a distance: An exploratory study of american presidents. *Behavioral Science*, *15*, 227–236. https://doi.org/10.1002/bs.3830150304

Entwisle, D. R. (1972). To Dispel Fantasies About Fantasy-Based Measures of Achievement Motivation. *Psychological Bulletin*, *77*, 377–391. https://doi.org/10.1037/h0020021

Eysenck, H. J. (1952). *The Scientific Study of Personality*. Routledge & K. Paul.

Fineman, S. (1977). The achievement motive construct and its measurement: Where are we now? *British Journal of Psychology*, *68*, 1–22. https://doi.org/10.1111/j.2044-8295.1977.tb01554.x

Heckhausen, H. (1963). *Hoffnung und Furcht in der Leistungsmotivation [Hope and fear in achievement motivation]*. Meisenheim am Glan, Germany: Verlag Anton Hain.

Hershcovis, M. S., Turner, N., Barling, J., Arnold, K. A., Dupré, K. E., Inness, M., … Sivanathan, N. (2007). Predicting workplace aggression: A meta-analysis. *Journal of Applied Psychology*, *92*, 228–238. https://doi.org/10.1037/0021-9010.92.1.228

Ilies, R., Scott, B. A., & Judge, T. A. (2006). The interactive effects of personal traits and experienced states on intraindividual patterns of citizenship behavior. *Academy of Management Journal*, *49*, 561–575. https://doi.org/10.5465/amj.2006.21794672

Jacobs, R. L., & McClelland, D. C. (1994). Moving up the corporate ladder: A longitudinal study of the leadership motive pattern and managerial success in women

and men. *Consulting Psychology Journal*, *46*, 32–41. https://doi.org/10.1037/1061-4087.46.1.32

Kanfer, R. (1992). Work motivation: New directions in theory and research. Work Science Center View project. In C. L. Cooper & I. . T. Robertson (Eds.), *International Review of Industrial and Organizational* (pp. 1–53). London: John Wiley & Sons, Ltd.

Kanfer, R., Frese, M., & Johnson, R. E. (2017). Motivation Related to Work: A Century of Progress. *Journal of Applied Psychology*, *102*, 338–355. https://doi.org/10.1037/apl0000133

Kehr, H. M. (2004a). Implicit/explicit motive discrepancies and volitional depletion among managers. *Personality and Social Psychology Bulletin*, *30*, 315–327. https://doi.org/10.1177/0146167203256967

Kehr, H. M. (2004b). Integrating implicit motives, explicit motives, and perceived abilities: The compensatory model of work motivation and volition. *Academy of Management Review*, *29*, 479–499. https://doi.org/10.5465/AMR.2004.13670963

Klinger, E. (1966). Fantasy need achievement as a motivational construct. *Psychological Bulletin*, *66*, 291–308. https://doi.org/10.1037/h0023820

Köllner, M. G., & Schultheiss, O. C. (2014). Meta-analytic evidence of low convergence between implicit and explicit measures of the needs for achievement, affiliation, and power. *Frontiers in Psychology*, *5*, 826. https://doi.org/10.3389/fpsyg.2014.00826

Kuhl, J. (1978). Situations-, reaktions- und personbezogene Konsistenz des Leistungsmotivs bei der Messung mittels des Heckhausen TAT [Situation-, reaction-, and person-related consistency of the achievement motive measured using Heckhausen's TAT]. *Archiv Für Psychologie*, *130*, 37–52.

Kuhl, J., & Scheffer, D. (2001). Der operante multi-motiv-test (OMT): Manual [The operant multi-motive-test (OMT): Manual]. Osnabrück, Germany: University of Osnabrück.

Lang, J. W. B. (2014). A dynamic Thurstonian item response theory of motive expression in the picture story exercise: Solving the internal consistency paradox of the PSE. *Psychological Review*, *121*, 481–500. https://doi.org/10.1037/a0037011

Lang, J. W. B., Zettler, I., Ewen, C., & Hülsheger, U. R. (2012). Implicit motives, explicit

traits, and task and contextual performance at work. *Journal of Applied Psychology*, *97*, 1201–1217. https://doi.org/10.1037/a0029556

Lee, K., & Ashton, M. C. (2018). Psychometric properties of the HEXACO-100. *Assessment*, *25*, 543–556. https://doi.org/10.1177/1073191116659134

Lievens, F., Ones, D. S., & Dilchert, S. (2009). Personality Scale Validities Increase Throughout Medical School. *Journal of Applied Psychology*, *94*, 1514–1535. https://doi.org/10.1037/a0016137

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The Scientific Status of Projective Techniques. *Psychological Science in the Public Interest*, *1*, 27–66. https://doi.org/10.1111/1529-1006.002

Lord, F. M., Novick, M. R., & Birnbaum, A. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Lowell, E. L. (1952). The Effect of Need for Achievement on Learning and Speed of Performance. *The Journal of Psychology*, *33*, 31–40. https://doi.org/10.1080/00223980.1952.9712815

McAdams, D. P. (1997). A Conceptual History of Personality Psychology. *Handbook of Personality Psychology*, 3–39. https://doi.org/10.1016/B978-012134645-4/50002-0

McClelland, D. C. (1961). *The achieving society.* Princeton N.J.: Van Nostrand.

McClelland, D. C. (1965). Toward a theory of motive acquisition. *American Psychologist*, (5), 321–333. https://doi.org/10.1037/h0022225

McClelland, D. C. (1984). *Motives, Personality, and Society: Selected Papers.* New York, NY, NY: Praeger Publishers.

McClelland, D. C., & Atkinson, J. W. (1948). The Projective Expression of Needs: I. the Effect of Different Intensities of the Hunger Drive on Perception. *The Journal of Psychology*, *25*, 205–222. https://doi.org/10.1080/00223980.1948.9917371

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). The measuring instrument. In *The Achievement Motive* (pp. 185–217). East Norwalk, CT, US: Appleton-Century-Crofts.

McClelland, D. C., & Boyatzis, R. E. (1982). Leadership motive pattern and long-term success in management. *Journal of Applied Psychology*, *67*, 737–743.

https://doi.org/10.1037/0021-9010.67.6.737

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*, 690–702. https://doi.org/10.1037/0033-295X.96.4.690

McClelland, D. C., & Winter, D. G. (1969). *Motivating Economic Achievement*. New York: Free Press.

McCrae, R. R., Costa, Jr., P. T., & Martin, T. A. (2005). The NEO–PI–3: A More Readable Revised NEO Personality Inventory. *Journal of Personality Assessment*, *84*, 261–270. https://doi.org/10.1207/s15327752jpa8403_05

McCrae, R. R., & John, O. P. (1992). An introduction to the Five-Factor Model and its applications. *Journal of Personality*, *60*, 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

Miron, D., & McClelland, D. C. (1979). The Impact of Achievement Motivation Training on Small Businesses. *California Management Review*, *21*, 13–28. https://doi.org/10.2307/41164830

Mitchell, J. V. (1961). An analysis of the factorial dimensions of the achievement motivation construct. *Journal of Educational Psychology*, *52*, 179–187. https://doi.org/10.1037/h0042486

Morgan, C. D., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology & Psychiatry*, *34*, 289–306. https://doi.org/10.1001/archneurpsyc.1935.02250200049005

Murray, H. A. (1938). *Explorations in Personality; A clinical and Experimental Study of Fifty Men of College Age.* New York: Oxfort University Press. https://doi.org/10.2307/2084329

Reuman, D. A., Alwin, D. F., & Veroff, J. (1984). Assessing the validity of the achievement motive in the presence of random measurement error. *Journal of Personality and Social Psychology*, *47*, 1347–1362. https://doi.org/10.1037/0022-3514.47.6.1347

Roberts, E. B. (1989). The personality and motivations of technological entrepreneurs. *Journal of Engineering and Technology Management*, *6*, 5–23.

https://doi.org/10.1016/0923-4748(89)90012-X

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 501–518. https://doi.org/10.1037/0278-7393.13.3.501

Schultheiss, O. C., & Brunstein, J. (2010a). Implicit Motives. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 603–633). New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195335156.001.0001

Schultheiss, O. C., & Brunstein, J. C. (2010b). Introduction. In *Implicit Motives2*.

Smith, C. P. (1992). *Motivation and personality: Handbook of thematic content analysis*. New York: Cambridge University Press.

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*, 117–143. https://doi.org/10.1037/pspp0000096

Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin*, *112*, 140–154. https://doi.org/10.1037/0033-2909.112.1.140

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychologysychology*, *15*, 72–101. https://doi.org/10.2307/1412159

Thissen, D., Steinberg, L., & Gerrard, M. (1986). Beyond group-mean differences: The concept of item bias. *Psychological Bulletin*, *99*, 118. https://doi.org/10.1037/a0027127

Thurstone, L. L. (1931). Multiple factor analysis. *Psychological Review*, *38*, 406–427. https://doi.org/10.1037/h0069792

Tuerlinckx, F., De Boeck, P., & Lens, W. (2002). Measuring needs with the Thematic Apperception Test: A psychometric study. *Journal of Personality and Social Psychology*, *82*, 448–461. https://doi.org/10.1037//0022-3514.82.3.448

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-

multimethod data. *Applied Psychological Measurement*, *9*, 1–26. https://doi.org/10.1177/014662168500900101

Winter, D. (1991). A motivational model of leadership: Predicting long-term management success from TAT measures of power motivation and responsibility. *The Leadership Quarterly*.

Winter, D. G. (1977). *Manual for Scoring Motive Imagery in Running Text*.

Winter, D. G. (1980). An exploratory study of the motives of southern african political leaders measured at a distance. *Political Psychology*, *2*, 75. https://doi.org/10.2307/3790818

Winter, D. G. (1994). *Manual for Scoring Motive Imagery in Running Text*. Ann Arbor: University of Michigan.

Winter, D. G., John, O. P., Stewart, A. J., Klohnen, E. C., & Duncan, L. E. (1998). Traits and motives: Toward an integration of two traditions in personality research. *Psychological Review*, *105*, 230–250. https://doi.org/10.1037/0033-295X.105.2.230

Woods, S. A., Lievens, F., De Fruyt, F., & Wille, B. (2013). Personality across working life: The longitudinal and reciprocal influences of personality on work. *Journal of Organizational Behavior*, *34*, S7–S25. https://doi.org/10.1002/job.1863

Wrzesniewski, A., Dutton, J. E., & Debebe, G. (2003). Interpersonal sensemaking and the meaning of work. *Research in Organizational Behavior*, *25*, 93–135. https://doi.org/10.1016/S0191-3085(03)25003-6

CHAPTER 2

A PSYCHOMETRIC ANALYSIS OF THE OPERANT MOTIVES TEST[1]

The Operant Motive Test (OMT) is a picture-based procedure that asks respondents to generate imaginative verbal behavior that is later coded for the presence of affiliation, power, and achievement-related motive content by trained coders. The OMT uses a larger number of pictures and asks respondents to provide more brief answers than earlier and more traditional picture-based implicit motive measures and has therefore become a frequently used measurement instrument in both research and practice. This article focuses on the psychometric response mechanism in the OMT and builds on recent advancements in the psychometric modeling of the response process in implicit motive measures using Thurstonian item response theory. The contribution of the article is twofold. First, the article builds on a recently developed dynamic Thurstonian model for more traditional implicit motive measures (Lang, 2014) and reports the first analysis of which we are aware that applies this model to OMT data (N = 633) and studies dynamic motive activation in the OMT. Results of this analysis yielded evidence for dynamic motive activation in the OMT and showed that simulated IRT reliabilities based on the dynamic model were .52, .62, and .73 for the affiliation, achievement, and power motive in the OMT, respectively. The second contribution of this article is a tutorial and R code that allows researchers to directly apply the dynamic Thurstonian IRT model to their data. The future use of the OMT in research and potential ways to improve the OMT are discussed.

---

[1] This chapter is based on: Runge, J. M., Lang, J. W. B., Engeser, S., Schüler, J., den Hartog, S. C., & Zettler, I. (2016). Modeling motive activation in the Operant Motive Test: A psychometric analysis using dynamic Thurstonian item response theory. *Motivation Science*, 2, 268-286. http://dx.doi.org/10.1037/mot0000041

Motivational researchers have long been interested in implicit motives that are typically described as stable affective preferences for certain goal states (McClelland, Koestner, & Weinberger, 1989; Schultheiss, Liening, & Schad, 2008; Winter, John, Stewart, Klohnen, & Duncan, 1998). Implicit motives predict performance in the laboratory (e.g., McAdams, Jackson, & Kirshnit, 1984), psychological well-being (e.g., Baumann, Kaschel, & Kuhl, 2005), and are useful in the prediction of outcome criteria like job performance (e.g., Lang, Zettler, Ewen, & Hülsheger, 2012), entrepreneurial success (e.g., Spangler, 1992), career success (McClelland & Boyatzis, 1982) or satisfaction in relationships (e.g., Winter et al., 1998). Implicit motives are also related to hormonal and neuronal indicators (Hall, Stanton, & Schultheiss, 2010; Kuhl & Kazén, 2008).

Most implicit motive measures are picture-based procedures (McClelland et al., 1989; Winter et al., 1998). These procedures consist of a series of pictures showing persons in different situations and ask respondents to describe what they see in the picture, what happened before, and what will happen next. The stories that respondents generate are then evaluated for their motive content by trained coders using standardized and detailed scoring manuals. Coders typically show good inter-rater agreement (Köllner & Schultheiss, 2014; Schüler, Brandstätter, Wegner, & Baumann, 2015).

The most common picture-based measure to assess implicit motives is the Picture Story Exercise (PSE). The PSE was developed in the 1950s (McClelland, Atkinson, Clark, & Lowell, 1953), consists of a total of six pictures, and asks respondents to write elaborate stories in response to each of the pictures. A more recently developed implicit motive measure is the Operant Motive Test (OMT). The OMT uses the same picture-based measurement approach like more traditional PSE-measures but uses a larger number (15) of simplified pictures and asks respondents to write only brief responses to these pictures. These characteristics may have contributed to the fact that the OMT is currently frequently used in implicit motive research (e.g., Baumann et al., 2005; Baumann & Scheffer, 2010; Hofer, Busch, Chasiotis, Kärtner, & Campos, 2008; Kazén & Kuhl, 2005; Lang et al., 2012; Schüler, Job, Fröhlich, & Brandstätter, 2008). The larger number of pictures in the OMT makes this measure also especially interesting for psychometric research.

One major problem of picture-based procedures like the PSE and the OMT has long been that the psychometric process underlying the responses in implicit motive measures is complex, and, in turn, has been difficult to capture (e.g., Entwisle, 1972). However, researchers have recently built on earlier theoretical work suggesting that people choose among different motivational alternatives that are not independent (Atkinson & Birch, 1970), and have applied Thurstonian item response theory based models (Thurstonian IRT) to model this process (Lang et al., 2012). More recently, research (Lang, 2014) has developed a dynamic Thurstonian IRT model for the more traditional PSE. This dynamic Thurstonian model additionally includes a dynamic IRT process (Verguts & De Boeck, 2000; Verhelst & Glas, 1993) and is inspired by the idea that motivational tendencies are dynamic and that their strength declines when they are acted out (Atkinson & Birch, 1970; Tuerlinckx, De Boeck, & Lens, 2002).

The first major contribution of this paper is to apply the dynamic Thurstonian IRT model (Lang, 2014) developed for the PSE also to the OMT. The OMT is now a frequently used measure and we accordingly believe that motivational research would benefit from additional psychometric insights into this measure. A second contribution of this paper is an accessible description and demonstration on how researchers can apply the dynamic Thurstonian IRT model to their OMT data. This demonstration should help motivational researchers in using the dynamic Thurstonian IRT model in their own research.

Implicit Motive Measures

Implicit motives are described as people's wishes and desires (Winter et al., 1998) or "the disposition to be concerned with and to strive for a certain class of incentives or goals" (Emmons, 1989, p. 32). The dispositions are implicit because they are usually not accessible through introspection (McClelland et al., 1989; Winter et al., 1998), and thus need to be inferred indirectly from coding imaginative verbal material. Implicit motive measures typically ask respondents to write a fantasy story based on ambiguous picture stimuli. These stories are then evaluated for motive content. Most research focuses on three basic motive contents: *need for affiliation*, the desire to build, maintain, and deepen social relationships; *need for achievement*, the desire to improve one's performance; *need for power*, the desire to influence other people (Schultheiss, 2008). For sake of simplicity, we refer to the three basic

implicit motives as affiliation, achievement and power and omit the term “need for” in the remainder of this chapter.

Implicit motive measures are typically not or only weakly correlated with explicit questionnaire measures of motives (Köllner & Schultheiss, 2014; Spangler, 1992). Several authors have suggested that this finding is not surprising because implicit motive measures and questionnaires tap different types of individual differences (McClelland et al., 1989; Winter et al., 1998) that predict different classes of behavior.

Morgan and Murray's (1935) Thematic Apperception Test (TAT) was the first measure to use picture stimuli in order to measure subconscious aspects of personality and influenced future measures of implicit motives. The first adaption developed to measure implicit achievement, affiliation, and power motivation is McClelland’s Picture-Story Exercise (McClelland, 1987; McClelland et al., 1953). A more recent implicit motive measure is the OMT (Kuhl & Scheffer, 2002).

The Operant Motive Test

The OMT uses a larger number of pictures (typically 15 pictures instead of six in the PSE) that depict scenes with one or several persons (Kuhl & Scheffer, 2002; Scheffer, 2005). Participants are instructed to choose one character shown in the picture as their protagonist. In contrast to other implicit motive measures that ask for full and long explanations like the PSE, the OMT asks for brief answers to the questions: (1) “What is important for the person in this situation and what is the person doing?” (2) “How does the person feel?” (3) “Why does the person feel this way?” (4) “How does the story end?”. The OMT developers chose this procedure in the hope to get a more spontaneous answer from participants, which is supposedly linked more directly to implicit motives (Baumann, Kaschel, & Kuhl, 2010). Another unique characteristic of the OMT is that the pictures are drawings instead of photographs so that characters are depicted in a neutral manner in order to facilitate respondent’s identification with the chosen character. Figure 1 shows a sample picture of the OMT.

The respondents’ answers are categorized into one of the following motivational themes: affiliation, achievement and power. Only one motive response is coded for each picture. Stories with a lack of these motivational themes are categorized as zero. Additionally,

the OMT manual differentiates between motive implementation strategies via three approach, one avoidance, and one in-between category within each motive (Kuhl & Scheffer, 1999, 2002). Researchers frequently combine the approach categories in order to get an overall approach factor (Baumann et al., 2005; Kazén & Kuhl, 2011; Lang et al., 2012; Schüler et al., 2008; Schüler, Job, Fröhlich, & Brandstätter, 2009). The approach codings for each picture are then added up to a *sum score* for each motive. The avoidance categories are usually analyzed separately.

Reliability of Implicit Motive Measures

The most frequent method to estimate reliability is internal consistency (Cronbach, 1951), which several researchers (Entwisle, 1972; Fineman, 1977; Lilienfeld, Wood, & Garb, 2000) found to be low—in the .20 to .40s—for implicit motive measures. However, internal consistency only equals the reliability when the underlying disposition is unidimensional. Internal consistency cannot deal with a scenario in which motive reactions depend on the activation of other motives and in which individuals react to motivational stimuli.

Modeling Motive Activation in the OMT Using IRT

Thurstonian IRT Models

Researchers have long recognized that respondents typically write stories that focus on one motive-related theme (e.g., either achievement or power). Motives may thus compete for activation in a particular story and respondents effectively choose between different motives when they respond to implicit motive measures (Atkinson & Birch, 1970; Atkinson, Bongort, & Price, 1977). An established tool for modeling complex choice behavior are Thurstonian models (e.g., Böckenholt, 2006; Brown & Maydeu-Olivares, 2011) and researchers have recognized that Thurstonian models can also be used to model choice behavior in implicit motive measures (Lang, 2014; Lang et al., 2012). Thurstonian models consist of pairwise comparisons and place these comparisons on a utility scale (Thurstone, 1927). This basic modeling approach can be used to also model more complex choice behavior among various response options. For instance, let *a*, *b*, *c*, and *d* be the options among which a person can choose. When a person is asked to bring *a*, *b*, *c*, and *d* into a rank order, this complex choice behavior can be recoded into a total of six pairwise comparisons (does the person prefer *a*

over *b*, *a* over *c*, *a* over *d*, *b* over *c*, *b* over *d*, and *c* over *d*). Full rank information is not always available.

However, Thurstonian models can also be used when not all pairwise comparisons are available due to design. For instance, in the OMT, let *a*, *b*, *c*, and *d* be the options in a picture *i*. The answer *a* in picture *i* informs about three of six possible comparisons: the option *a* in picture *i* is stronger than option *b*, *c*, and *d*. The information that these pairwise comparisons yield is that these options were stronger than the respective other options.

The original Thurstonian model was a model that was designed to study the rank order of choices within a population. For instance, Thurstone was interested in people's general preferences for specific stimuli in apperception (Thurstone, 1927). Standard Thurstonian models thus do not take individual differences into account. Modern applications of Thurstonian models recognized that individuals can also differ in their preference for a particular option and thus transform Thurstonian models into Thurstonian IRT (Böckenholt, 2004, 2006, A. Brown & Maydeu-Olivares, 2011, 2013).

Applying the Thurstonian IRT model to the OMT

In this section we describe how OMT data can be analyzed using the dynamic Thurstonian IRT model for implicit motive measures that has earlier been described in the literature and applied to data for the PSE (Lang, 2014). To fully analyze data for an implicit motive measure using this model, it is frequently useful to test a series of nested models that includes more simple models. The most basic Thurstonian IRT model (M_1) includes a level 1 specification similar to classic Thurstonian models that do not differentiate between persons (Critchlow & Fligner, 1991). This model can be extended by adding level 2 specifications to include individual differences between persons (M_1; Böckenholt, 2001, 2006) and picture effects (M_2). The basic models (M_1 and M_2) provide the basis for specifying and testing dynamic effects in the Thurstonian IRT Model (Lang, 2014). We test two variants of dynamic effects in a replication of previous research (Lang, 2014). In the first variant, the dynamic effect is *temporary* and only lasts for one picture (M_3), in the second variant, the dynamic effect is a *sustained* effect that adds up over the complete test length (M_4). The original empirical analysis comparing these different conceptualizations of dynamic effects found that the latter (the sustained dynamic effect) described PSE data better than the former. Below, we

provide additional detail on these models. We start with a non-technical description and follow up with more statistical details. Available estimation methods are described in the demonstration below.

Level 1. The basis for applying Thurstonian IRT to the OMT and other implicit motive measures, is the insight that a motive response to a particular picture does not only provide information about the motive that is chosen in a given picture but also about the magnitude of the other motives. For instance, when an answer from person *s* in picture *i* is coded as affiliation, one gets information about comparisons between three categories: The preference for affiliation in option *i* is higher than preference for achievement, power, and other content. This reflects the idea of motive competition and is represented in level 1 of the generalized linear mixed-effects model.

Thurstonian IRT Model: M_1. In order to account for individual differences in preferring one motive category (*aff* for affiliation, *ach* for achievement, *pow* for power) to another this model needs a person parameter, also known as θ in IRT. The person parameter in M_1 makes the Thurstonian model an IRT model (Boeck, 2008; Doran, Bates, Bliese, & Dowling, 2007).

Thurstonian IRT Model: M_2. A main feature of IRT models is the analysis of picture effects, which are studied as cue strength in implicit motive research. To account for differences in mean evaluations of the motivational choice behavior options between pictures, a picture covariate is added in M_2.

Dynamic Thurstonian IRT Models: M_3 and M_4. Motivational researchers have long suggested that the strength of motives changes over time (Atkinson & Birch, 1970; Atkinson et al., 1977; Murray, 1938). One core idea has been the notion of dynamic effects. When a motive is acted out, the motive temporally loses some of its strength and behavior motivated by the respective motive becomes less likely. Behavior motivated by other motives becomes then more likely because the motives compete to be acted out.[2] After a refractory period, the strength of a motive may grow back to its previous strength. On the basis of Atkinson's theory, researchers have developed dynamic IRT models for implicit motive measures

[2] A different parametrization of these ideas was suggested by Revelle (1986).

(Tuerlinckx et al., 2003, Lang, 2014). The work on dynamic Thurstonian IRT models on which we build (Lang, 2014) includes two different dynamic Thurstonian IRT models that differ in the length of their refractory period.

In the first dynamic Thurstonian model, M_3, the dynamic effect is *temporary* and lasts only for one picture – motives recover quickly. We consider only the motive response from the last picture for this temporary dynamic effect.

In the second dynamic Thurstonian model, M_4, the dynamic effect is *sustained* and lasts for the complete test length. All previous motive related responses are added up in the sustained dynamic effect. This model represents the idea that motive strength does not recover during the test administration.

Statistical details. In the following part of the paper we describe the statistical details and formulations of the Thurstonian IRT models discussed earlier. We first describe the level 1 specifications that reflects the classic Thurstonian model that does not differentiate between people. Then we describe the level 2 specifications that add person parameters (M_1), picture parameters (M_2), and the dynamic effects (M_3 and M_4). The model formulation for the level 2 specifications are provided in Table 1.

The probability that person *s* prefers one motive category (*aff, ach, pow*, and *zero* for other content*)* to another category can be written as $\pi_{affachs}$, $\pi_{affpows}$, $\pi_{affzeros}$, $\pi_{achpows}$, $\pi_{achzeros}$, and $\pi_{powzeros}$. These are a function of differences between the mean evaluations (latent utilities) of motive categories. The mean evaluations are denoted as μ_{affs} and μ_{achs}, μ_{pows}, μ_{zeros}, and are linked to the binomial outcome using a probit link. The probit model can then be written as:

$$\begin{pmatrix} probit\left(\pi_{affachs}\right) \\ probit\left(\pi_{affpows}\right) \\ probit\left(\pi_{affzeros}\right) \\ probit\left(\pi_{achpows}\right) \\ probit\left(\pi_{achzeros}\right) \\ probit\left(\pi_{powzeros}\right) \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{affs} \\ \mu_{achs} \\ \mu_{pows} \\ 0 \end{pmatrix} = \boldsymbol{D}\mu_s.$$

The design matrix $\boldsymbol{D}$ consists of four dummy variables that refer to the four categories (the three motives and zero motive responses) and is multiplicatively related to the vector of

mean evaluations. The last row of dummy variables can be omitted without a loss of generality and the equivalent row of mean evaluations is constrained to zero. This probit model does not include an intercept.

Model M_1 adds random effects v to the motive categories to capture the variance of person *s* for the respective motive, they are normally distributed around 0. The model formulation is provided in Table 1.

Model M_2 adds a picture covariate x_{ms} $(m = 1, \dots, M)$ with *M* being the total number of pictures. If person *s* writes a story with motive content for the affiliation motive in picture *m* then $x_{affms} = 1$, if the response shows no motive content for the affiliation motive, then $x_{affms} = 0$. The mean evaluation of picture *m* for the corresponding motive category is denoted as β_m.

In applying Model M_2 to the OMT, it is important to consider some specific characteristics of this test. The OMT includes a larger number of pictures. However, some pictures in the OMT evoke stories with content from mostly one or two motives. For the other motives, the response probabilities can then be very low. IRT models that are effectively a version of the generalized linear mixed-effects model generally have difficulties to estimate effects when response probabilities are very low because the information included in these responses is limited which leads to complete separations in model estimation (Fox & Weisberg, 2011; Heinze & Schemper, 2002). The standard approach in the IRT literature is typically to exclude stimuli with extremely low or high response probabilities unless the researcher has a strong theoretical reason to assume that the stimuli actually measures the characteristic of interest (Debeer & Janssen, 2013; Wainer, Bradlow, & Wang, 2007; Wright & Stone, 1979). For the OMT, it is theoretically reasonable to assume that some pictures simply do not invoke certain motives. Therefore, it makes sense to not include pairwise comparisons for specific motives and pictures that show low response probabilities when estimating the respective latent motive.

A temporary dynamic reduction in motive strength is modelled in M_3. Therefore, another fixed effect TDE (temporary dynamic effect) can be added to model M_2 for each motive. TDE_s is simply denoting the response of the previous picture for the respective motive (0 = no response in the previous picture, 1 = one response in the previous picture).

M_4 models the sustained dynamic effect. A fixed effect SDE (sustained dynamic effect) is used instead of the temporary dynamic effect TDE of model M_3. $\mathrm{SDE_s}$ is simply denoting the total number of previous responses for the respective motive (0 = no previous response, 1 = one previous response, 2 = two previous responses, …).

Estimation Methods. Thurstonian IRT models can be estimated using broadly available software designed to model multilevel data with binary outcomes (generalized linear mixed-effects models; Böckenholt, 2004, 2006). In this paper we demonstrate the implementation of these models in the R environment (R Core Team, 2015) using the glmer function for general linear mixed-effect models (Bates, Mächler, Bolker, & Walker, 2015) in the lme4 1.1-12 package. The glmer function uses the Laplace approximation. However, other frequentist or Bayesian software packages for generalized linear-mixed effects models that allow for multiple random effects are equally capable to estimate these types of models. For instance, Lang (2014) also used Bayesian generalized linear mixed-effects models as implemented in the software package MCMCglmm (Hadfield, 2010). Furthermore, Thurstonian models can also be implemented using structural equation modeling software (Brown & Maydeu-Olivares, 2011). From our experience, generalized linear mixed-effects models software can more easily deal with missing pairwise comparisons for individual responses because it uses a long format. As one response to an OMT picture has always three missing pairwise comparisons ("take the best" response format), generalized linear mixed-effects models software may therefore be more suitable for the OMT.

IRT Reliability

One effective strategy to understand reliability in complex IRT models is to use simulation studies (Brown & Maydeu-Olivares, 2011; Lang, 2014; Reise & Yu, 1990; Weiss, 1982). The goal of this approach is to investigate how well model parameters and underlying latent dispositions can be recovered from simulated data by simulating the entire response process. Such a simulation study consists of three steps. In the first step, one simulates datasets generated on basis of the empiric sample and the Thurstonian IRT model. The basis for the simulated datasets are the estimated model parameters from the empiric dataset which are used as the underlying true population estimates. In the second step, the IRT model is fitted to these simulated datasets. In the third step, one correlates the true population estimates

with the estimates that one gets from fitting the simulated data. The goal of this step is to examine how well the IRT model is able to recover the true population estimates from the simulated data. Correlations between true motive scores and simulated IRT scores are called fidelity correlations (Weiss, 1982). Squared fidelity correlations equal the squared correlation reliability, which is known as a common definition of the reliability (Lord, Novick, & Birnbaum, 1968). We first run the simulation 100 times to get average fidelity correlations. It is then necessary to repeat the procedure again and correct for potential systematic estimation bias, which is commonly necessary for complex and multidimensional IRT models (Kuk, 1995; Ng, Carpenter, Goldstein, & Rasbash, 2006). For the bias correction, we first identify the bias as the difference between the actual picture parameters from the empirical sample and the average picture parameters from the first 100 simulation runs. In the bias corrected 100 simulation runs we subtract this bias from the empiric picture parameters used in the first step of this procedure and thereby correct for the bias.

Empirical Study

To evaluate the viability of applying the dynamic Thurstonian IRT approach (Lang, 2014) to the OMT, we analyze a dataset composed of three smaller samples (sample one n=250; sample two n = 275; sample three n = 112). We combined data to one larger set, because psychometric analysis become more robust with a higher number of respondents. Subsets of two of the samples have been published earlier (sample one: Schüler et al. (2015); sample two: Lang et al. (2012)).

The empirical study has two goals. We first want to investigate whether or not the models fit our data and evaluate which model fits best. Based on previous research (Lang, 2014) we expect model M_4, which accounts for motive competition and sustained dynamic effects, to show the best model fit. Second, we want to study the IRT reliability of the OMT and compare it to traditional reliability estimates. As the Thurstonian IRT model accounts for underlying response processes, we expect to find higher reliability coefficients than in an analysis of the sum scores with Cronbach's alpha or the greatest lower bound, which is less conservative than Cronbach's alpha in estimating reliability (Sijtsma, 2008).

Method

Participants. From the overall sample of 637, four participants with more than five not motive related or missing codings were excluded from the analysis. This approach is typically recommended to exclude participants who did not understand the test instructions and is also in line with the measure recommendations in the manual. The final sample of 633 OMTs comprises 403 women and 220 men (ten participants did not indicate their gender) with an age range from 18 to 69 (M = 31,71, SD = 12,43; 16 participants did not indicate their age) years.

Measures. Six coders evaluated the sample of 637 OMTs, using detailed information provided by the OMT manual (Kuhl & Scheffer, 1999, 2002). In addition, the coders were instructed and trained through various OMT coding seminars. In order to evaluate interrater reliability between the coders we calculated Gwet's $\boldsymbol{AC_1}$ statistic (Gwet, 2008a, 2008b), which is an improved version of Cohen's $\boldsymbol{\kappa}$ (kappa). Coder A and B coded 250 OMTs from the first sample and showed good interrater reliability (0.98 for affiliation, 0.97 for achievement, and .87 for power). A third coder C has coded the remaining 275 OMTs from the second sample. To ensure coding quality, 165 OMTs coded by coder C have also been coded by a coder D, showing again sufficient interrater reliability (0.88 for affiliation, 0.87 for achievement, and 0.76 for power). Coder E and F coded 112 OMTs from the third sample and showed sufficient interrater reliability (0.92 for affiliation, 0.90 for achievement, and .83 for power). Disagreements between coders were solved through discussions in additional sessions in the first sample. The codings of the main coders C and E were used in samples two and three. To get the most distinct approach category, we integrated the first three OMT categories for each motive into one approach factor.

Results

Table 2 shows descriptive information for the OMT picture set. Pictures clearly differ in their activation strength for the motives. Some pictures show very few responses for a particular motive. As discussed earlier, we excluded motive comparisons that had a low response probability (lower than 1%). We eliminated the comparisons for specific motives in seven pictures (pictures 1,2,4,6,7,8 and 15; see Table 5) in models M_2, M_3 and M_4.

Table 3 shows the fit statistics for all model tests. A lower log-likelihood indicates a better model fit. To analyze the first question (which model shows the best fit) we started by comparing the model fit of M_1, the most basic model, to the fit of M_2, that additionally accounts for picture effects. Picture effects seem to improve model fit. M_2 shows better model fit in comparison with M_1. This is in line with the observation of the descriptive information given in Table 2 (i.e., motive scores differ substantially between pictures). The model fit improved further by adding dynamic effects (M_3 and M_4). Both models with dynamic effects improved model fit, but the extent to which they improved model fit differs. A direct test of difference in fit for the two dynamic Thurstonian models is not possible because the models are not nested. However, the comparison of all information indices for the two models with different dynamic effects (M_3 and M_4) shows that model M_4 (sustained dynamic effects) fits best. All three sustained dynamic effects are negative (Table 5), which confirms our expectation of a dynamic implicit motive system and is in line with existing literature (Lang, 2014). Because M_4 fits best to the data, we used it as the standard model to run all following analysis.

Table 4 provides the IRT reliability estimates from the simulation study. Affiliation, achievement, and power were $r^2_{\theta\hat{\theta}}$ = .52, $r^2_{\theta\hat{\theta}}$ = .62, and $r^2_{\theta\hat{\theta}}$ = .73 respectively. Table 4 also provides Cronbach's alpha and greatest lower bound. Both were considerably lower, in line with our expectations.

Table 5 shows picture effects for model M_4 from the empiric sample and from the simulated datasets as well as the corrected simulated dataset. They indicate how likely it is to write a story with respective motive contents. For example, in picture 9, an achievement related response is more likely than an affiliation related response and an affiliation related response is more likely than a power related response. Table 5 further provides the sustained dynamic effects *SDE*. A negative value indicates that a previous story with this motive content lowers the probability of another story with the same motive content.

Table 5 also provides correlations between the IRT scores for each motive, which were low. Correlations between the sum and IRT θ scores were $r(631) = .96$ for affiliation, $r(631) = .96$ for achievement and $r(631) = .96$ for power.

Tutorial/Demonstration: Fitting Dynamic Thurstonian IRT Models Using Generalized Linear Mixed Effects Modeling Software

To enable researchers to reproduce the application of our dynamic Thurstonian IRT model on OMT data, it is possible to use the script in the Appendix, that generates the OMT dataset as the first step of the reliability estimation. Note that the OMT generated dataset are based on the parameters of the model with the best fit, M_4. It is, however, possible to fit the other models to these data for demonstrational purposes.

Fitting the Thurstonian Models

Recoding the data. As described above, one receives information about three pairwise comparisons for each motive response in the OMT. Therefore, it is necessary to recode the OMT data first. One needs three rows in the OMT data matrix for a response of one person to a particular picture. Three dummy variables (DAFF, DACH, DPOW) indicate the three motives. Table 6 shows an example of an OMT data matrix recoded into a pairwise comparison matrix. The first picture is coded as affiliation. The first row represents the comparison between affiliation and achievement, the second row between affiliation and power, and the third row between affiliation and the zero category. Similarly, picture two is an example for an achievement response, picture three is an example for a power response and picture four is an example for a zero response. Each picture consists of three rows representing the three comparisons. The dynamic effects are also specified in the recoded OMT data matrix. TDACH, TDAFF and TDPOW indicate the temporary dynamic effect. In the sample matrix presented in Table 6, the first picture is an affiliation response. It follows that TDAFF is 1 for the second picture. The second picture is an achievement response, so that TDACH is 1 for the third picture – TDAFF resets to 0. SDAFF, SDACH, and SDPOW indicate all potential previous responses of the respective motive, they reflect the sustained dynamic effects. In the sample matrix, the first picture is an affiliation response. It follows that SDAFF is 1 for the second picture. The second picture is an achievement response, so that SDACH is 1 for the third picture, SDAFF is still 1.

Fitting the data. After recoding the data, models M_1, M_2, M_3 and M_4 can be fitted using the lme4 1.1-12 package in the R environment (R Core Team, 2015). The lme 4 package fits linear mixed-effect models with the glmer function. The glmer function is a two-sided

linear formula object with the response variables `response1` and `response2` on the left side of the ~ operator and the terms of the model on the right side, separated by + operators. For model M_1, latent motive traits are modelled as random effects and written in parenthesis. Vertical bars separate the expressions for the design matrices from the grouping factor `persons`:

```
glmer(cbind(response1,response2) ~
0
+(0+DAFF+DACH+DPOW|person),
family=binomial("probit"),tdata,
control=glmerControl(optimizer="nloptwrap",optCtrl=list(maxfun=1e10),
calc.derivs=F))
```
[3]

In model M_2, fixed effects are added for the pictures:

```
tdata$PICAFF<-as.factor(ifelse(tdata$DAFF!=0,tdata$picture,1))
tdata$PICACH<-as.factor(ifelse(tdata$DACH!=0,tdata$picture,3))
tdata$PICPOW<-as.factor(ifelse(tdata$DPOW!=0,tdata$picture,1))

glmer(cbind(response1,response2) ~
0
+(0+DAFF+DACH+DPOW|person)
+PICAFF:DAFF
+PICACH:DACH
+PICPOW:DPOW,
family=binomial("probit"),tdata,
control=glmerControl(optimizer="nloptwrap",optCtrl=list(maxfun=1e10),calc.derivs=F))
```

In model M_3, temporary dynamic effects are added:

```
tdata$PICAFF<-as.factor(ifelse(tdata$DAFF!=0,tdata$picture,1))
tdata$PICACH<-as.factor(ifelse(tdata$DACH!=0,tdata$picture,3))
tdata$PICPOW<-as.factor(ifelse(tdata$DPOW!=0,tdata$picture,1))

M3<-glmer(cbind(response1,response2) ~
0
```

[3] We use calc.derivs=F because in the current lme4 version the warnings are still under development. We followed the suggestions of the lme4 authors and refitted the model with different optimizers to check the results.

```
+(0+DAFF+DACH+DPOW|person)
+PICAFF:DAFF
+PICACH:DACH
+PICPOW:DPOW
+DAFF:TDAFF+DACH:TDACH+DPOW:TDPOW,
family=binomial("probit"),tdata,
control=glmerControl(optimizer="nloptwrap",optCtrl=list(maxfun=1e10),calc.derivs=F)
```

In model M_4, sustained dynamic effects are added instead of the temporary dynamic effects from M_3:

```
tdata$PICAFF<-as.factor(ifelse(tdata$DAFF!=0,tdata$picture,1))
tdata$PICACH<-as.factor(ifelse(tdata$DACH!=0,tdata$picture,3))
tdata$PICPOW<-as.factor(ifelse(tdata$DPOW!=0,tdata$picture,1))

glmer(cbind(response1,response2) ~
0
+(0+DAFF+DACH+DPOW|person)
+PICAFF:DAFF
+PICACH:DACH
+PICPOW:DPOW
+DAFF:SDAFF+DACH:SDACH+DPOW:SDPOW,
family=binomial("probit"),tdata,
control=glmerControl(optimizer="nloptwrap",optCtrl=list(maxfun=1e10),calc.derivs=F))
```

The glmer function uses the Laplace approximation (Böckenholt, 2001) and estimates random effects with the maximum a posteriori method (De Boeck et al., 2011). In the empiric demonstration of the model we use the bobyqa optimizer from the nloptr package in R, because this optimizer is typically the fastest. The substantive conclusions did not differ using other lme4 optimizers (Bates et al., 2015).

Estimating IRT Reliability

In this part of the paper we describe the implementation of the simulation approach to estimate the IRT reliability of the OMT in the R environment. As described above, the approach determines the degree to which the model can recover true scores from simulated datasets. The approach includes three steps.

First Step. The goal of the first step is to simulate a dataset based on the empirical model. The estimated model parameters from the empirical sample are used as the underlying true population estimates to generate the new dataset. The code for this step is provided in the Appendix. The estimated model parameters used in this demonstration are from model M_4 and the empirical study that we describe in the applied part.

In the `tv` object we specify the following parameters from our empirical OMT sample. The `pictureAff, pictureAch,` and `picturePow` vectors contain the information about the empiric picture parameters. `NA` indicates that this picture does not measure the respective motive and reflects the resolution of low response probabilities. The `de` vector contains the fixed dynamic effects, the `ranefsd` vector stores the standard deviations of the random effects, which are the estimated latent traits, and the `ranefcor` object contains the correlations between those latent traits. `n` describes the number of respondents.

The `genomtdat` function generates the dataset based on the parameters stored in the `tv` object. We use the mvtnorm package in R (Genz et al., 2016), that was made to generate multivariate normal and t distributions. The resulting simulated dataset is then stored in the `tdata` matrix. The matrix is similar to the matrix provided in Table 6.

Second Step. In the second step, the IRT model, in this demonstration M_4, is fitted to the simulated dataset. The code is identical to the code provided above for the fit of model M_4.

Third Step. In the third step, one examines how well the IRT model is able to recover the true underlying latent motives with model M_4 from the simulated datasets. Therefore, we correlate the estimated latent trait scores from the simulated dataset—the outcome of the second step—with the latent trait scores from the empiric sample. The square of this correlation is called fidelity correlation (Weiss, 1982) and is equal to the squared correlation reliability, a common definition of reliability (Lord et al., 1968).

```
diag(cor(ranef(mod1)$person,unique(tdata[,2:4]))^2)
```

`ranef(mod1)$person` indicates the estimated latent trait scores from the simulated dataset and `unique(tdata[,2:4])` refers to the latent trait scores from the empiric sample.

One then simulates 100 datasets and repeats the procedure to correct for potential systematic estimation biases as described in the introduction.

Discussion

Researchers have long discussed the measurement properties of implicit motive measures (e.g., Entwisle, 1972) and in recent years have made significant progress in understanding the response processes in these types of measures (Tuerlinckx et al., 2003; Lang et al., 2012; Lang, 2014). The present paper contributes to this emerging line of research. The specific goals of our study were to (a) apply recently developed dynamic Thurstonian IRT models (Lang, 2014) to the OMT and to (b) demonstrate the use of these models with OMT data. We found that the response process in the OMT can be described using a Thurstonian IRT model with a sustained dynamic effect and this study thus extends and replicates earlier findings. Like in previous research, we also found that IRT reliabilities were much higher than internal consistency estimates. The IRT reliability for power reached .70 and the reliability for achievement was still in a range that is typically sufficient for research purposes (Ellis, 2013).

Theoretical Implications

This research has two theoretical implications for the understanding of implicit motive measures. One important finding of our study is that the Thurstonian IRT model with a sustained dynamic effect describes the response process in implicit motives measures better than the alternative Thurstonian models we examined. This finding thus generalizes from Lang (2014). A possible implication of this finding is that dynamic processes in implicit motive measures capture a fundamental mechanism in the human motivation system. Dynamic processes have always been an important assumption in dynamic system models of motivation (Atkinson & Birch, 1970; Atkinson et al., 1977; Revelle, 1986). Dynamic Thurstonian IRT makes it possible to study and model these processes. We suggest that these models could possibly also serve as a general tool to study motivational behavior. For instance, the model could be applied to social interactions in laboratory situations or be used to study motive change in diary studies to understand daily motivational processes.

This study also contributes to a deeper understanding of the exact nature of dynamic motivational processes in implicit motive measures. Atkinson and Birch (1970) suggested that

motives are temporally reduced by action for a limited period, so that motivational behavior typically can be described by a saw tooth pattern (also see Carver & Scheier, 2002; Wright, 2016). This study replicates previous research (Lang, 2014) suggesting that the dynamic Thurstonian model in which consumption of motive strength is sustained (M_4) provides a better fit than a saw tooth model (M_3). Implications of this finding is that motives have a greater inertia than originally assumed by theoretical work (Atkinson, 1950; Atkinson et al., 1977; Reitman & Atkinson, 1958). A reviewer noted that the two models we examined are not the only possible dynamic models and that in-between models are possible. Specifically, it is possible that the dynamic effects last longer than one picture (M_3) but shorter than the duration of the test (M_4). We therefore also fitted two other models, in which the dynamic effects lasted not only one but two or three pictures, before motive strength recovers. The model with a dynamic effect that lasted two pictures (deviance = 19,957; AIC = 20,047; BIC = 20,408) and the model with a dynamic effect that lasted three pictures (deviance = 19,924, AIC = 20,014, BIC = 20,375), had a somewhat better fit than M_3 (deviance = 19,970; AIC = 20,060; BIC = 20,422) but did not fit as well as M_4 (deviance = 19,204; AIC = 19,294, BIC = 19,656). Another possibility is that a general fatigue model in which the dynamic effect is not motive-specific could possibly fit the data equally good or better than M_4. We tested a model, in which the response probability for all three motives declines with each motive-related response to a picture no matter what type of motive-related response occurred. This model did not provide a better fit than M_4 (deviance = 19,964; AIC = 20,054; BIC = 20,415.

Measurement Implications

Our research has three measurement implications for research on implicit motives. One important implication of our research is that the OMT is an implicit motive measure that can be used to measure implicit motives in research. Although the IRT reliability estimates for affiliation and achievement in our study were smaller than the minimum requirements for questionnaire measures typically recommended in the literature (e.g., Nunnally, 1978), estimates around .60 are typically sufficient for research purposes (Ellis, 2013). The OMT is a procedure that is relatively easy to learn because coders do not need to separate texts in different parts and only one motive response for each picture is coded. The OMT may therefore have practical advantages for research purposes. The Thurstonian approach suggests

that it is beneficial to present pictures that stimulate multiple motives and code those motives to assess the motive of interest even if one is only interested in a single motive. This idea is also in line with other recent recommendations in the literature (Ramsay & Pang, 2013).

A second measurement implication of our study is that researchers should not interpret internal consistency estimates for the OMT. As our analyses show, Cronbach's alpha, as well as the greatest lower bound (Woodhouse & Jackson, 1977), likely do not adequately estimate the true underlying reliability of the response process.

A third measurement implication of our study is that sum score estimates based on OMT responses are likely useful for research and can be used in future research. We correlated sum scores and IRT θ scores from the dynamic Thurstonian IRT model and our results suggested that these estimates are very high. Researchers who seek to use the OMT and do not have sample sizes available that allow them to estimate IRT models, can likely confidentially use the OMT sum scores as an approximation of IRT θ.

Limitations and Future Directions

One limitation of the OMT that has become apparent in our study is the fact that some pictures do not capture some of the motives, which shortens test length. From a psychometric perspective, one potential way to further develop the OMT and increase its IRT reliability may be to include more pictures that simultaneously capture all three motives and thereby increase test length. Including ambiguous pictures in measures of implicit motives has previously been recommended (Murstein, 1965; Pang, 2010; Smith, 1992). However, it is possible that the pictures that only capture a specific motive do so particularly well. It is likely useful to examine the implications of different types of pictures for motive measurement.

Another limitation of the OMT and other implicit motive measures is the fact that these measures cannot be lengthened to an unlimited amount (Reitman & Atkinson, 1958; Schultheiss & Pang, 2007). The dynamic Thurstonian model provides a formal theory on why this is the case and in so doing also demonstrates a natural limitation of measurement procedures like the OMT. We nevertheless believe that future research could systematically investigate to what degree the OMT can be lengthened and at what point the dynamic effects become so strong that adding additional pictures is not useful anymore.

References

Atkinson, J. W. (1950). Studies in projective measurement of achievement motivation. In *Microfilm Abstracts*.

Atkinson, J. W., & Birch, D. (1970). *Dynamics of Action*. New York: Wiley.

Atkinson, J. W., Bongort, K., & Price, L. H. (1977). Explorations using computer simulation to comprehend thematic apperceptive measurement of motivation. *Motivation and Emotion*, *1*, 1–27. https://doi.org/10.1007/BF00997578

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *1*. https://doi.org/10.18637/jss.v067.i01

Baumann, N., Kaschel, R., & Kuhl, J. (2005). Striving for Unwanted Goals: Stress-Dependent Discrepancies Between Explicit and Implicit Achievement Motives Reduce Subjective Well-Being and Increase Psychosomatic Symptoms. *Journal of Personality and Social Psychology*, *89*, 781–799. https://doi.org/10.1037/0022-3514.89.5.781

Baumann, N., Kaschel, R., & Kuhl, J. (2010). Implicit Motives: A Look from Personality Systems Interaction Theory. In O. C. Schultheiss & J. C. Brunstein (Eds.), *Implicit Motives*. New York: Oxford University Press.

Baumann, N., & Scheffer, D. (2010). Seeing and mastering difficulty: The role of affective change in achievement flow. *Cognition & Emotion*, *24*, 1304–1328. https://doi.org/10.1080/02699930903319911

Böckenholt, U. (2001). Hierarchical Modeling of Paired Comparison Data. *Psychological Methods*, *6*, 49–66. https://doi.org/10.1037//1082-989X.6.1.49

Böckenholt, U. (2004). Comparative judgments as an alternative to ratings: identifying the scale origin. *Psychological Methods*, *9*, 453–65. https://doi.org/10.1037/1082-989X.9.4.453

Böckenholt, U. (2006). Thurstonian-Based Analyses: Past, Present, and Future Utilities. *Psychometrika*, *71*, 615–629. https://doi.org/10.1007/s11336-006-1598-5

Boeck, P. de. (2008). Random Item IRT Models. *Psychometrika*, *73*, 533–559. https://doi.org/10.1007/s11336-008-9092-x

Brown, A., & Maydeu-Olivares, A. (2011). Item Response Modeling of Forced-Choice

Questionnaires. *Educational and Psychological Measurement*, *71*, 460–502. https://doi.org/10.1177/0013164410375112

Brown, A., & Maydeu-Olivares, A. (2013). How IRT can solve problems of ipsative data in forced-choice questionnaires. *Psychological Methods*, *18*, 36–52. https://doi.org/10.1037/a0030641

Carver, C. S., & Scheier, M. F. (2002). Control processes and self-organization as complementary principles underlying behavior. *Personality and Social Psychology Review*, *6*(4), 304–315.

Critchlow, D. E., & Fligner, M. A. (1991). Paired comparison, triple comparison, and ranking experiments as generalized linear models, and their implementation on GLIM. *Psychometrika*, *56*, 517–533. https://doi.org/10.1007/BF02294488

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334. https://doi.org/10.1007/BF02310555

De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., Tuerlinckx, F., & Partchev, I. (2011). The Estimation of Item Response Models with the lmer Function from the lme4 Package in R. *Journal Of Statistical Software*, *39*, 1–28.

Debeer, D., & Janssen, R. (2013). Modeling Item-Position Effects Within an IRT Framework. *Journal of Educational Measurement*, *50*, 164–185.

Doran, H., Bates, D., Bliese, P., & Dowling, M. (2007). Estimating the Multilevel Rasch Model : With the lme4 package. *Journal of Statistical Software*, *20*, 1–18. https://doi.org/10.1111/j.1467-9868.2007.00600.x

Ellis, J. L. (2013). A standard for test reliability in group research. *Behavior Research Methods*, *45*, 16–24. https://doi.org/10.3758/s13428-012-0223-z

Emmons, R. A. (1989). Exploring the Relations Between Motives and Traits: The Case of Narcissism. In D. Buss & N. Cantor (Eds.), *Personality Psychology SE - 3* (pp. 32–44). Springer US. https://doi.org/10.1007/978-1-4684-0634-4_3

Entwisle, D. R. (1972). To Dispel Fantasies About Fantasy-Based Measures of Achievement Motivation. *Psychological Bulletin*, *77*, 377–391. https://doi.org/10.1037/h0020021

Fineman, S. (1977). The achievement motive construct and its measurement: Where are we now? *British Journal of Psychology*, *68*, 1–22. https://doi.org/10.1111/j.2044-

8295.1977.tb01554.x

Fox, J., & Weisberg, S. (2011). *An R Companion to Applied Regression*. SAGE Publications. Retrieved from https://books.google.be/books?id=YH6NotdvzF0C

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., & Hothorn, T. (2016). mvtnorm: Multivariate Normal and t Distributions. Retrieved from http://cran.r-project.org/package=mvtnorm

Gwet, K. L. (2008a). Computing inter-rater reliability and its variance in the presence of high agreement. *The British Journal of Mathematical and Statistical Psychology*, *61*, 29–48. https://doi.org/10.1348/000711006X126600

Gwet, K. L. (2008b). Variance Estimation of Nominal-Scale Inter-Rater Reliability with Random Selection of Raters. *Psychometrika*, *73*, 407–430. https://doi.org/10.1007/s11336-007-9054-8

Hadfield, J. D. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software*, *1*(2). https://doi.org/10.18637/jss.v033.i02

Hall, J. L., Stanton, S. J., & Schultheiss, O. C. (2010). Biopsychosocial and Neural Processes of Implicit Motivation. In O. C. Schultheiss & J. C. Brunstein (Eds.), *Implicit Motives* (pp. 279–307). New York: Oxford University Press.

Heinze, G., & Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine*, *21*(16), 2409–2419. https://doi.org/10.1002/sim.1047

Hofer, J., Busch, H., Chasiotis, A., Kärtner, J., & Campos, D. (2008). Concern for Generativity and Its Relation to Implicit Pro-Social Power Motivation, Generative Goals, and Satisfaction With Life: A Cross-Cultural Investigation. *Journal of Personality*, *76*, 1–30. https://doi.org/10.1111/j.1467-6494.2007.00478.x

Kazén, M., & Kuhl, J. (2005). Intention memory and achievement motivation: volitional facilitation and inhibition as a function of affective contents of need-related stimuli. *Journal of Personality and Social Psychology*, *89*, 426–448. https://doi.org/10.1037/0022-3514.89.3.426

Kazén, M., & Kuhl, J. (2011). Directional discrepancy between implicit and explicit power motives is related to well-being among managers. *Motivation and Emotion*, *35*, 317–327.

https://doi.org/10.1007/s11031-011-9219-8

Köllner, M. G., & Schultheiss, O. C. (2014). Meta-analytic evidence of low convergence between implicit and explicit measures of the needs for achievement, affiliation, and power. *Frontiers in Psychology, 5*, 826. https://doi.org/10.3389/fpsyg.2014.00826

Kuhl, J., & Kazén, M. (2008). Motivation, affect, and hemispheric asymmetry: power versus affiliation. *Journal of Personality and Social Psychology, 95*, 456–469. https://doi.org/10.1037/0022-3514.95.2.456

Kuhl, J., & Scheffer, D. (1999). Auswertungsmanual für den Operanten Motiv-Test (OMT).

Kuhl, J., & Scheffer, D. (2002). Der operante Multi-Motiv-Test (OMT): Manual [The operant multi-motive-test (OMT): Manual]. Osnabrück, Germany: University of Osnabrück.

Kuk, Y. C. A. (1995). Asymptotically unbiased estimation in generalized linear models with random effects. *Journal of the Royal Statistical Society: Series B (Methodological), 57*, 395–407.

Lang, J. W. B. (2014). A dynamic Thurstonian item response theory of motive expression in the picture story exercise: Solving the internal consistency paradox of the PSE. *Psychological Review, 121*, 481–500. https://doi.org/10.1037/a0037011

Lang, J. W. B., Zettler, I., Ewen, C., & Hülsheger, U. R. (2012). Implicit motives, explicit traits, and task and contextual performance at work. *Journal of Applied Psychology, 97*, 1201–1217. https://doi.org/10.1037/a0029556

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The Scientific Status of Projective Techniques. *Psychological Science in the Public Interest, 1*, 27–66. https://doi.org/10.1111/1529-1006.002

Lord, F. M., Novick, M. R., & Birnbaum, A. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

McAdams, D. P., Jackson, R. J., & Kirshnit, C. (1984). Looking, laughing, and smiling in dyads as a function of intimacy motivation and reciprocity. *Journal of Personality, 52*, 261–273. https://doi.org/10.1111/j.1467-6494.1984.tb00881.x

McClelland, D. C. (1987). *Human motivation*. CUP Archive.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). The measuring instrument. In *The achievement motive* (pp. 185–217). CHAP, East Norwalk, CT, US:

Appleton-Century-Crofts. https://doi.org/10.1037/11144-007

McClelland, D. C., & Boyatzis, R. E. (1982). Leadership motive pattern and long-term success in management. *Journal of Applied Psychology*, *67*, 737–743. https://doi.org/10.1037/0021-9010.67.6.737

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*, 690–702. https://doi.org/10.1037/0033-295X.96.4.690

Morgan, C. D., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology & Psychiatry*, *34*, 289–306. https://doi.org/10.1001/archneurpsyc.1935.02250200049005

Murray, H. A. (1938). *Explorations in Personality; A Clinical and Experimental Study of Fifty Men of College Age. American Sociological Review*. New York: Oxfort University Press. https://doi.org/10.2307/2084329

Murstein, B. I. (1965). Projection of hostility on the TAT as a function of stimulus, background, and personality variables. *Journal of Consulting Psychology*, *29*, 43–48. https://doi.org/10.1037/h0021666

Ng, E. S. W., Carpenter, J. R., Goldstein, H., & Rasbash, J. (2006). Estimation in generalised linear mixed models with binary outcomes by simulated maximum likelihood. *Statistical Modelling*, *6*, 23–42. https://doi.org/10.1191/1471082X06st106oa

Nunnally, J. C. (1978). *Psychometric Theory*. McGraw-Hill Book Company.

Pang, J. S. (2010). Content coding methods in implicit motive assessment: Standards of measurement and best practices for the picture story exercise. In O. C. Schultheiss & J. C. Brunstein (Eds.), *Implicit Motives* (pp. 119–1550). New York: Oxford University Press.

R Core Team. (2015). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, *v3.0.1*, 3604. https://doi.org/10.1007/978-3-540-74686-7

Ramsay, J. E., & Pang, J. S. (2013). Set ambiguity: A key determinant of reliability and validity in the picture story exercise. *Motivation and Emotion*, *37*, 661–674. https://doi.org/10.1007/s11031-012-9339-9

Reise, S. P., & Yu, J. (1990). Parameter Recovery in the Graded Response Model Using MULTILOG. *Journal of Educational Measurement*, *27*, 133–144. https://doi.org/10.1111/j.1745-3984.1990.tb00738.x

Reitman, W. R., & Atkinson, J. W. (1958). Some methodological problems in the use of thematic apperceptive measures of human motives. In J. W. Atkinson (Ed.), *Motives in fantasy, action, and society* (pp. 664–683). Princeton, NJ: Van Nostrand.

Revelle, W. (1986). Motivation and Efficiency of Cognitive Performance BT - Frontiers of Motivational Psychology: Essays in Honor of John W. Atkinson. In D. R. Brown & J. Veroff (Eds.) (pp. 107–131). New York, NY: Springer US. https://doi.org/10.1007/978-1-4684-6341-5_7

Revelle, W. (2016). psych: Procedures for Psychological, Psychometric, and Personality Research. Evanston, Illinois. Retrieved from https://cran.r-project.org/package=psych

Scheffer, D. (2005). *Implizite Motive [Implicit motives]*. Göttingen, Germany: Hogrefe.

Schüler, J., Brandstätter, V., Wegner, M., & Baumann, N. (2015). Testing the convergent and discriminant validity of three implicit motive measures: PSE, OMT, and MMG. *Motivation and Emotion*, *39*, 839–857. https://doi.org/10.1007/s11031-015-9502-1

Schüler, J., Job, V., Fröhlich, S. M., & Brandstätter, V. (2008). A high implicit affiliation motive does not always make you happy: A corresponding explicit motive and corresponding behavior are further needed. *Motivation and Emotion*, *32*, 231–242. https://doi.org/10.1007/s11031-008-9096-y

Schüler, J., Job, V., Fröhlich, S. M., & Brandstätter, V. (2009). Dealing with a `hidden stressor': Emotional disclosure as a coping strategy to overcome the negative effects of motive incongruence on health. *Stress and Health*, *25*, 221–233. https://doi.org/10.1002/smi.1241

Schultheiss, O. C. (2008). Implicit motives. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 603–633). New York: The Guilford Press.

Schultheiss, O. C., Liening, S. H., & Schad, D. (2008). The reliability of a Picture Story Exercise measure of implicit motives: Estimates of internal consistency, retest reliability, and ipsative stability. *Journal of Research in Personality*, *42*, 1560–1571.

https://doi.org/10.1016/j.jrp.2008.07.008

Schultheiss, O. C., & Pang, J. S. (2007). Measuring implicit motives. In R. W. Robins, R. C. Fraley, & R. Krueger (Eds.), *Handbook of Research Methods in Personality Psychology* (pp. 322–344). New York: Guilford.

Sijtsma, K. (2008). On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha. *Psychometrika*, *74*, 107–120. https://doi.org/10.1007/s11336-008-9101-0

Smith, C. P. (1992). *Motivation and personality: Handbook of thematic content analysis*. New York: Cambridge University Press.

Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin*, *112*, 140–154. https://doi.org/10.1037/0033-2909.112.1.140

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, *34*, 273–286. https://doi.org/10.1037/h0070288

Tuerlinckx, F., De Boeck, P., & Lens, W. (2002). Measuring needs with the Thematic Apperception Test: A psychometric study. *Journal of Personality and Social Psychology*, *82*, 448–461. https://doi.org/10.1037//0022-3514.82.3.448

Verguts, T., & De Boeck, P. (2000). A Rasch Model for Detecting Learning While Solving an Intelligence Test. *Applied Psychological Measurement*, *24*, 151–162. https://doi.org/10.1177/01466210022031589

Verhelst, N. D., & Glas, C. A. W. (1993). A dynamic generalization of the Rasch model. *Psychometrika*, *58*, 395–414. https://doi.org/10.1007/BF02294648

Wainer, H., Bradlow, E. T., & Wang, X. (2007). *Testlet response theory and its applications*. Cambridge University Press.

Weiss, D. J. (1982). Improving Measurement Quality and Efficiency with Adaptive Testing. *Applied Psychological Measurement*, *6*, 473–492. https://doi.org/10.1177/014662168200600408

Winter, D. G., John, O. P., Stewart, A. J., Klohnen, E. C., & Duncan, L. E. (1998). Traits and motives: Toward an integration of two traditions in personality research. *Psychological Review*, *105*, 230–250. https://doi.org/10.1037/0033-295X.105.2.230

Woodhouse, B., & Jackson, P. (1977). Lower bounds for the reliability of the total score on a

test composed of non-homogeneous items: II: A search procedure to locate the greatest lower bound. *Psychometrika*, *42*, 579–591. https://doi.org/10.1007/BF02295980

Wright, B. D., & Stone, M. H. (1979). *Best Test Design*. Univ of Chicago Social Research.

Wright, R. A. (2016). Motivation theory essentials: Understanding motives and their conversion into effortful goal pursuit. *Motivation and Emotion*, *40*(1), 16–21. https://doi.org/10.1007/s11031-015-9536-4

Figure

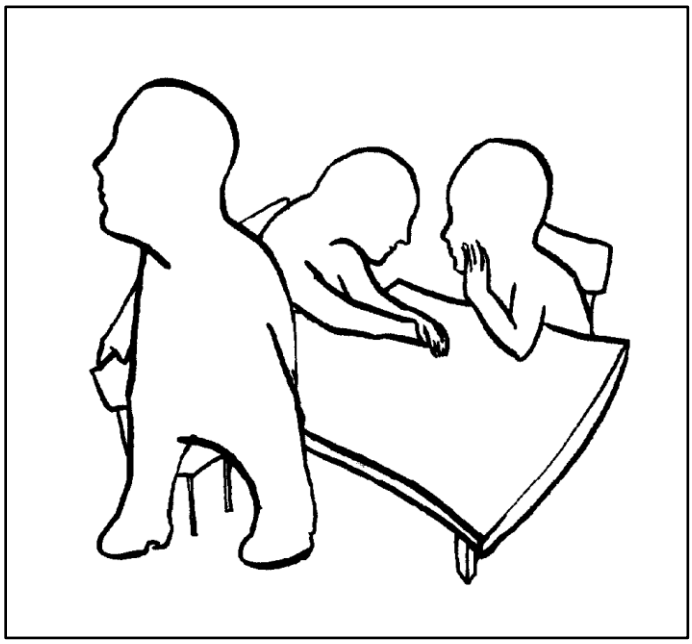

Figure 1: Example picture of the OMT.

Tables

Table 1

Formulas Model M_1-M_4

Latent utility		Effects added by model M_1	M_2	M_3	or	M_4
μ_{aff}	=	ν_{affs} +	$\sum_{m=1}^{M} x_{affms}\,\beta_{affm}$	+ $\mathrm{TDE}_{affs}\beta_{affm}$	or	$\mathrm{SDE}_{affs}\beta_{affm}$
μ_{ach}	=	ν_{achs} +	$\sum_{m=1}^{M} x_{achms}\,\beta_{achm}$	+ $\mathrm{TDE}_{achs}\beta_{achm}$	or	$\mathrm{SDE}_{achs}\beta_{achm}$
μ_{pow}	=	ν_{pows} +	$\sum_{m=1}^{M} x_{powms}\,\beta_{powm}$	+ $\mathrm{TDE}_{pows}\beta_{powm}$	or	$\mathrm{SDE}_{pows}\beta_{powm}$
μ_{zero}	=	0				

Table 2

Descriptive Information

Picture	Affiliation %	Achievement %	Power %	Other %
1	31.4	0.0	36.9	31.7
2	70.5	0.2	18.4	10.9
3	18.1	3.6	13.3	65.0
4	6.3	0.3	56.0	37.4
5	4.2	9.6	1.4	84.8
6	0	65.0	0.6	34.4
7	0.9	23.1	41.9	34.1
8	0.5	2.4	38.0	59.1
9	2.7	43.0	6.6	47.7
10	1.4	3.0	40.2	55.4
11	3.5	6.0	72.8	17.7
12	5.8	15.1	66.7	12.4
13	10.2	5.7	40.3	43.8
14	16.2	5.2	39.7	38.9
15	0.5	0.5	39.6	59.4

Table 3

Model Comparisons

Thurstonian Model	*df*	*logLik*	Deviance	AIC	BIC	Δ*df*	χ^2
M_1: Motivational Conflict	7	-13,942.9	27,886	27,900	27,956		
M_2: Picture effects	42	-10,020.1	20,040	20,124	20,461		
M_1vs. M_2						35	7,846*
M_3: Picture effects with TDE	45	-9,985.2	19,970	20,060	20,422		
M_2 vs. M_3						3	69.85*
M_4: Picture effects with SDE	45	-9,602.2	19,204	19,294	19,656		
M_2 vs. M_4						3	835.9*

Note. TDE = temporary dynamic effect. SDE = sustained dynamic effect k = 22,601 pair-wise comparisons nested in n = 633 persons and 15 pictures. *p <. 0.0001.

Table 4

Squared Correlation Reliability ($r^2_{\theta\hat{\theta}}$) for M_4 and Traditional Measures of Reliability in Bias-Corrected Simulated Data

Motive	$r^2_{\theta\hat{\theta}}$	Cronbach's α	glb
Affiliation	.52	.17	.29
Achievement	.62	.28	.40
Power	.73	.46	.56

Note. $r^2_{\theta\hat{\theta}}$ = squared correlation reliability estimated by squaring the correlation between the true underlying latent motives θ and the estimated maximum a-posteriori estimates of the latent motives $\hat{\theta}$ from M_3. glb = greatest lower bound estimate of reliability (Woodhouse & Jackson, 1977) estimated using the glb.algebraic function in the psych package (Revelle, 2016; version 1.6.9) in the R environment (R Core Team, 2014).

Table 5

Parameter Estimates for Model M4

	M_4			Average across 100 simulated datasets based on M_4			Average across 100 simulated datasets based on M_4 and bias-corrected		
Estimate	Aff	Ach	Pow	Aff	Ach	Pow	Aff	Ach	Pow
Fixed effects (β)									
Picture 1	-0.38	-	0.05	-0.77	-	0.00	-0.36	-	0.01
Picture 2	2.07	-	0.42	2.17	-	0.35	2.31	-	0.44
Picture 3	0.01	-2.84	-1.02	0.79	-3.25	-0.79	0.39	-2.87	-1.00
Picture 4	-0.03	-	0.57	0.76	-	0.55	0.22	-	0.56
Picture 5	-0.72	-2.15	-2.22	0.40	-2-24	-1.70	-0.41	-2.14	-
Picture 6	-	0.77	-	-	0.76	-	-	0.81	-
Picture 7	-	0.17	0.60	-	0.38	0.67	-	0.27	0.60
Picture 8	-	-1.53	0.25	-	-1.00	0.48	-	-1.45	0.29
Picture 9	-0.51	0.85	-0.66	0.45	1.30	-0.30	-0.27	1.00	-0.61
Picture 10	-0.94	-0.81	0.51	0.15	0.00	0.77	-0.73	-0.64	0.54
Picture 11	0.52	0.51	1.90	1.46	1.14	1.90	0.72	0.67	1.88
Picture 12	1.11	1.41	2.30	2.04	2.04	2.34	1.39	1.60	2.33
Picture 13	0.77	0.01	1.30	1.94	0.87	1.65	1.09	0.21	1.38
Picture 14	1.33	0.14	1.50	2.61	1.02	1.87	1.65	0.35	1.58
Picture 15	-	-	1.32	-	-	1.82	-	-	1.45
SDE	-1.34	-0.91	-0.34	-2.16	-1.38	-0.43	-1.55	-1.03	-0.36
Random effects									
SD	1.66	1.49	0.98	2.43	1.92	1.02	1.80	1.58	0.99
Correlations									
Aff	-			-			-		
Ach	0.19	-		0.11	-		0.09	-	
Pow	0.02	0.31	-	-0.10	0.20	-	-0.08	0.26	-

Note. k = 22,601 pair-wise comparisons nested in n = 633 persons and 15 pictures. Aff = affiliation vs other. Ach = achievement vs other. Pow = power vs other. SDE = sustained dynamic effect operationalized as the number of previous motive related responses.

Table 6

Recoded OMT data matrix

Person	Picture	Code	DAFF	DACH	DPOW	TDAFF	TDACH	TDPOW	SDAFF	SDACH	SDPOW	Response1	Response2
1	1	AFF	1	-1	0	0	0	0	0	0	0	1	0
1	1	AFF	1	0	-1	0	0	0	0	0	0	1	0
1	1	AFF	1	0	0	0	0	0	0	0	0	1	0
1	2	ACH	1	-1	0	1	0	0	1	0	0	0	1
1	2	ACH	0	1	-1	1	0	0	1	0	0	1	0
1	2	ACH	0	1	0	1	0	0	1	0	0	1	0
1	3	POW	1	0	-1	0	1	0	1	1	0	0	1
1	3	POW	0	1	-1	0	1	0	1	1	0	0	1
1	3	POW	0	0	1	0	1	0	1	1	0	1	0
1	4	ZERO	1	0	0	0	0	1	1	1	1	0	1
1	4	ZERO	0	1	0	0	0	1	1	1	1	0	1
1	4	ZERO	0	0	1	0	0	1	1	1	1	0	1

Note. DAFF, DACH, and DPOW are dummy variables to indicate the three comparisons for each coding. TDAFF, TDACH and TDPOW indicate the answer to the previous picture for the respective motive. SDAFF, SDACH, and SDPOW indicate all previous answers for the respective motive. Response 1 and Response 2 are binary coded response vectors.

Appendix

```
library(mvtnorm)

tv<-c(
  picturenAff=c(-0.38,2.07,0.01,0.03,-0.72,NA,NA,NA,-0.51,-0.94,0.52,1.11,0.77,1.33,NA),
  picturenAch=c(NA,NA,-2.84,NA,-2.15,0.77,0.17,-1.53,0.85,-0.81,0.51,1.41,0.01,0.14,NA),
  picturenPow=c(0.05,0.42,-1.02,0.57,-2.22,NA,0.60,0.25,-0.66,0.51,1.90,2.30,1.30,1.50,1.32),
  de=c(-1.34 ,-0.91,-0.34),
  ranefsd=c(1.66,1.49,0.98),
  ranefcor=c(0.19,0.02,0.31),
  n=633
)

genomtdat<-function(tv){
  rp=tv[paste("ranefcor",1:3,sep="")]
  sp=tv[paste("ranefsd",1:3,sep="")]
  res <-  rmvnorm(tv[c("n")], mean = rep(0,3),
                  sigma = matrix(c(1,
        rp[1],rp[2],rp[1],1,rp[3],rp[2],rp[3],1),3,byrow=T)*sp%*%t(sp))

  pes=matrix(tv[1:45],nrow=15,ncol=3)
  idata<-matrix(c(rep(1:tv[c("n")],each=15*6),rep(as.vector(res),each=15*6),

        rep(rep(1:15,each=6),tv[c("n")]),rep(NA,tv[c("n")]*15*6*12)),ncol=17,nrow=tv[c("n")]*15
        *6)
  for (i in 1:tv[c("n")]) {
    for (j in 1:15) {

      if (j==1) { tonext<-c(0,0,0) }
      if (j==1) { tonext2<-c(0,0,0) }
      uvnr<-c(res[i,]+pes[j,]+tonext*tv[46:48],0)

      index=(i*15*6-(15*6-1))+(j-1)*6
      a<-ifelse(is.na(uvnr),-1e5,uvnr)
      ra=rnorm(4,a,1)

      aff=ifelse(ra[1]==max(ra),1,0)
      ach=ifelse(ra[2]==max(ra),1,0)
      pow=ifelse(ra[3]==max(ra),1,0)
      rest=ifelse(ra[4]==max(ra),1,0)

      idata[index,6:17]<-c(ifelse(is.na(uvnr[1]),NA,aff),ifelse(is.na(uvnr[2]),NA,ach),1,-
        1,0,0,tonext,tonext2)
      idata[index+1,6:17]<-c(ifelse(is.na(uvnr[1]),NA,aff),ifelse(is.na(uvnr[3]),NA,pow),1,0,-
        1,0,tonext,tonext2)
      idata[index+2,6:17]<-c(ifelse(is.na(uvnr[2]),NA,ach),ifelse(is.na(uvnr[3]),NA,pow),0,1,-
        1,0, tonext,tonext2)
      idata[index+3,6:17]<-c(ifelse(is.na(uvnr[1]),NA,aff),rest,1,0,0,-1, tonext,tonext2)
      idata[index+4,6:17]<-c(ifelse(is.na(uvnr[2]),NA,ach),rest,0,1,0,-1, tonext,tonext2)
      idata[index+5,6:17]<-c(ifelse(is.na(uvnr[3]),NA,pow),rest,0,0,1,-1,tonext,tonext2)

      tonext<-c(ifelse(aff>0,tonext[1]+1,tonext[1]),ifelse(ach>0,tonext[2]+1,tonext[2]),
                ifelse(pow>0,tonext[3]+1,tonext[3]))
      tonext2<-c(ifelse(aff>0,1,0),ifelse(ach>0,1,0),
                 ifelse(pow>0,1,0))
    }
  }
  idata<-data.frame(idata[idata[,6]!=0|idata[,7]!=0,])
  idata<-na.exclude(idata)
  names(idata)<-c("person","trueaff","trueach","truepow","picture",

        "response1","response2","DAFF","DACH","DPOW","d4","SDAFF","SDACH","SDPOW","TDAFF","TDAC
        H","TDPOW")
  return(idata)
}
tdata<-genomtdat(tv)
```

CHAPTER 3

IMPROVING THE ASSESSMENT OF IMPLICIT MOTIVES USING IRT: CULTURAL DIFFERENCES AND DIFFERENTIAL ITEM FUNCTIONING[4]

Researchers have long been interested in studying differences in implicit motive between different groups. Implicit motives are typically measured by scoring text that respondents have written in response to picture cues. Recently, research on the measurement of implicit motives has made progress through the application of a dynamic Thurstonian item response theory model (DTM; Lang, 2014) that captures two basic motivational processes in motivational research: Motive competition and dynamic reduction of motive strength after a motive has been acted out. In this paper, the authors use the DTM to investigate Differential Item Functioning (DIF) in implicit motive measures. The chapter first discusses DIF in the context of the DTM. The authors then conduct a DIF analysis of data from a study that used a picture set of the Operant Motive Test (OMT; Kuhl & Scheffer, 2002) with participants from Cameroon, Germany, and Costa Rica. Results showed no evidence of DIF in 9 and some evidence for DIF in 3 pictures. The authors show a partial invariance model can be specified and use this partial invariance model to study latent mean differences between Cameroon, Germany, and Costa Rica. The discussion focuses on the use of IRT DIF methods in future research on implicit motives.

---

[4] This chapter is based on: Runge, J. M., Lang, J.W.B., Chasiotis, A., & Hofer, J. (2018). Improving the assessment of implicit motives using IRT: Cultural differences and differential item functioning. *Journal of Personality Assessment*. http://dx.doi.org/10.1080/00223891.2017.1418748

Motivational researchers have long been interested in cultural differences (and similarities) in the development of implicit motives (Chasiotis, Bender, & Hofer, 2014) and cultural differences in the link between implicit motives and behavioral outcomes (e.g. Hofer & Bond, 2008; Hofer, Kärtner, Chasiotis, Busch, & Kiessling, 2007). One core challenge in cultural group comparisons using implicit motive measures is that respondents from different cultural groups may differ in the way they interpret picture cues. Differences between cultures in implicit motive measures would then not result from real differences in implicit motives but from cultural bias in picture perception. A methodological tool designed to distinguish between real existing group differences (impact) and test bias (Thissen, Steinberg, & Gerrard, 1986) is Differential Item Functioning (DIF). The core goal of DIF is to identify response options that function differently across groups.

DIF is typically studied using Item Response Theory (IRT). However, IRT-DIF analyses require a viable IRT measurement model for the psychological individual differences construct of interest. The response process in implicit motive measures is complex because the motives do not act independently (Atkinson & Birch, 1970). Recently, researchers developed an IRT measurement model—the dynamic Thurstonian IRT model (DTM; Lang, 2014)—that accounts for the dependency of the motives and applied variants of this model to implicit motive measures (Lang, 2014; Lang, Zettler, Ewen, & Hülsheger, 2012; Runge et al., 2016). This measurement model builds on recent advancements in the modelling of forced choice data using Thurstonian IRT (Brown & Maydeu-Olivares, 2013) and additionally accounts for a dynamic reduction in motive strength after motive enactment (Tuerlinckx, De Boeck, & Lens, 2002; Verhelst & Glas, 1993).

This chapter seeks to contribute to the literature in two ways. First, we contribute to the emerging literature on Thurstonian IRT modeling for implicit motive measures and show that the DTM can be used to study implicit motives across cultures. We describe how a frequently used method to asses DIF in IRT models—the likelihood-ratio method—can be used to analyze and identify DIF in DTM models for implicit motive measures. We also describe how the results of a likelihood-ratio analysis can be used to build a partially invariant DTM model that accounts for pictures with DIF in estimating latent motive scores.

The second contribution is a demonstration of DTM-based cross-cultural analyses using a dataset of the Operant Motive Test (OMT; Kuhl & Scheffer, 2002), a measure for implicit motives. This dataset includes participants from three cultures: Cameroon, Germany, and Costa Rica (Chasiotis, Hofer, & Campos, 2006). Our demonstration shows how DTM-based analyses can provide insights into differences in picture perception and implicit motives across cultures.

Implicit Motives

Implicit motives are described as associative networks that connect situational cues with basic affective reactions and implicit behavioral tendencies. These behavioral tendencies are nonconscious dispositions to seek or avoid certain classes of incentives (McClelland, 1984; McClelland, Atkinson, Clark, & Lowell, 1953). The incentives are typically classified in three categories: need for affiliation, need for achievement, and need for power. The implicit *affiliation* motive is described as an interest to establish, maintain, or restore positive relationships with others, the implicit *achievement* motive as an interest to improve one's performance and the *power* motive as the interest to impress, influence and control others (Schultheiss & Brunstein, 2010).

Cultural Differences

In the early years of research on implicit motives, McClelland and colleagues studied cultural aspects of the need for achievement (McClelland, 1961; McClelland et al., 1953). In later years, research was typically limited to Euro−American cultural contexts. Over a decade ago, researchers started to revive cross-cultural studies on implicit motives.

Cross-cultural research on implicit motives has focused on two lines of research. The first line of research focused on cultural differences in implicit motives and on culture-specific differences in behavioral predictions of implicit motives. Contributing to this first line of research, researchers found cultural differences in need for achievement (e.g. Cameroon vs. Costa Rica: $d = 0.88$; Busch, Hofer, Chasiotis, & Campos, 2013) and need for power (e.g., Germany vs. Cameroon: $d = .58$; Hofer, Busch, Chasiotis, Kärtner, & Campos, 2008). A sample including managers from 24 countries also provided evidence for systematic differences in need for achievement and need for affiliation (van Emmerik, Gardner, Wendt, & Fischer, 2010).

The second line of research focused on universal characteristics and showed similarities in the development and behavioral predictions of implicit motives among people with various cultural backgrounds. Studies contributing to this second line of research have, for instance, shown that a high congruency between implicit and explicit achievement and affiliation motives was related to higher life satisfaction of participants recruited in Euro-American, Latin American, and African cultures (Hofer & Chasiotis, 2003; Hofer, Chasiotis, & Campos, 2006). Research in this area has also documented that people—across various cultures—who are capable of accessing their implicit motives have a higher congruency between explicit and implicit motives, and commit stronger to self-congruent goals (Hofer et al., 2010). Finally, cross-cultural research has also found support for the universality of the developmental link between a pro-social power orientation and childhood context variables like parental Socioeconomic Status and number of siblings (Chasiotis et al., 2014, 2006). Although both lines of cultural research on implicit motives have made considerable advances, knowledge on the influence of cultural contexts on implicit motives is still limited (Hofer & Chasiotis, 2011).

Bias in Cross Cultural Research of Implicit Motive Measures

Cross-cultural studies on implicit motive measures come along with various methodological challenges. Van de Vijver (2000) addressed some of those challenges and formulated a theoretical background to evaluate Thematic Apperception Test (TAT) type measures used in cross-cultural research. The key concept in this evaluation is *bias,* which is defined as "a lack of similarity of psychological meaning of test scores across cultural groups" (Van de Vijver, 2000, p.88). He describes three types of bias: *construct* bias, *method* bias, and *item* bias.

Construct bias exists if the measured construct differs across cultural groups. Sources of construct bias can be only partially overlapping construct definitions or incomplete assessment of various aspects of the construct. Method bias occurs when test samples differ in relevant characteristics, when test characteristics such as familiarity with stimulus material differs between countries, or when the person administering the test influences the test score.

Item bias is present when people with the same level in a certain trait that come from different cultures have a different probability to endorse an item. Sources for item bias can be

that stimulus material, such as a picture, invokes different psychological functions or has a specific meaning only in one of the tested cultures. Thus, item bias is systematic and not the same as noise, or measurement error. Item bias is typically studied using DIF analysis. DIF analysis can be done with IRT and non-IRT techniques.

Researchers have successfully identified DIF pictures in implicit motive tests using non-IRT techniques such as the Mantel–Haenszel procedure (Hofer & Chasiotis, 2004; Hofer, Chasiotis, Friedlmeier, Busch, & Campos, 2005) or loglinear model analysis (Busch et al., 2013; Hofer et al., 2005). In the Mantel–Haenszel procedure one applies chi-square statistics to test whether the number of motive responses is equally distributed among two cultures for a picture (Hofer & Chasiotis, 2004). In the loglinear model analysis one tests hierarchically related models that successively include the following effects: score level, culture, and interaction of culture and score level (Hofer et al., 2005). If the model with only score level fits best to the data, an item is considered as not biased. Both procedures categorize respondents score levels into low–medium–high.

In this study, we analyze item bias using an IRT based approach. The procedure we suggest expands non-IRT techniques and enables precise parameter estimations using a model-based separation of latent motives and item effects. A central feature of the IRT approach is that it enables partial invariance models. In partial invariance models, one estimates item parameters separately for each group in case DIF was detected. Therefore, one does not need to fully exclude DIF items in order to be able to estimate meaningful group differences in implicit motives.

DIF in Implicit Motive Measures Using DTM

DIF is present when an item has a different probability to be solved (or acknowledged) by respondents with the same latent trait who belong to qualitatively different groups, such as groups with different cultural backgrounds. DIF analysis is a statistical method—typically applied in IRT—to detect biased test items. The goal of a DIF analysis is to distinguish between real existing group differences in latent traits and test bias (Thissen, Steinberg, & Gerrard, 1986).

Dynamic Thurstonian IRT

The most prevalent measurement model for implicit motive tests has always been the sum score model (classical test theory). Because researchers found low internal consistencies (which are part of classical test theory methodology), but also higher retest reliabilities and strong validity evidence, it seems unlikely that the sum score model is adequate (Atkinson & Birch, 1970; Tuerlinckx, De Boeck, & Lens, 2002). A prominent explanation of why classical test theory fails to adequately measure reliability is given in the *Dynamics of Action* theory (DoA; Atkinson & Birch, 1970). The authors describe—among other processes—two response processes that influence the response to stimuli in implicit motive measures. The first process is termed *change of activity* and describes that an individual changes an action, when the strength of another action tendency is higher, which implies that action tendencies compete for enactment. The second process is the *consummatory force*, which describes the reduction of an action tendency after enactment. For instance, when a person is hungry, the action strength of eating is high, so that the action eating will win the competition against other action tendencies. When the person has finished eating, the action tendency for eating will be reduced and other action tendencies will be higher. The typical explanation of this consummatory force is, that the need has been satisfied and temporarily loses some of its strength. After a refractory period, the motive strength returns to its original strength. This leads, according to DoA, to a waning and waxing of motive specific behavior.

With advancements in IRT modelling it became possible to test these processes. Tuerlinckx, De Boeck, and Lens (2002) tested the assumed reduction of motive strength in a series of models. The first model was a basic IRT model, the second model was a dynamic model that included a reduction of motive strength after enactment, and the third model was a stochastic drop-out model that accounts for behavior which was not led by the studied motive. The best fitting model was the third model. These IRT models included only one motive so that motive competition was not examined.

The most recent and refined measurement model for implicit motives is the Dynamic Thurstonian IRT (DTM) model (Lang, 2014) that builds on recent advancements in the modelling of paired comparison data (Böckenholt, 2001) and forced choice questionnaires (Brown & Maydeu-Olivares, 2013) to also model motive competition. The DTM also models dynamic reduction of motive strength. This model enabled researchers to test the process of

reduction in motive strength after enactment suggested in DoA. Studies for the Picture Story Exercise (PSE; Lang, 2014) as well as the OMT (Runge et al., 2016) have used the DTM as a framework to test different conceptualizations of motive reduction after enactment: A temporary effect as suggested in DoA, a longer but still temporary effect that lasts for two or three pictures, as well as a sustained effect that lasts for the whole test length. These studies showed that a sustained effect fit best to the PSE and OMT datasets that were studied. Lang (2014) speculates that the recovery of implicit motives may take more time than originally anticipated and may happen after respondents finish working on a picture set.

Using DTM for DIF analyses

The advantage of using an IRT framework for studying motive responses is that IRT allows for advanced DIF analysis. The fundamental difference between the previously applied DIF detection procedures and IRT-based DIF detection procedures is that IRT models estimate person and picture effects separately. IRT based techniques are commonly considered to be less prone to confound real mean differences with bias (Lord, 1977) and researchers have suggested that IRT methods are "the most generally valid of all biased item detection methods" (Osterlind, 1987, p. 69). Analyzing DIF in implicit motive measures using DTM offers various additional advantages over non-IRT methods.

First, the DTM estimates latent traits based on two underlying response processes (motive competition and dynamic reduction of motive strength) and, thus, allows for a more precise parameter estimation both for person as well as for picture parameters. Both previous approaches build groups of low-medium-high motive levels (score level) and depend on sum scores. As a consequence, those test only really function well when most pictures are unbiased (Hofer et al., 2005). On the contrary, IRT based DIF analysis is able to perform DIF analysis even when only one item is unbiased (Thissen et al., 1986).

Second, IRT DIF analyses are parametric and therefore allow to compare more than two groups simultaneously, which is especially interesting in studies in more than two cultures. The IRT approach shares this feature with the previously applied loglinear approach. The Mantel–Haenszel procedure can only compare two groups and is, thus, less suitable for cross-cultural research based on three or more cultures.

Third, in case that DIF is found, it is possible to build a partial invariance model that estimates picture parameters of the DIF categories separately for each group. One does not need to exclude pictures in which DIF was found, thus, more information is included in the trait estimation.

DIF Analysis Using Likelihood-Ratio Tests

A common strategy to analyze items for DIF in IRT uses likelihood-ratio tests to compare nested models (Thissen et al., 1986). The typical procedure is to compare a basic IRT model in which item parameter estimation is constrained to be equal for all groups with a DIF model that allows the item parameter for the focused item to be estimated separately for the groups of interest (De Boeck et al., 2011; Thissen et al., 1986). If the DIF model improves model fit, the item is flagged as DIF item.

In this strategy, one assumes that all other items in the DIF model are unbiased, which is in most cases unknown. Because a set of unbiased items is typically unknown, one starts the model comparisons for the first item with the basic IRT model. If the first item shows DIF, all following model comparisons estimate this item separately for each group, if the item shows no DIF it is estimated equally for all groups. In this procedure, one does not need to make assumptions for already tested items. We then proceed to test all following items in the same way. We also tested a more complex iterative strategy to establish a set of unbiased items, results were highly similar.[5] The DTM is a Rasch type model with a fixed *a* parameter. Thus,

[5] The iterative procedure to establish a set of unbiased items is performed in three steps (Edelen, Thissen, Teresi, Kleinman, & Ocepek-Welikson, 2006). In a first step, one identifies a set of temporary anchor items. Therefore, one tests every item for DIF under the assumption that all other items are non-DIF. All items that do not show DIF in the first step are then used as potential anchor items, the other items are considered as potential DIF items. In the second step, one establishes the final anchor set of unbiased items. Therefore, one removes all potential DIF items and retests all potential anchor items, under the assumption that the other potential anchor items are unbiased. This step is repeated until no item is identified as DIF item. The remaining items can then be used as final anchor set. In the final step, each of the potential DIF items identified in step one and two is retested relative to the final set of anchor items. Only items that show DIF in the last steps are considered as DIF items.

only uniform DIF can be tested. However, non-uniform DIF is relatively rare (Dorans & Holland, 1992; Van de Vijver & Leung, 1997).

When DIF has been identified, one can either remove those pictures from the test or specify a partial invariance model. In the partial invariance model, all picture parameters were DIF has been found are estimated separately for each group and all non-DIF parameters are estimated jointly. This procedure is typically used in SEM modeling (Reise, Widaman, & Pugh, 1993) and has the advantage that information from the DIF pictures is not lost.

The Present Research

We analyze a dataset of an OMT picture set that has been developed and applied for implicit motive research in three different cultures: Cameroon, Germany, and Costa Rica (Chasiotis et al., 2006). We were interested in differences in latent motives between those cultures and OMT pictures that potentially show DIF. We analyze the OMT picture set for DIF in three cultures using the steps described above and determine unbiased latent mean differences using a partial invariance model.

Method

Selection of Cultures

A sample needs to show a high cultural diversity to test universal assumptions. The cultures in this study show high cultural diversity, being recruited from Africa (Cameroon), Europe (Germany), and Latin-America (Costa Rica). The cultures in this study differ in socio-economical background, socio-cultural norms, values, and orientations (Hofstede, 2001). The Costa Rican and German sample is relatively representative for the nation, but Cameroon is a multi-ethnic nation. To control for cultural differences among African participants, the Cameroonian sample is restricted to Nso, which is a large ethnic group in the Anglophone North-West province of Cameroon (Yovsi, 2003).

Participants

The sample of 369 OMTs is part of a research project that aimed at measuring implicit motives in three different cultures (Chasiotis et al., 2006; Hofer et al., 2005) and comprises 180 women and 189 men with an age range from 18 to 75 (M=36,42, SD=14,25) years. The dataset includes participants from Cameroon (n=125), Germany (n=124), and Costa Rica

(n=120). Participants from Costa Rica and Cameroon have been interviewed at their homes, whereas participants from Germany were interviewed at the University of Osnabrück. The sample is balanced for gender, socioeconomic status, and rural versus urban regions within each of the countries. The OMT data from this research project has previously been published in Busch, Hofer, Chasiotis, and Campos (2013), in Chasiotis, Bender, and Hofer (2014), in Chasiotis, Hofer, and Campos (2006), and in Hofer, Busch, Chasiotis, Kärtner, and Campos (2008).

The Operant Motive Test

The OMT is a recent measure of implicit motives that is based on more traditional picture based implicit motive test such as PSE and TAT (e.g. Smith, 1992). In the OMT assessment procedure, researchers present a series of pictures depicting social scenes with one or more persons. Respondents are asked to choose one of the persons as main character of their story. The OMT only asks for brief answers to four questions: (1) "What is important for the person in this situation and what is the person doing?" (2) "How does the person feel?" (3) "Why does the person feel this way?" (4) "How does the story end?". A distinctive feature is that the OMT uses drawings as picture cues. The drawings are less detailed than photographs and leave certain characteristics such as gender, clothing, and ethnic group open. Neutral drawings reduce the probability that respondents choose a character based on cultural characteristics, which could lead to DIF.

The responses are categorized into one of the three motives, a story with no motive content is coded as zero. The OMT coding system additionally differentiates between five categories within each motive: three approach, one avoidance and one in between category (Kuhl & Scheffer, 2001). Researchers studied the differentiation between approach and avoidance motivation since the early years of research on implicit motives, especially for achievement motivation (McClelland and Liberman, 1949). Different researchers have developed procedures to measure hope of success and fear of failure (Birney, Burdick, & Teevan, 1969; McClelland et al., 1953; Pang, 2006). Veroff and Veroff (1972) suggested that the power motive divides into a hope of power and a fear of weakness component and Boyatzis (1973) suggested that the affiliation motive divides into hope for closeness and fear

of rejection. A differentiated review is provided by Schultheiss (2010). The OMT builds on this research and differentiates between approach and avoidance motivation.

To test whether the theoretically based differentiation into approach and avoidance components applies to this data set, we correlated latent approach and avoidance motivations for each motive. A high correlation would indicate that a differentiation between approach and avoidance motivation is not justified for the OMT. No correlation or a negative correlation would indicate that the approach and avoidance components are measuring different aspects. We found negative correlations between approach and avoidance motivation for affiliation (r = -.35), achievement (r = -.37), and power (r = -.15) motivation when approach motivation included level one to three. This finding gives evidence that separating approach and avoidance motivation is reasonable for this OMT data set. We used a strict definition of approach motivation and included OMT levels one to three into an approach factor, because this model showed higher parameter recovery compared to a broader approach factor that includes level one to four (see below for IRT reliability estimation).[6]

The OMTs were filled out in three different languages: English in Cameroon, German in Germany and Spanish in Costa Rica. Four coders at the University of Osnabrück that speak both English and German evaluated the OMTs from Cameroon and Germany. Five Spanish speaking coders from the University of Costa Rica evaluated the OMTs from Costa Rica. All coders used the detailed information provided in the OMT manual (Kuhl & Scheffer, 2001)—in the Spanish translation for coders from Costa Rica—and were instructed and trained through OMT coding seminars. All coders reached an agreement of Cohen's $\kappa > .85$ on training material prior to the coding of the study. The rater agreement in the study has been evaluated on 20 double coded and translated OMTs from Costa Rica. Cohen's κ for the 20 translated stories from Costa Rica between the raters from Costa Rica and Germany was .86 for affiliation, .81 for achievement, and .84 for power (Chasiotis & Hofer, 2003). Because German and Costa Rican coders reached a sufficient agreement, all following OMTs have

[6] All analysis can also be applied to the less strict definition of approach using OMT levels one to four that also correlate negative with avoidance motivation (affiliation r = -.30, achievement r = -.47, and power r = -.09).

been single coded and randomly distributed between all coders. A detailed description of the coding procedure in this research project is described by Chasiotis and Hofer (2003).

We estimated reliability using an IRT simulation approach that describes how well trait estimates are recovered in 100 simulated and bias corrected simulation runs. A description of this standard procedure for reliability estimation in DTM is provided by Lang (2014). The reliability estimates were $r^2_{\theta\hat{\theta}} = .63$ for affiliation, $r^2_{\theta\hat{\theta}} = .50$ for achievement, and $r^2_{\theta\hat{\theta}} = .69$ for power. The reliabilities for affiliation and power are similar to OMT IRT reliabilities in the literature (Lang et al., 2012; Runge et al., 2016). For research purposes, reliabilities in this range are typically sufficient (Ellis, 2013). The reliability for achievement is lower compared to previous OMT studies.

Analyses

Dynamic Thurstonian IRT model for the OMT. Researchers have recently applied modern Thurstonian IRT models to implicit motive tests and modelled motive competition as choice behavior (Lang et al., 2012) and added the dynamic reduction to the model (Lang, 2014; Runge et al., 2016). In the following section, we describe how the Thurstonian IRT model can be used to analyze OMT data using general linear mixed-effects models.

Motive competition is modelled as pairwise motive comparisons on level 1. Let $\pi_{affachs}$, $\pi_{affpows}$, $\pi_{affzeros}$, $\pi_{achpows}$, $\pi_{achzeros}$, and $\pi_{powzeros}$ denote the probability that person *s* prefers one motive category (aff, ach, pow, and zero for other content) to another. These probabilities are a function of difference between the latent utilities for each motive category. The latent utilities for each motive can be denoted as μ_{affs}, μ_{achs}, μ_{pows}, and μ_{zeros}. They are linked with a probit link to the binomial outcome. The Level 1 probit model can be written as:

$$\begin{pmatrix} probit\left(\pi_{affachs}\right) \\ probit\left(\pi_{affpows}\right) \\ probit\left(\pi_{affzeros}\right) \\ probit\left(\pi_{achpows}\right) \\ probit\left(\pi_{achzeros}\right) \\ probit\left(\pi_{powzeros}\right) \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{affs} \\ \mu_{achs} \\ \mu_{pows} \\ 0 \end{pmatrix} = \boldsymbol{D}\mu_s.$$

In this denotation, μ_s is a vector that contains the latent utilities and $\boldsymbol{D}$ is the design matrix that contains three dummy variables that refer to three of the four motive categories. The fourth row of dummy variables can be omitted without losing generality, the equivalent row of mean evaluations is constrained to zero.

In level 2 of the linear mixed-effects model, the latent utilities μ_s are specified by adding a person parameter, a picture parameter and additionally a dynamic parameter that accounts for the dynamic motive reduction. The complete level 2 specifications of the DTM model can be written as follows:

$$\mu_{affs} = \sum_{m=1}^{M} x_{affms}\,\beta_{affm} + DE_{affs}\beta_{aff} + \nu_{affs}$$

$$\mu_{achs} = \sum_{m=1}^{M} x_{achms}\,\beta_{achm} + DE_{achs}\beta_{ach} + \nu_{achs}$$

$$\mu_{pows} = \sum_{m=1}^{M} x_{powms}\,\beta_{powm} + DE_{pows}\beta_{pow} + \nu_{pows}$$

$$\mu_{zeros} = 0.$$

In this specification, ν_s is a random person effect (theta) for each motive that captures the latent motive of the person. x_{ms} $(m = 1, \dots, M)$ is a picture covariate with *M* being the total picture number and a fixed effect β_m that denotes the mean evaluation of picture *m* for the corresponding motive category (in contrast to standard IRT β_m denotes easiness and not difficulty). For each picture, the DTM includes up to three motivational categories (affiliation, achievement, and power) and a zero category that includes all non-motive relevant responses including all avoidance categories. These response options for each picture are captured by the fixed effect β_m and are conceptually the same as items in other (Thurstonian) IRT models. When we refer to items in the context of this article, we thus refer to the endorsement of a specific motivational response category for a specific picture in the context of the DTM. DE denotes the dynamic effect and is the sum of all previous responses of the respective motive. A detailed model description for the OMT is provided by Runge et al. (2016).

Model estimation. DTM models can be fitted in the lme4 1.1-12 package using the Bayesian glmer function (Bates, Mächler, Bolker, & Walker, 2015) in the R environment (R

Core Team, 2015). Some of the pictures showed almost no motive responses for one or two motives. Because information in responses with very low response probabilities is limited, it is difficult for generalized linear mixed-effects models (GLMM) like the DTM to estimate effects. The standard approach is to exclude the item (e.g. Debeer & Janssen, 2013). It is reasonable to assume that some pictures in the OMT do not invoke a certain motive. Thus, we excluded items that show very few responses from model estimation. In detail, items with a response frequency lower than 3% for specific motives in seven cases (pictures 1-5 and 12, see Table 1) were eliminated before model estimation. The glmer function uses the Laplace approximation (Böckenholt, 2001) and estimates random effects with the maximum a posteriori method (De Boeck et al., 2011). We used the BOBYQA algorithm in NLOPT as optimizer for model estimation. We rerun the analysis with other optimizers, results were highly similar. A detailed description of these procedures including an R code tutorial is provided by Runge et al. (2016). The basic R code for a DIF model that tests DIF for affiliation in picture 1 is provided in a footnote.[7]

Results

Table 1 shows the overall response frequencies for each motive. Table 2 provides the result of the DIF analyses. As indicators of model fit we report the Bayesian Information Criterion (BIC) as it is common in the literature on generalized linear mixed-effects models (Bates et al., 2015; Doran, Bates, Bliese, & Dowling, 2007). The BIC is an index that is smaller for models that provides better fit to the data and penalizes complex models. With a high sample size, the BIC is consistent in selecting the true model (Vrieze, 2012). Picture

[7] The following R code fits this model:

```
glmer(response~0+
AFF:PIC1:CAMEROON+AFF:PIC1:GERMANY+AFF:PIC1:COSTARICA+
AFF:PIC2+AFF:PIC3+AFF:PIC6+AFF:PIC7+AFF:PIC8+AFF:PIC9+AFF:PIC10+AFF:PIC11+
ACH:PIC4+ACH:PIC5+ACH:PIC6+ACH:PIC7+ACH:PIC8+ACH:PIC9+ACH:PIC10+ACH:PIC11+A
CH:PIC12+
POW:PIC1+POW:PIC2+POW:PIC3+POW:PIC5+POW:PIC6+POW:PIC7+POW:PIC8+POW:PIC9+POW
:PIC10
+POW:PIC11+POW:PIC12+
AFF:DEAFF+ACH:DEACH+POW:DEPOW+
(AFF+ACH+POW+0|person)
,family=binomial("probit"),tdata,control=glmerControl(optimizer="nloptwrap"
,calc.derivs=F))
```

1/affiliation (BIC change: 15.6), picture 2/affiliation (BIC change: 8.6), and Picture 6/power (BIC change: 3.5) were identified as DIF items. The first DIF item was picture 1/affiliation that showed a much lower item parameter for Germany compared to Cameroon and Costa Rica (see Table 3). The picture shows two people embracing each other. The second DIF item was picture 2/affiliation that also showed a much lower item parameter for Germany compared to Cameroon and Costa Rica. This picture shows two people that speak behind the back of another person. The third DIF item was picture 6/power that had a higher item parameter for Costa Rica compared to Cameroon and Germany. This picture shows two people drawing circles.

To evaluate latent mean differences across the tested cultures, we analyzed a partial invariance model with item parameters estimated separately for each culture for the three identified DIF items. In this model, we allowed latent mean differences in implicit motives to vary between the cultures. The parameter estimates of the final model are presented in Table 3. Higher values for a picture parameter indicate a higher picture pull. The dynamic effects are negative indicating that a response for a certain motive is less likely for each previous response to this motive. This finding is in line with previous studies (Lang, 2014; Runge et al., 2016).[8] The information from a DTM can be directly translated to predictions for responses for a particular picture. For instance, consider a hypothetical scenario in which a person with one previous affiliation response (a dynamic effect DE of 1 for affiliation), but no previous achievement responses (a dynamic effect DE of 0 for achievement) and true latent traits of theta = 2 and theta = 1 for affiliation and achievement, respectively, would answer picture 7. The predicted probit that this person chooses affiliation over achievement in this picture is $\pi_{affachs} = (x_{aff7s}\beta_{aff7} + DE_{aff}\beta_{aff} + \nu_{affs}) - (x_{ach7s}\beta_{ach7} + DE_{ach}\beta_{ach} + \nu_{achs}) = (-1.29 + (1 \times (-1.75)) + 2) - (-0.57 + (0 \times (-0.79)) + 1) = -1.61$. The probit link in GLMM follows the (inverse) cumulative standard normal distribution. Thus, probits can be

[8] We have also tested a model using a temporary reduction in motive strength that is similar to the conceptualization of motive consumption in DoA. We found that, consistent with the literature (Lang, 2014; Runge et al. 2016), the model with temporary reduction in motive strength (BIC = 10241.9) did not improve model fit over the model with sustained reduction in motive strength (BIC = 9778.4).

translated to a probability that the person chooses affiliation over achievement using a normal distribution table or a statistics software. -1.61 is equal to a probability of 5%.

Table 4 presents the latent mean differences of implicit motives for each culture from the partial invariance model as well as the original model without any DIF analysis. The values indicate the difference between the cultures in the average response probabilities for each motive. For example, German respondents were least likely and people from Costa Rica were most likely to respond with an affiliation story. People from Germany have a lower latent affiliation motivation than people from Cameroon and people from Cameroon have a lower latent affiliation motivation than people from Costa Rica. Considerable differences between the normal and the DIF corrected model were found for affiliation motivation in the German part of the sample. The difference reflects the corrected item easiness parameters in the two identified DIF items for affiliation.

Discussion

In this study, we showed how the DTM model can be used to study implicit motives across cultures and identify DIF items in implicit motive measures. We demonstrated the DIF model using a picture set of the OMT. Three items were flagged as DIF. Parameter estimates for these items showed that the first two pictures elicited less affiliation responses in the German subsample compared to the other countries and picture six elicited more power responses for Costa Rica compared to the other countries. We built a partial invariance model in which item parameters were estimated separately for each country for the DIF items. Latent mean differences showed that implicit motives differ between the countries. A comparison with the latent mean differences obtained from the standard model showed that the partial invariance model corrected for the DIF items. However, the rank order for latent mean implicit motive scores between the countries did not change and the DIF correction was not large.

The first implication of the study is that this picture set of the OMT is useful for the comparison of implicit motives in the three tested cultures. Although three items were identified as DIF, pictures in this OMT set were mostly unbiased. This picture set can be used for future studies if researchers correct for the identified DIF items, or exclude those pictures from model estimations. On a more general level, this study emphasizes earlier research on

implicit motive tests that demonstrate that implicit motive tests can be used to compare implicit motives universally across cultures (Hofer et al., 2005).

The second implication of the study is that implicit motives vary between people from different countries with different cultural backgrounds. Our findings suggest, that people from Germany had the lowest motive scores among the three groups for all motives. People from Costa Rica had higher affiliation scores than people from Cameroon. This finding is in line with previous research that found differences in implicit motives between people with various cultural backgrounds (Busch et al., 2013; Hofer et al., 2008; van Emmerik et al., 2010) and, thus, indicates the existence of culture specific variations in implicit motives. Germany having a lower power motivation compared to Costa Rica and Cameroon has previously been reported (Chasiotis et al., 2006). More recent studies (Aydinli, Bender, Chasiotis, van de Vijver, & Cemalcilar, 2015; Chasiotis et al., 2014) replicated the findings—including more than the three countries analyzed in this study—that two factors explain the cultural variance in the development of implicit power motivation: the number of siblings and parental socio-economic status. Participants with more siblings have a developmental context in which they are exposed to more power motivated behavior such as caretaking compared to participants with less or no siblings. Busch et al. (2013) found, focusing on the flow aspect of achievement, that Germany scored lower, compared to Costa Rica and Cameroon. Achievement motivation can be seen as a materialistic value which is less important in postindustrial cultures like Germany, compared to self-expression and subjective well-being (Inglehart & Baker, 2000). For affiliation motivation, a previous study showed that Costa Rica has a higher need for affiliation compared to Germany, however this study did not find a lower affiliation motivation for Germany compared to Cameroon (Hofer et al., 2006). Overall, our findings are largely in line with previous studies.

It is difficult to interpret the DIF items because the DIF test does not reveal the source of DIF. One possible explanation for DIF in this study is a systematic bias for coders between Costa Rica and Germany. However, such a bias could only apply to the DIF found in picture 6, which was between Costa Rica and the other cultures. Although a coder bias could be responsible for the differences in picture 6, we cannot test this interpretation because the cultures and coders are confounded. None of the other pictures showed DIF between the

coders so a general coder bias is unlikely. Another potential post-hoc explanation for the DIF in pictures 1 and 2 is that Germany is an individualistic country, whereas Cameroon and Costa Rica are more collectivistic countries (Hofstede, 2001). It may be that pictures 1 and 2 are interpreted as more individualistic in Germany compared to Cameroon and Costa Rica, independent of individual motive strength.

The third implication of the study relates to the scope of the presented DIF detection procedure. This study demonstrated the use of IRT methods in studying DIF for cultures in the OMT. The usefulness of the presented method is, however, not limited to the OMT or cultural research. Another commonly used implicit motive measure is the PSE. Because DTM was developed for the PSE (Lang, 2014) it is straightforward to adapt the presented DIF detection methods for the use in PSE research. Researchers have also been interested in gender (Drescher & Schultheiss, 2016) and age differences (Denzinger, Backes, Job, & Brandstätter, 2016) in implicit motives. The suggested method can be used to study DIF for those comparisons as well.

Strength and Limitations

A limitation of our demonstration using the OMT is that the OMT is an implicit motive test that is less frequently used than the PSE. Although the OMT likely measures implicit motives, the implicit motive definition of the OMT is not necessarily highly correlated with the motive definition in the Winter coding manual (Winter, 1994). One study found that only power motivation correlated significantly, but lower (r = .15), between OMT and PSE scores. Another study found modest positive correlations of PSE and OMT scores (affiliation r = .31, achievement r = .47, and power r = .47) that were measured three months apart (Scheffer, 2001). A strength of the PSE is that some of the pictures have been developed in the context of laboratory experiments.

Another limitation of this study is that the IRT reliability of the achievement motivation measured with this picture set is relatively low. One potential reason could be that the picture set uses less pictures (12) compared to the standard picture set of the OMT (15) and the first three pictures show almost no achievement responses, thus the effective test length for the achievement motive is reduced to 9 pictures. Researchers recommend a minimum test length of four to five pictures for the PSE (Pang, 2010a; Schultheiss & Pang,

2007). The OMT uses more pictures but responses are brief and coders assign only one motive for a picture. Thus, the PSE recommendation is not one-to-one applicable to the OMT. Reliability is also not a linear function of test length, other factors such as cue ambiguity, variance, and dynamic effects can play a role. Thus, it is difficult to give a general advice on minimum picture number in the OMT based on this study. Future studies should systematically evaluate the influence of test length on reliability and validity. Researchers have noted that implicit motive measures with a high number of pictures lead to a decrease in reliability and validity as a result of fatigue (Reitman & Atkinson, 1958) or motive consumption (Lang, 2014).

Another limitation is that other response processes that have been suggested are not modeled with the DTM. For instance, Schultheiss, Liening, and Schad (2008) suggested that situations which arouse a specific motive response may differ for individuals, based on differential learning histories. Another process that is not modelled in this study is joint motive expression. A response to a picture may be influenced by multiple motives. For instance, helping behavior can satisfy the need for power as well as the need for affiliation. It is generally possible to model several codings in a story using the DTM (see Lang, 2014), but in the OMT the response questions are designed for brief answers and coders are instructed to only code one motive in each picture on the basis of the idea that most responses are dominated by one motive theme (also see Tuerlinckx et al., 2003). In future studies, researchers could modify the response and coding format of the OMT and model multiple motive expressions in the OMT using DTM. Statisticians have argued that all models are wrong, but some models are useful (Box, 1976). The DTM is a useful model to test response processes in implicit motive measures. Future research may continue this work and model other response processes.

Conclusion

In this study, we showed how the DTM can be used to identify item bias in implicit motive measures and suggested the use of a partial invariance model that corrects for bias, to study implicit motives across culture. We demonstrated the procedure on an OMT data set with participants from Cameroon, Costa Rica, and Germany. This study further contributes to the emerging IRT literature on implicit motives by displaying an additional application.

References

Atkinson, J. W., & Birch, D. (1970). *Dynamics of Action*. New York: Wiley.

Aydinli, A., Bender, M., Chasiotis, A., van de Vijver, F. J. R., & Cemalcilar, Z. (2015). Implicit and explicit prosocial motivation as antecedents of volunteering: The moderating role of parenthood. *Personality and Individual Differences*, *74*, 127–132. doi:10.1016/j.paid.2014.10.011

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *1*. doi:10.18637/jss.v067.i01

Birney, R. C., Burdick, H., & Teevan, R. C. (1969). *Fear of failure*. New York: Van Nostrand-Reinhold Co.

Böckenholt, U. (2001). Hierarchical Modeling of Paired Comparison Data. *Psychological Methods*, *6*, 49–66. doi:10.1037//1082-989X.6.1.49

Borsboom, D. (2006). The attack of the psychometricans. *Psychometrika*, *71*, 425–440. doi:10.1007/s11336-006-1447-6

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*, 203–219. doi:10.1037/0033-295X.110.2.203

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The Concept of Validity. *Psychological Review*, *111*, 1061–1071. doi:10.1037/0033-295X.111.4.1061

Box, G. E. P. (1976). Science and Statistics. *Journal of the American Statistical Association*, *71*, 791–799. doi:10.1080/01621459.1976.10480949

Boyatzis, R. E. (1973). Affiliation motivation. In D. C. McClelland & R. S. Steele (Eds.), *Human motivation -- a book of readings* (pp. 252–276). Morristown, NJ: General Learning Corporation.

Brown, A., & Maydeu-Olivares, A. (2013). How IRT can solve problems of ipsative data in forced-choice questionnaires. *Psychological Methods*, *18*, 36–52. doi:10.1037/a0030641

Busch, H., Hofer, J., Chasiotis, A., & Campos, D. (2013). The achievement flow motive as an element of the autotelic personality: predicting educational attainment in three cultures. *European Journal of Psychology of Education*, *28*, 239–254. doi:10.1007/s10212-012-0112-y

Chasiotis, A., Bender, M., & Hofer, J. (2014). Childhood context explains cultural variance in implicit parenting motivation: Results from two studies with six samples from Cameroon, Costa Rica, Germany, and PR China. *Evolutionary Psychology*, *12*, 295–317. doi:10.1177/147470491401200203

Chasiotis, A., & Hofer, J. (2003). *Abschlußbericht zum DFG-Projekt „Die Messung impliziter Motive in Deutschland, Costa Rica und Kamerun – Methodische Voruntersuchungen zu einer kulturinformierten Motivationspsychologie [Final Report of the DFG Project "The Measurement of implicit Motives in Germany, Costa Rica, and Cameroon - Methodological Preliminary Investigation for a Culturally Informed Motivational Psychology]*. Osnabrück, Germany.

Chasiotis, A., Hofer, J., & Campos, D. (2006). When does Liking Children Lead to Parenthood? Younger Siblings, Implicit Prosocial Power Motivation, and Explicit Love for Children Predict Parenthood across Cultures. *Journal of Cultural and Evolutionary Psychology*, *4*, 95–123. doi:10.1556/JCEP.4.2006.2.2

De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., Tuerlinckx, F., & Partchev, I. (2011). The Estimation of Item Response Models with the lmer Function from the lme4 Package in R. *Journal Of Statistical Software*, *39*, 1–28. doi:10.18637/jss.v039.i12

Debeer, D., & Janssen, R. (2013). Modeling Item-Position Effects Within an IRT Framework. *Journal of Educational Measurement*, *50*, 164–185.

Denzinger, F., Backes, S., Job, V., & Brandstätter, V. (2016). Age and Gender Differences in Implicit Motives. *Journal of Research in Personality*. doi:10.1016/j.jrp.2016.09.003

Doran, H., Bates, D., Bliese, P., & Dowling, M. (2007). Estimating the Multilevel Rasch Model : With the lme4 package. *Journal of Statistical Software*, *20*, 1–18. doi:10.1111/j.1467-9868.2007.00600.x

Dorans, N. J., & Holland, P. W. (1992). DIF Detection and Description: Mantel–Haenszel and Standardization. *ETS Research Report Series*, *1992*.

Drescher, A., & Schultheiss, O. C. (2016). Meta-analytic evidence for higher implicit affiliation and intimacy motivation scores in women, compared to men. *Journal of Research in Personality*, *64*, 1–10. doi:10.1016/j.jrp.2016.06.019

Edelen, M. O., Thissen, D., Teresi, J. A., Kleinman, M., & Ocepek-Welikson, K. (2006).

Identification of differential item functioning using item response theory and the likelihood-based model comparison approach. Application to the Mini-Mental State Examination. *Medical Care*, *44*. doi:10.1097/01.mlr.0000245251.83359.8c

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, *5*, 155–174. doi:10.1037/1082-989X.5.2.155

Ellis, J. L. (2013). A standard for test reliability in group research. *Behavior Research Methods*, *45*, 16–24. doi:10.3758/s13428-012-0223-z

Embretson, S. E. (1994). Applications of cognitive design systems to test development. In C. R. Reynolds (Ed.), *Cognitive assessment: A multi- disciplinary perspective.* (pp. 107–135). New York: Plenum Press.

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and Individual Differences: A Natural Science Approach*. New York: Plenum Press.

Glymour, C. N. (2001). *The mind's arrows : Bayes nets and graphical causal models in psychology*. MIT Press.

Hofer, J., & Bond, M. H. (2008). Do implicit motives add to our understanding of psychological and behavioral outcomes within and across cultures? In R. M. Sorrentino & S. Yamaguchi (Eds.), *The handbook of motivation and cognition within and across cultures* (pp. 95–118). Amsterdam: Elsevier Academic Press.

Hofer, J., Busch, H., Bond, M. H., Kärtner, J., Kiessling, F., & Law, R. (2010). Is Self-Determined Functioning a Universal Prerequisite for Motive-Goal Congruence? Examining the Domain of Achievement in Three Cultures. *Journal of Personality*, *78*, 747–780. doi:10.1111/j.1467-6494.2010.00632.x

Hofer, J., Busch, H., Chasiotis, A., Kärtner, J., & Campos, D. (2008). Concern for Generativity and Its Relation to Implicit Pro-Social Power Motivation, Generative Goals, and Satisfaction With Life: A Cross-Cultural Investigation. *Journal of Personality*, *76*, 1–30. doi:10.1111/j.1467-6494.2007.00478.x

Hofer, J., & Chasiotis, A. (2003). Congruence of life goals and implicit motives as predictors of life satisfaction: Cross-cultural implications of a study of Zambian male adolescents. *Motivation and Emotion*, *27*, 251–272.

Hofer, J., & Chasiotis, A. (2004). Methodological Considerations of Applying a Tat-Type Picture-Story Test in Cross-Cultural Research A Comparison of German and Zambian Adolescents. *Journal of Cross-Cultural Psychology*, *35*, 224–241. doi:10.1177/0022022103262246

Hofer, J., & Chasiotis, A. (2011). Implicit motives across cultures. *Online Readings in Psychology and Culture*, *4*, 5.

Hofer, J., Chasiotis, A., & Campos, D. (2006). Congruence between social values and implicit motives: effects on life satisfaction across three cultures. *European Journal of Personality*, *20*, 305–324. doi:10.1002/per.590

Hofer, J., Chasiotis, A., Friedlmeier, W., Busch, H., & Campos, D. (2005). The measurement of implicit motives in three cultures power and affiliation in Cameroon, Costa Rica, and Germany. *Journal of Cross-Cultural Psychology*, *36*, 689–716. doi:10.1177/0022022105280510

Hofer, J., Kärtner, J., Chasiotis, A., Busch, H., & Kiessling, F. (2007). Socio-cultural Aspects of Identity Formation: The Relationship between Commitment and Well-Being in Student Samples from Cameroon and Germany. *Identity: An International Journal of Theory and Research*, *7*, 265–288. doi:10.1080/15283480701600744

Hofstede, G. H. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations across Cultures*. Thousand Oakes, CA: Sage.

Inglehart, R., & Baker, W. E. (2000). Modernization, Cultural Change, and the Persistence of Traditional Values. *American Sociological Review*, *65*, 19. doi:10.2307/2657288

Kuhl, J., & Scheffer, D. (2001). Der operante Multi-Motiv-Test (OMT): Manual [The operant multi-motive-test (OMT): Manual]. Osnabrück, Germany: University of Osnabrück.

Lang, J. W. B. (2014). A dynamic Thurstonian item response theory of motive expression in the picture story exercise: Solving the internal consistency paradox of the PSE. *Psychological Review*, *121*, 481–500. doi:10.1037/a0037011

Lang, J. W. B., Zettler, I., Ewen, C., & Hülsheger, U. R. (2012). Implicit motives, explicit traits, and task and contextual performance at work. *Journal of Applied Psychology*, *97*, 1201–1217. doi:10.1037/a0029556

Lord, F. M. (1977). A study of item bias using item characteristic curve theory. In Y. H.

Poortinga (Ed.), *Bias problems in cross-cultural psychology.* (pp. 19–29). Amsterdam: Swets and Zeitlinger.

McClelland, D. C. (1961). *The achieving society.* Princeton N.J.: Van Nostrand.

McClelland, D. C. (1984). *Motives, Personality, and Society: Selected Papers.* New York, NY, NY: Praeger Publishers.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). The measuring instrument. In *The achievement motive* (pp. 185–217). East Norwalk, CT, US: Appleton-Century-Crofts. doi:10.1037/11144-007

McClelland, D. C., & Liberman, A. M. (1949). The Effect of Need for Achievement on Need-Related Words. *Journal of Personality*, *18*, 236–251. doi:10.1111/j.1467-6494.1949.tb01243.x

Osterlind, S. J. (1987). *Test item bias: Quantitative applications in social sciences.* Beverly Hills, CA: Sage Publications.

Pang, J. S. (2006). *A revised content-coding measure for hope of success(HS) and fear of failure (FF). Unpublished Dissertation.* University of Michigan, Ann Arbor.

Pang, J. S. (2010). Content coding methods in implicit motive assessment: Standards of measurement and best practices for the picture story exercise. In O. C. Schultheiss & J. C. Brunstein (Eds.), *Implicit Motives* (pp. 119–1550). New York: Oxford University Press.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge, England: Cambridge University Press.

R Core Team. (2015). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, *v3.0.1*, 3604. doi:10.1007/978-3-540-74686-7

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance. *Psychological Bulletin*, *114*, 552–566. doi:10.1037/0033-2909.114.3.552

Reitman, W. R., & Atkinson, J. W. (1958). Some methodological problems in the use of thematic apperceptive measures of human motives. In J. W. Atkinson (Ed.), *Motives in fantasy, action, and society* (pp. 664–683). Princeton, NJ: Van Nostrand.

Revelle, W., & Condon, D. M. (2015). A model for personality at three levels. *Journal of*

Research in Personality, *56*, 70–81. doi:10.1016/j.jrp.2014.12.006

Runge, J. M., Lang, J. W. B., Engeser, S., Schüler, J., den Hartog, S. C., & Zettler, I. (2016). Modeling Motive Activation in the Operant Motives Test: A Psychometric Analysis Using Dynamic Thurstonian Item Response Theory. *Motivation Science*, *2*, 268–286. doi:10.1037/mot0000041

Schultheiss, O. C. (2010). Implicit motives. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of Personality: Theory and Research* (pp. 603–633). New York: Guilford Press.

Schultheiss, O. C., & Brunstein, J. (2010). Implicit Motives. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 603–633). New York: Oxford University Press. doi:10.1093/acprof:oso/9780195335156.001.0001

Schultheiss, O. C., Liening, S. H., & Schad, D. (2008). The reliability of a Picture Story Exercise measure of implicit motives: Estimates of internal consistency, retest reliability, and ipsative stability. *Journal of Research in Personality*, *42*, 1560–1571. doi:10.1016/j.jrp.2008.07.008

Schultheiss, O. C., & Pang, J. S. (2007). Measuring implicit motives. In R. W. Robins, R. C. Fraley, & R. Krueger (Eds.), *Handbook of Research Methods in Personality Psychology* (pp. 322–344). New York: Guilford.

Smith, C. P. (1992). *Motivation and personality: Handbook of thematic content analysis*. New York: Cambridge University Press.

Thissen, D., Steinberg, L., & Gerrard, M. (1986). Beyond group-mean differences: The concept of item bias. *Psychological Bulletin*, *99*, 118. doi:10.1037/a0027127

Tuerlinckx, F., De Boeck, P., & Lens, W. (2002). Measuring needs with the Thematic Apperception Test: A psychometric study. *Journal of Personality and Social Psychology*, *82*, 448–461. doi:10.1037//0022-3514.82.3.448

Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research* (Vol. 1). Sage.

van Emmerik, H., Gardner, W. L., Wendt, H., & Fischer, D. (2010). Associations of Culture and Personality With McClelland's Motives: A Cross-Cultural Study of Managers in 24

Countries. *Group & Organization Management*, *35*, 329–367. doi:10.1177/1059601110370782

Verhelst, N. D., & Glas, C. A. W. (1993). A dynamic generalization of the Rasch model. *Psychometrika*, *58*, 395–414. doi:10.1007/BF02294648

Veroff, J., & Veroff, J. B. (1972). Reconsideration of a measure of power motivation. *Psychological Bulletin*, *78*, 279–291. doi:10.1037/h0033184

Vrieze, S. I. (2012). Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*, *17*, 228. doi:10.1037/a0027127

Winter, D. G. (1994). *Manual for scoring motive imagery in running text*. Ann Arbor: University of Michigan.

Yovsi, R. D. (2003). *Ethnotheories about Breastfeeding and Mother-infant Interaction*. New Jersey: Transaction Publishers.

Tables

Table 1

Frequency of motive responses for each picture

Picture	Aff %	Ach %	Pow %	Other %
1. Embracing couple	44.7	0.3	23.6	31.4
2. Talking behind the back	7.0	2.4	23.8	66.7
3. Giant in a group	6.0	2.4	44.2	47.4
4. Laying stones	0.0	37.7	0.8	61.5
5. People behind a desk	0.5	7.0	23.8	68.6
6. Drawing circles	1.6	39.8	10.8	47.7
7. Opposite with folded arms	5.1	9.2	36.0	49.6
8. Big and small person	3.5	12.2	10.8	73.4
9. Women sitting behind a man	6.2	3.5	12.8	79.4
10. Group of people	12.7	15.7	27.4	44.2
11. Women and man talking	23.6	3.3	24.1	49.1
12. Man with a fist	0.5	15.7	18.4	65.3

Table 2

Change in BIC *in the sequential DIF testing*

Model	Aff	Ach	Pow
Picture 1: Embracing couple	**15.6**	-	-9.5
Picture 2: Talking behind the back	**8.6**	-	-13.1
Picture 3: Giant in a group	-3.0	-	-4.6
Picture 4: Laying stones	-	-0.7	-
Picture 5: People behind a desk	-	-14.5	-10.9
Picture 6: Drawing circles	-11.2	-11.0	**3.5**
Picture 7: Opposite with folded arms	-18.3	-17.7	-17.4
Picture 8: Big and small person	-19.6	-15.8	-11.3
Picture 9: Women sitting behind a man	-14.6	-16.4	-11.7
Picture 10: Group of people	-15.2	-18.1	-13.3
Picture 11: Women and man talking	-15.4	-14.2	-13.0
Picture 12: Man with a fist	-	-10.7	-18.0

Note. Bold BIC indicates DIF.

Table 3

Parameter Estimates for the partial invariance model

Estimate	Aff		Ach		Pow	
	Fixed effects	*SE*	Fixed effects	*SE*	Fixed effects	*SE*
Picture 1: Embracing couple					-0.10	0.14
Cameroon	0.24	0.27				
Germany	-0.71	0.30				
Costa Rica	0.69	0.29				
Picture 2: Talking behind the back					-0.49	0.14
Cameroon	-1.32	0.29				
Germany	-3.68	0.60				
Costa Rica	-1.89	0.33				
Picture 3: Giant in a group	-1.45	0.27			0.47	0.14
Picture 4: Laying stones			-0.26	0.15		
Picture 5: People behind a desk			-1.55	0.17	-0.03	0.14
Picture 6: Drawing circles	-2.23	0.31	0.37	0.15		
Cameroon					-0.45	0.21
Germany					-0.58	0.22
Costa Rica					0.34	0.19
Picture 7: Opposite with folded arms	-1.29	0.26	-0.57	0.17	0.76	0.15
Picture 8: Big and small person	-1.64	0.27	-0.47	0.18	-0.10	0.17
Picture 9: Women sitting behind a man	-1.13	0.26	-1.21	0.20	-0.09	0.17
Picture 10: Group of people	-0.01	0.25	0.23	0.18	1.03	0.17
Picture 11: Women and man talking	0.86	0.26	-0.74	0.22	1.06	0.18
Picture 12: Man with a fist			0.18	0.20	0.85	0.19
DE	-1.75	0.13	-0.79	0.10	-0.61	0.05
Random effects		2.20		1.30		1.14
Correlations						
Aff	-					
Ach	0.17		-			
Pow	0.03		0.06		-	

Note. k = 10,627 pair-wise comparisons nested in n = 369 persons and 12 pictures. Aff = affiliation vs other. Ach = achievement vs other. Pow = power vs other. DE = dynamic effect operationalized as the number of previous motive related responses. *SE*s are obtained from the summary output in glmer.

Table 4

Differences in Latent Motives by Country

	Latent Mean Differences							
	Standard DTM				Partial Invariance DTM			
Picture	Germany	*SE*	Costa Rica	*SE*	Germany	*SE*	Costa Rica	*SE*
Affiliation	-1.26	0.33	0.60	0.30	-1.06	0.34	0.51	0.32
Achievement	-0.84	0.19	0.18	0.18	-0.85	0.19	0.18	0.18
Power	-0.56	0.16	-0.17	0.16	-0.57	0.16	-0.26	0.16

Note. The latent mean differences are centered around Cameroon that has 0.00 as latent mean motive score for all three motives. *SE*s are obtained from the summary output in glmer.

CHAPTER 4

PREDICTING COUNTERPRODUCTIVE WORK BEHAVIOR: DO IMPLICIT MOTIVES HAVE INCREMENTAL VALIDITY BEYOND EXPLICIT TRAITS?[9]

The present chapter extends previous research on the link between personality and CWB and studies whether the implicit Power and Affiliation motives contribute to the prediction of CWB above personality traits. Employees high in implicit motives may avoid CWB, because potential negative consequences conflict with implicit motives. We report two studies (N = 263 and N = 121) in which respondents filled out the Operant Motive Test, a modern implicit motive measure, and personality trait measures. Results of Study 1 showed that implicit motives—the implicit Affiliation motive in particular—predicted self-rated CWB when personality traits are considered. Study 2 replicated the findings of Study 1 using supervisor ratings of CWBs and found that the implicit Power motive predicted CWB.

[9] This chapter is based on: Runge, J. M. & Lang, J. W. B, Zettler, I., & Lievens, F. (under review). Predicting counterproductive work behavior: Do implicit motives have incremental validity beyond explicit traits?

An important goal of applied psychological research has long been the prediction of counterproductive work behavior (CWB). CWB is commonly defined as voluntary behaviors infringing important social and organizational rules, norms and values (Collins & Griffin, 1998; Robinson & Bennett, 1995; Spector & Fox, 2005). They include acts directed against other individuals in the organization such as violence, gossip, theft from coworkers, as well as acts directed towards the organization directly such as damaging company property or theft, intentionally working slowly, failing to follow instructions, or sharing confidential company information (Berry, Ones, & Sackett, 2007). CWB can be costly to organizations (Marcus, Taylor, Hastings, Sturm, & Weigelt, 2016), for instance, Hollinger and Davis (2002) showed that employee theft costs retailers US$40.7 million every day in the United States. The costs of other CWB (e.g., waste of resources, property damage) is more difficult to estimate, but researchers have calculated damage in the billions of dollars annually (Robinson & Greenberg, 1999).

Prior research typically studied cognitive abilities, demographic characteristics, or other personality characteristics as antecedents of CWB (Salgado, 2002). Personality research on CWB mostly focused on personality traits based on the Five-Factor Model (FFM) and the HEXACO model (Hershcovis et al., 2007; Marcus, Lee, & Ashton, 2007; Salgado, 2002; Spector, 2011; Zettler & Hilbig, 2010). Focusing solely on personality traits, however, may limit the theoretical understanding and the possibilities to predict CWB with personality characteristics because there is considerable evidence that personality does not only consist of explicit personality traits, but also implicit personality dispositions (Brunstein & Maier, 2005; Kehr, 2004; Winter, John, Stewart, Klohnen, & Duncan, 1998). Researchers generally studied various implicit personality constructs, but a large number of studies focused on implicit motives. Implicit motives are described as a predisposition to follow classes of incentives and goals such as Affiliation, Power, and Achievement (Emmons, 1993). Recent research suggests considering implicit motives frequently has incremental validity over considering explicit personality traits alone in predicting important outcomes like job performance and career success (Apers, Lang, & Derous, 2018; Lang, Zettler, Ewen, & Hülsheger, 2012; Winter et al., 1998). Explicit personality traits and implicit motives are theoretically distinct, do not necessarily correlate, and are operationalized with different assessment methods (McClelland,

Koestner, & Weinberger, 1989; Spangler, 1992; Winter et al., 1998). Therefore, both explicit and implicit dispositions may be useful as distinct predictors of CWB. Consequently, focusing only on the role of explicit personality dispositions in CWB and omitting implicit dispositions may lead to lower explanatory and predictive power.

The aim of the present paper is to extend prior research on the link between personality and CWB (Salgado, 2002; Salgado, Moscoso, & Anderson, 2013; Spector, 2011; Zettler, in press) and study whether implicit motives contribute to the prediction of CWB above personality traits. This is, to our knowledge, the first study linking implicit motives to CWB. Building on earlier research, we test the idea that people high in an implicit motive show less CWB because potential negative consequences may conflict with future goal fulfillment.

Explicit Personality and Counterproductive Workplace Behavior

The FFM (McCrae & John, 1992) is the most prevalent taxonomy to study individual differences in personality. Researchers showed in meta-analyses that three of those five traits—Emotional Stability, Conscientiousness, and Agreeableness—are related to different forms of CWB. Salgado (2002) found corrected correlations ranging from -.06 to -.26 between those three traits and absenteeism, accidents, deviant behavior, and turnover. Berry et al. (2007) reported that employees high in Agreeableness show less interpersonal deviance and employees high in Conscientiousness show less organizational deviance. Moreover, researchers have also expanded research on antecedents of CWB to other explicit personality constructs. For instance, Hershcovis et al. (2007) reported in a meta-analysis that trait anger relates to interpersonal and organizational aggression, and Marcus, Lee, and Ashton (2007) showed in a large scale study that Honesty-Humility (from the HEXACO model) buffers CWB. These findings indicate that some explicit personality constructs are related to CWB. However, the relationships are at best moderate and it may thus be useful to study other individual difference constructs such as implicit motives.

Implicit Motives and Counterproductive Workplace Behavior

Personality traits based on the FFM are commonly described as "probabilistic descriptions of relatively stable patterns of emotion, motivation, cognition, and behavior, in response to classes *of stimuli* that have been present in human cultures over evolutionary time" (DeYoung, 2015, p. 35), whereas implicit motives are typically characterized as stable

preferences for classes of affectively charged incentives or goals (Brunstein & Maier, 2005). The capacity to affectively charge motive relevant incentives leads people high in an implicit motive to follow motive relevant incentives and select behaviors leading to motive satisfaction (Stanton, Hall, & Schultheiss, 2010). Implicit motives typically predict a class of behaviors that McClelland et al. (1989) named *operant* behaviors. Operant behaviors are generated spontaneously and it is difficult to specify the stimuli in the environment that elicit the behavior (Spangler, 1992). In contrast, personality traits typically predict *respondent* behavior with known stimuli and immediate choices (McClelland et al., 1989).

Researchers typically study the *Big Three* implicit motives: Affiliation, Power, and Achievement (Kehr, 2004). *Affiliation*-motivated people want to build and maintain positive relationships with others. *Power*-motivated people want to influence others, and search for opportunities to lead and move upwards in organizational hierarchies. *Achievement*-motivated people want to reach performance goals and improve their skills. Implicit motives are not accessible through introspection and researchers therefore commonly measure them with the analysis of imaginary verbal responses to picture cues (Kuhl & Scheffer, 2001; McClelland, Atkinson, Clark, & Lowell, 1953; Winter, 1994).

Researchers found that implicit motives predict career success (Spangler, 1992; Winter et al., 1998), entrepreneurial success (Rauch & Frese, 2007; Wainer & Rubin, 1969), income (Apers et al., 2018), and task and contextual performance in interaction with personality traits (Lang et al., 2012). In this study, we want to investigate whether implicit motives also contribute to the understanding of CWB over and above personality traits. Spector (2011) described that CWB is often—in addition to planned and controlled acts—immediate, reactive, and impulsive in response to events in the work environment like provocations. Spontaneous and operant behaviors are typically predicted by implicit motives and less by personality traits (Spangler, 1992). Therefore, we suggest that implicit motives may have incremental predictive validity above personality traits in the prediction of CWB.

CWB can have negative consequences for an organization and individuals in the organization (Spector & Fox, 2005). There may be, however, also negative consequences for people who engage in CWB, because some types of CWB are visible such as intentionally working slowly or directly harming others by gossiping. A possible negative consequence of

CWB is that coworkers may have negative impressions of employees who engage in CWB (Farley, 2011). This can result in negative reactions from the work environment towards the employees who engaged in CWB. For instance, if someone feels threatened by interpersonal aggressive behavior, the person may engage in retaliation and revenge acts harming the original perpetrator. This can lead to a downward spiral of conflict escalation (Andersson & Pearson, 1999; Glomb & Liao, 2003). Similarly, when employees directly harm an organization through theft or sharing confidential organizational information and it is discovered, it can have serious consequences for the employee. Negative social and organizational consequences of CWB may make it more difficult to successfully strive for implicit motives, because those consequences can conflict with goal attainment. For instance, an employee receiving negative reactions to her/his behavior from coworkers will have difficulties striving for Affiliation goals with those coworkers. CWB may, therefore, be a behavior that is relevant for goal attainment. Employees high in an implicit motive may actively avoid engaging in CWB. As traits and motives describe different aspects of personality, we expect that implicit motives predict CWB over and above personality traits and, thus, have incremental validity. This leads us to our first hypothesis:

Hypothesis 1: Employees' implicit motives improve the prediction of counterproductive work behavior when personality traits are taken into account.

In the following sections, we theorize more specifically for the implicit Affiliation and Power motive how CWB may interfere with a successful motive thriving, and, consequently, why they may act as a buffer for CWB.

The Implicit Affiliation Motive and Counterproductive Workplace Behavior

The implicit Affiliation motive is defined as a concern to establish, maintain, or restore positive emotional relationships with other people or a group (Koestner & McClelland, 1992). People high in the implicit Affiliation motive try to actively find affiliative activities, spend their time interacting with other people (McClelland, 1985), and try to maintain the goodwill of their social interaction partner (McClelland, 1975). They avoid conflict (Exline, 1962), are sympathetic towards others (Koestner & McClelland, 1992), and adapt their social behavior to achieve a peaceful relationship (Walker & Heyns, 1962). Historical and experimental research also showed that Affiliation motivated people facilitate concessions and

compromises to reach peaceful solutions in international crises (Langner & Winter, 2001). Taken together, people high in the implicit Affiliation motive invest their resources into positive emotional relationships with others and value them highly.

Behaviors and goals of Affiliation motivated employees typically do not include behaviors that harm others. On the contrary, consequences of CWB may actually lead to need frustration for Affiliation motivated employees. Employees engaging in negative gossiping, bullying, or violence may be seen as less likable by others and leave negative impressions. Research has found that high-frequency gossipers are liked less (Farley, 2011), bullies in school are rejected by most peers (Pellegrini, Bartini, & Brooks, 1999), and are more likely to be isolated in teams (Coyne, Craig, & Smith-Lee Chong, 2004). Social relationships are important for people high in the implicit Affiliation motive, thus, they have strong reasons to disengage from CWB. People high in the implicit Affiliation motive generally value positive relationships, devalue conflict, and possible consequences of CWB may conflict with those goals. Therefore, we expect a negative relationship between the implicit Affiliation motive and CWB.

Hypothesis 2: People high in the implicit Affiliation motive show less CWB compared to people low in the implicit Affiliation motive.

The Implicit Power Motive and Counterproductive Workplace Behaviors

Researchers suggested that the implicit Power motive has a dual nature, it can inspire great leaders, but also lead to corruption and destruction (McClelland, 1970; Winter, 2016). In this study, we solely focus on *socialized* implicit Power—or need for social influence—and do not study *personalized* implicit Power focusing on the desire for dominance leading to profligate and impulsive behavior (Hofer et al., 2010; Winter, 1973). The implicit Power motive is characterized by an implicit desire to help and influence others, to control the environment, and maintain reputation, honor, and prestige (Winter, 1973). In an organizational context, Power motivated employees search for situations in which they can take the lead and look for opportunities to reach higher and more prestigious positions in an organization. Researchers suggested that charismatic leaders have a high implicit Power motive (House & Howell, 1992) and found evidence in a study linking the implicit Power motive of US presidents to charismatic leadership behavior (House, Spangler, & Woycke,

1991). The implicit Power motive also predicted socially responsible behavior over a 10-year period (Winter, McClelland, & Stewart, 1981).

A central characteristic that people look for in a leader is integrity and trustworthiness (Kouzes & Posner, 2002; Lord, Foti, & De Vader, 1984). Whether or not someone is perceived as a credible leader may depend, among other factors, also on perceived integrity. Therefore, Hogan and Kaiser (2005, p. 173) wrote that the most important question one asks of a potential leader is "Can we trust you not to abuse the privilege of authority?". People showing CWB are less likely to be perceived as trustworthy and integer. It may, therefore, be difficult for an employee to become a leader if decision-makers in an organization think the employee may potentially harm the organization. Employees high in the implicit Power motive search for possibilities to take the lead. We, therefore, expect that they control their behavior because they see the potential harm that CWB could inflict on their position within the organization. Employees who are not high in the implicit Power motive do not seek a higher position within a company and may not be affected by negative consequences of CWB. We, therefore, expect a negative relationship between the implicit Power motive and CWB.

Hypothesis 3: People high in the implicit Power motive show less CWB compared to people low in the implicit Power motive.

Implicit Achievement Motivation and Counterproductive Workplace Behaviors

The implicit Achievement motive is defined as a concern to improve one's performance, to do well in challenging tasks, and meet a standard of excellence (McClelland et al., 1989). Employees high in the implicit Achievement motive get satisfied by mastering tasks independently in Achievement-related situations (Brunstein & Maier, 2005). They seek, benefit, and learn from performance feedback (Fodor & Carver, 2000; McClelland, 1985). Employees engaging in CWB may receive less help and feedback from an organization. Having less support would make it more difficult for them to successfully strive for their Achievement motive. This line of reasoning suggests that employees high in the implicit Achievement motive may show less CWB. However, Achievement motivated employees may not always be successful in striving for their goals. Failure can lead to reduced positive affect, reduced performance and to frustration (Baumann, Kaschel, & Kuhl, 2005; Kehr, 2004; Lang

et al., 2012). Young (2009) argued that employees high in the Achievement motive are more likely to react strongly to possible job frustrations which may ultimately lead to turnover. A typical reaction to job frustration is to engage in CWB (Fox, Spector, & Miles, 2001). It is, thus, also possible that the Achievement motive could lead to more CWB. Given that research on Achievement motivation suggests different possible relationships to CWB, we decided to investigate the link between the implicit Achievement motive and CWB as a research question.

Research question 1: Is there a relation between the implicit Achievement motive and CWB?

The Present Study

Researchers have long argued that CWB ratings can differ depending on the rating source. Self-ratings and supervisor ratings may cover different aspects of CWB (Berry, Carpenter, & Barratt, 2012). In this paper, we present two studies with different rating sources. In the first study[10], we used employee self-ratings for CWB. An advantage of CWB self-ratings is that CWB may include covert behaviors so that the only complete source of employee engagement in CWB is the employee itself (Berry et al., 2012). In the second study, we asked supervisors to rate employee CWB. Supervisor ratings have the advantage that they seem less likely to underreport CWB due to social desirability biases and also ensure that predictor and criterion measures come from different sources (Berry et al., 2012). Studying CWB with self- and supervisor reports combines the advantages of both approaches.

Method

Study 1

Participants and Procedure. The participants in the first study were a diverse sample of German employees from various organizations and occupational backgrounds. Three research assistants contacted people they knew personally and through social networking sites to participate in the study. In addition, they went to public offices and shops in city centers of larger cities and asked employees whether they wanted to participate in the study. The instructions asked the employees to fill out a questionnaire booklet and return it using a

[10] The data for Study 1 overlaps with data in Lang et al. (2012).

stamped return envelope. Participants could provide their e-mail addresses to receive a summary of the findings of the study via e-mail. A total of 520 employees agreed to participate. We received 272 questionnaires back. A total of 263 questionnaires provided complete information on the study variables. The employees (158 female, 105 male) were between 18 and 69 years of age (M = 35.25, SD = 10.88), worked M = 36.48 hrs per week on average (SD = 12.04 hrs), and had been in their organizations for an average of M = 7.67 years (SD = 9.33 years). We asked the participants to indicate their job with an open question and provide an overview in the Appendix.

Measures.

Implicit motives. Implicit motives are not accessible through introspection and can therefore not be measured via questionnaires (McClelland et al., 1989). Researchers therefore developed picture based procedures where respondents write a story that gets coded for motivational content (McClelland et al., 1953; Morgan & Murray, 1935). We measured implicit motives with the Operant Motive Test (OMT; Kuhl & Scheffer, 2001). The OMT is a modernized version of earlier picture-based measures like the Thematic Apperception Test (Morgan & Murray, 1935) and the Picture Story Exercise (McClelland et al., 1953; Winter, 1994) with more pictures. In the OMT, respondents see 15 pictures with ambiguous social scenes (see Runge et al., 2016, for an example). Respondents are told to imagine a story that spontaneously comes to their mind when they see the pictures. They are then instructed to indicate the main person of their story (in pictures with more than one person) and answer the following two questions "What is important for the person in this situation and what is the person doing?" and "How does the person feel and why does the person feel this way?". The verbal responses are then coded into five different levels of Affiliation, Power, and Achievement motives and non-motivational content. The five levels describe different modes of motive enactment and researchers typically combine the first three levels to an approach motive (Runge et al., 2016; Runge, Lang, Chasiotis, & Hofer, 2018). In this study, we wanted to capture approach oriented motivational behavior and therefore also combined the first three levels for each motive. This also ensured that the implicit Power motive measure only included socialized Power content and not personalized Power. The responses were coded by a trained expert coder who received an extensive coder training and reached a high agreement

with coworkers of the original author of the OMT. To analyze inter-rater agreement, two trained student assistants coded a subsample of the responses. The first assistant coded 65 and the second 81 OMT responses. To determine rater agreement, we calculated Gwet's AC_1 statistic (Gwet, 2008a, 2008b). Gwet's AC_1 is an improved version of Cohen's Kappa. Rater agreement between the main coder and the first assistant was .87 for Affiliation, .76 for Power, and .83 for Achievement and rater agreement between the main coder and the second assistant was .89 for Affiliation, .76 for Power, and .90 for Achievement.

Personality. We measured the five factors from the FFM with the German adaptation of the Big Five Inventory (John, Naumann, & Soto, 2008), a 45-item scale. Participants responded using a 5-point Likert-type scale ranging from (1) *totally disagree* to (5) *totally agree*. Cronbach's α was α = .85, α = .78, α = .68, α = .80, and α = .75, for Extraversion, Conscientiousness, Agreeableness, Neuroticism, and Openness, respectively.

Counterproductive Work Behavior. CWB was rated by employees using the workplace deviance measure by Bennett and Robinson (2000). The scale consists of 19 items describing deviant behaviors such as mobbing or property theft. Respondents indicated on a 7-point Likert scale ranging from 1 (*never*) to 7 (*daily*) the extent to which they had engaged in each of the behaviors in the last year The CWB score had a Cronbach's α of .80.

Study 2

Participants and Procedure. The participants in this study were a diverse sample of Dutch and German employees and their respective supervisors from different organizations with various occupational backgrounds. Three research assistants contacted people they knew personally or through social networking sites and employees in public offices and shops. We first asked employees whether they wanted to participate in a scientific study on personality and workplace behavior. If they agreed, we then directly asked the supervisor to participate. If both agreed to participate, we gave them an employee and a supervisor questionnaire booklet including two stamped return envelopes. We instructed the employees to hand the supervisor questionnaire booklet and one envelope to a supervisor and fill out the employee questionnaire booklet. In addition, the instructions asked the employee and the supervisor to then return the questionnaire via mail. Participants had the option to provide their email address to receive a summary of the findings of the study via e-mail. We received envelopes from 128 employees

and 125 supervisors, and 121 employee-supervisor dyads were complete. The employees (72 female, 45 male; 4 did not indicate their gender) in these dyads were between 20 and 61 years of age (M = 34.36, SD = 11.14), worked M = 37.39 hours per week on average (SD = 10.07), and had been in their organizations for an average of M = 5.74 years (SD = 9.60 years). Additionally, we asked the participants to indicate their job with an open question and provide an overview in the Appendix. Supervisors had indicated that they had been working with their employees for on average M = 4.02 years (SD = 6.05). Most of the supervisors reported that they had frequent interactions with their employees (n = 88 at least once a day; n = 28 at least once a week).

Measures.

Implicit motives. We used the same measure for implicit motives as in study 1. The responses were coded by three student assistants that received an extensive coder training. To analyze rater agreement, we used a recoding of the data by a fourth trained student assistant that coded all stories using a variant of the OMT coding manual including autonomy as a fourth motive. Gwet's AC_1 was .93 for Affiliation, .83 for Power, and .92 for Achievement for the first rater (n=32), .92 for Affiliation, .78 for Power, and .93 for Achievement for the second rater (n=50), and .96 for Affiliation, .84 for Power, and .95 for Achievement for the third rater (n=39).

Personality. We measured the six HECAXO personality dimensions with the 100-item version of the HEXACO Personality Inventory-Revised 100 (Lee & Ashton, 2018). Participants responded using a 5-point Likert-type scale ranging from (1) *totally disagree* to (5) *totally agree*. Internal consistency was analyzed with Cronbach's α and was α = .87, α = .87, α = .83, α = .83, α = .80, and α = .81, for Honesty-Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness to Experience, respectively.

Counterproductive Work Behavior. CWB was measured using the same scale by Bennett and Robinson as in Study 1. However, in contrast to Study 1 where the scale was rated by the employees themselves, in Study 2, the employees' supervisors answered the questions. The CWB scale had a Cronbach's α of .92.

Results

Descriptive Statistics and Correlations

Table 1 provides descriptive information and intercorrelations of the variables in Study 1. CWB correlates significant and negatively with Conscientiousness, Agreeableness, and the implicit Affiliation motive. Descriptive information and intercorrelations for Study 2 are provided in Table 2. CWB correlates significant and negatively with the implicit Affiliation motive. CWB ratings have a relatively low mean in both studies and the mean for self-reported CWB is higher than for supervisor-rated CWB. This finding is consistent with literature suggesting that supervisors may have less opportunity to observe employees engaging in CWB (Berry et al., 2012).

Multiple Regression Analysis

To test our hypothesis, we conducted multiple regression analysis to predict CWB with personality and the three implicit motives. All analyses were conducted using R. The upper part of Table 3 shows the multiple regression analysis testing our hypotheses for Study 1. To examine our three hypotheses and the research question, we first fitted a model including only personality traits (M1). We then fitted a model including only the three implicit motive (M2). Finally, we fitted a third model including both personality traits and implicit motives (M3). The upper part of Table 3 shows the results of all three models for Study 1. We started our analyses with Hypothesis 1 stating that implicit motives explain additional variance in CWB over personality traits. To examine this hypothesis, we compared the explained variance of the three models using multiple R^2 values. As indicated in Table 3, results showed that the personality only model (M1; $R^2 = .16$) explained more variance than the implicit motives only model (M2; $R^2 = .05$). When implicit motives were added to the personality model, explained variance (M3; $R^2 = .20$) increased significantly ($F(257) = 4.02$; $p = 0.001$). Thus, Hypothesis 1 is supported in Study 1, suggesting that implicit motives show incremental validity over personality for self-rated CWB. Hypothesis 2 stated that Affiliation motivated employees show less CWB. In both Model 2 (ß = -0.25; $t = -3.67$; $p < 0.01$) and Model 3 (ß = -0.20; $t = -3.12$; $p < 0.01$), Affiliation was a significant predictor of CWB. Hypothesis 2 is therefore supported in Study 1. Hypothesis 3 stated that Power is negatively related to CWB. In both Model 2 (ß = -0.10; $t = -1.48$; $p = 0.14$) and Model 3 (ß = -0.05; $t = -0.75$; $p = 0.45$), Power was not related to CWB. Hypothesis 2 is therefore not supported in Study 1. For our research

question, we were interested in the relationship between the Achievement motive and CWB. Results show no relationship between them in Study 1.

We analyzed Study 2 using the same strategy as in Study 1. In Study 2, personality was measured with the HEXACO and CWB was rated by supervisors. The lower part of Table 3 shows the multiple regression analysis testing our hypothesis for Study 2. For Hypothesis 1, the analysis showed that the personality only model (M1; R^2 = .02) explained less variance than the implicit motives only model (M2; R^2 = .08). When implicit motives were added to the personality model, explained variance increases (M3; R^2 = .11), significantly ($F(114) = 3.75$; p = .0.01). Thus, Hypothesis 1 is also supported in Study 2. Testing Hypothesis 2, Affiliation motivation was a significant predictor of CWB in Model 2 (ß = -0.19; t = -2.04; p = 0.04) and Model 3 (ß = -0.19; t = -2.16; p = 0.02). Hypothesis 2 was therefore also supported in Study 2. For Hypothesis 3, we found that Power was a significant predictor for CWB in Model 2 (ß = -0.20; t = -2.21; p = 0.03) and Model 3 (ß = -0.22; t = -2.29; p = 0.02). Contrary to Study 1, Hypothesis 3 was supported in Study 2. Results for the research questions also showed no relationship between Achievement and CWB in Study 2.[11]

Discussion

The general aim of this study was to advance the literature on CWB by considering implicit motives as a key albeit unexplored antecedent. We specifically investigated whether implicit motives have incremental validity over personality traits in predicting CWB. We investigated this in two studies with broad occupational samples measuring CWB with self-ratings in the first study and supervisor ratings in the second study. We found that implicit motives show incremental validity in the prediction of CWB in addition to personality traits in both studies. More specifically, we found that employees high in the implicit Affiliation

[11] The Bennett and Robinson (2000) measure includes two subscales: Interpersonal CWB (targeted at members of an organization) and organizational CWB (targeted at the organization itself). We therefore also ran our analyses with interpersonal and organizational CWB. The results were largely similar: interpersonal CWB, Affiliation Study 1 (ß = -0.16; t = -2.50; p = 0.01) Study 2 (ß = -0.19; t = -2.07; p = 0.04), Power Study 1 (ß = -0.03; t = -0.47; p = 0.64) Study 2 (ß = -0.20; t = -2.05; p = 0.04); organizational CWB, Affiliation Study 1 (ß = -0.16; t = -2.57; p = 0.01) Study 2 (ß = -0.17; t = -1.85; p = 0.07), Power Study 1 (ß = -0.05; t = -0.70; p = 0.49) Study 2 (ß = -0.19; t = -2.06; p = 0.04). All values refer to Model 3 respectively.

motive show less CWB in both studies and employees high in the implicit Power motive show less CWB in the second study. Below, we discuss implications of those findings.

Implications

This paper contributes to the literature by showing that implicit motives improve the prediction of CWB. Recent studies have shown that implicit motives predict various work-related outcomes in combination with personality traits (Apers et al., 2018; Kehr, 2004; Lang et al., 2012), but we are not aware of any research linking implicit motives to CWB. Our findings suggest that conclusions regarding the role of personality in CWB may need to be extended. Thus far, most research has focused on explicit personality traits (Hershcovis et al., 2007; Salgado, 2002; Zettler, in press). Over the two studies, we demonstrated that implicit motives play an important role in predicting CWB. This finding is in line with researchers (Winter et al., 1998) arguing that implicit motives, as in the tradition of Murray (1938) and McClelland (1984), play an equally important role in the description of personality than traits, as in the tradition of Allport (1931). We believe that these findings complement previous research on individual difference variables on CWB by also considering implicit motives.

We found in both studies, with self- and supervisor-rated CWB that the implicit Affiliation motive buffers CWB. This finding is supportive of theoretical ideas and qualitative findings that social motives and interpersonal interaction can play an important role at work (Wrzesniewski, Dutton, & Debebe, 2003). It is also in line with previous research that studied CWB and personality traits that are important in social situations. Researchers found evidence that Agreeableness and Emotional Stability predict CWB (Berry et al., 2007; Hershcovis et al., 2007). Both traits are linked to Affiliation motives (Barrick, Mount, & Li, 2013). We conclude that both theoretical ideas and empirical studies support the notion that employees who strive for Affiliation goals are typically not harming coworkers or the organization they work in.

We found that the implicit Power motive buffers CWB only in Study 2 using supervisor ratings of CWB, but not in Study 1 using self-ratings of CWB. A possible explanation is that ambitious employees striving for higher positions in an organization are especially hesitant to show CWB that is visible to supervisors. Previous research revealed that this may be a successful strategy. Supervisor-focused impression management strategies are related to career success (Judge & Bretz, 1994), higher ratings of organizational citizenship

behavior, liking by the supervisor (Bolino, Varela, Bande, & Turnley, 2006), and supervisor-rated job performance (Wayne & Liden, 1995). It is, thus, possible that employees high in the implicit Power motive use supervisor-focused impression management strategies and avoid negative impressions by their supervisor through disengaging in CWB.

For our research question, we did not find any relationship between the implicit Achievement motivation and CWB. In contrast to the other two implicit motives (Affiliation and Power), the implicit Achievement motive has no social focus and does not require the presence of others or the organization. Research has shown that a major situational predictor of CWB is rooted in interpersonal conflicts and perceptions of organizational and interpersonal injustice (Berry et al., 2007; Hershcovis et al., 2007). As the interpersonal context is less important for the implicit Achievement motive, it seems plausible that we did not find any relationship.

Limitations

As a first limitation, both of our studies use cross-sectional research designs. Accordingly, it is not possible to infer causality. Future research should, therefore, also explore effects of implicit motives on CWB over time. Another potential limitation of the investigation is that Study 1 used only self-reports of CWB. However, we replicated and extended the findings in Study 2 using supervisor-ratings of CWB. At the same time, supervisor-rated CWB might also come with some specific aspects. For instance, correlations between personality traits and CWB are low in this sample (e.g., from the HEXACO model, Honesty-Humility, Agreeableness, and Conscientiousness have typically been found to correlate with self- and colleague-ratings of CWB [e.g. Marcus et al., 2007]). Note, though, that the findings of Study 2 are still within the range of typical effect sizes for observer-ratings reported by Berry et al. (2012), but on the low end. Irrespective of the similarities and differences between the rating sources (see Berry et al., 2012), we believe that the reliance of self- and supervisor-ratings of CWB in the investigation overall is rather a strength, than a limitation.

The sample in both studies is from broad occupational backgrounds. It is unclear to what extent the observed effects can be generalized to a subsample with more specific jobs.

Future research should study whether job contexts may have an influence on the effects in the present paper. For instance, in jobs with few social interactions, it seems plausible that the effect for the implicit Affiliation motive on CWB may be weaker.

Conclusion

Studying and understanding antecedents of CWB is important because CWB can be costly for organizations (Marcus et al., 2016). One line of research has therefore studied individual difference variables as predictors of CWB (Salgado, 2002; Spector, 2011). The present study contributes to this literature on individual difference variables by examining the effects of implicit motives on CWB. Overall, our findings support the idea that implicit motives have incremental validity in the prediction of CWB over personality traits. Our study further suggests that the implicit Power motive and especially the implicit Affiliation motive predict CWB. In conclusion, we suggest that it may be beneficial for researchers studying CWB to consider implicit motives and more specifically the implicit Affiliation and the implicit Power motive as buffers of CWB.

References

Allport, G. W. (1931). What is a trait of personality? *Journal of Abnormal and Social Psychology*, *25*, 368–372.

Andersson, L. M., & Pearson, C. M. (1999). Tit for tat? The spiraling effect of incivility in the workplace. *Academy of Management Review, 24*, 452–471. https://doi.org/10.5465/amr.1999.2202131

Apers, C., Lang, J. W. B., & Derous, E. (2018). Who earns more? Explicit traits, implicit motives and income growth trajectories. *Journal of Vocational Behavior*. https://doi.org/10.1016/J.JVB.2018.12.004

Barrick, M. R., Mount, M. K., & Li, N. (2013). The theory of purposeful work behavior: The role of personality, job characteristics, and experienced meaningfulness. *Academy of Management Review, 38*, 1–51. https://doi.org/10.5465/amr.2010.0479

Baumann, N., Kaschel, R., & Kuhl, J. (2005). Striving for unwanted goals: Stress-dependent discrepancies between explicit and implicit achievement motives reduce subjective well-being and increase psychosomatic symptoms. *Journal of Personality and Social Psychology*, *89*, 781–799. https://doi.org/10.1037/0022-3514.89.5.781

Bennett, R. J., & Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of Applied Psychology, 85*, 349–360. https://doi.org/10.1037/0021-9010.85.3.349

Berry, C. M., Carpenter, N. C., & Barratt, C. L. (2012). Do other-reports of counterproductive work behavior provide an incremental contribution over self-reports? A meta-analytic comparison. *Journal of Applied Psychology, 97*, 613–636. https://doi.org/10.1037/a0026739

Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal deviance, organizational deviance, and their common correlates: A review and meta-analysis. *Journal of Applied Psychology*, *92*, 410–424. https://doi.org/10.1037/0021-9010.92.2.410

Bolino, M. C., Varela, J. A., Bande, B., & Turnley, W. H. (2006). The impact of impression-management tactics on supervisor ratings of organizational citizenship behavior. *Journal of Organizational Behavior, 27*, 281–297. https://doi.org/10.1002/job.379

Brunstein, J. C., & Maier, G. W. (2005). Implicit and self-attributed motives to achieve: Two

separate but interacting needs. *Journal of Personality and Social Psychology, 89*, 205–222. https://doi.org/10.1037/0022-3514.89.2.205

Collins, J. M., & Griffin, R. W. (1998). The psychology of counterproductive job performance. In R. W. Griffin, A. O'Leary-Kelly, & J. M. Collins (Eds.), *Dysfunctional behavior in organizations: Non-violent dysfunctional behavior. Monographs in organizational behavior and relations* (pp. 219–242).

Coyne, I., Craig, J., & Smith-Lee Chong, P. (2004). Workplace bullying in a group context. *British Journal of Guidance & Counselling*, *32*, 301–317. https://doi.org/10.1080/03069880410001723530

DeYoung, C. G. (2015). Cybernetic Big Five theory. *Journal of Research in Personality*, *56*, 33–58. https://doi.org/10.1016/j.jrp.2014.07.004

Emmons, R. A. (1993). The current status of the motive concept. In K. H. Craik, R. Hogan, & R. N. Wolfe (Eds.), *Fifty Years of Personality Psychology* (pp. 187–196). New York: Plenum.

Exline, R. V. (1962). Need affiliation and initial communication behavior in problem-solving groups characterized by low interpersonal visibility. *Psychological Reports*, *10*, 79–89. https://doi.org/10.2466/pr0.1962.10.1.79

Farley, S. D. (2011). Is gossip power? The inverse relationships between gossip, power, and likability. *European Journal of Social Psychology*, *41*, 574–579. https://doi.org/10.1002/ejsp.821

Fodor, E. M., & Carver, R. A. (2000). Achievement and power motives, performance feedback, and creativity. *Journal of Research in Personality*, *34*, 380–396. https://doi.org/10.1006/jrpe.2000.2289

Fox, S., Spector, P. E., & Miles, D. (2001). Counterproductive work behavior (CWB) in response to job stressors and organizational justice: Some mediator and moderator tests for autonomy and emotions. *Journal of Vocational Behavior*, *59*, 291–309. https://doi.org/10.1006/jvbe.2001.1803

Glomb, T. M., & Liao, H. (2003). Interpersonal aggression in work groups: Social influence, reciprocal, and individual fffects. *Academy of Management Journal*, *46*, 486–496. https://doi.org/10.5465/30040640

Gwet, K. L. (2008a). Computing inter-rater reliability and its variance in the presence of high agreement. *The British Journal of Mathematical and Statistical Psychology, 61*, 29–48. https://doi.org/10.1348/000711006X126600

Gwet, K. L. (2008b). Variance estimation of nominal-scale inter-rater reliability with random selection of raters. *Psychometrika, 73*, 407–430. https://doi.org/10.1007/s11336-007-9054-8

Hershcovis, M. S., Turner, N., Barling, J., Arnold, K. A., Dupré, K. E., Inness, M., … Sivanathan, N. (2007). Predicting workplace aggression: A meta-analysis. *Journal of Applied Psychology, 92*, 228–238. https://doi.org/10.1037/0021-9010.92.1.228

Hofer, J., Busch, H., Bond, M. H., Campos, D., Li, M., & Law, R. (2010). The implicit power motive and sociosexuality in men and women: Pancultural effects of responsibility. *Journal of Personality and Social Psychology, 99*, 380–394. https://doi.org/10.1037/a0020053

Hogan, R., & Kaiser, R. B. (2005). What we know about leadership. *Review of General Psychology, 9*, 169–180.

Hollinger, R. C., & Davis, J. L. (2002). *National retail security survey 2001*. Gainesville.

House, R. J., & Howell, J. M. (1992). Personality and charismatic leadership. *The Leadership Quarterly, 3*, 81–108. https://doi.org/10.1016/1048-9843(92)90028-E

House, R. J., Spangler, W. D., & Woycke, J. (1991). Personality and charisma in the U.S. presidency: A psychological theory of leader effectiveness. *Administrative Science Quarterly, 36*, 364. https://doi.org/10.2307/2393201

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of Personality: Theory and Research* (3rd ed., pp. 114–158). New York: The Guilford Press.

Judge, T. A., & Bretz, R. D. (1994). Political influence behavior and career success. *Journal of Management, 20*(1), 43–65. Retrieved from https://www.sciencedirect.com/science/article/pii/S0149206305800042

Kehr, H. M. (2004). Integrating implicit motives, explicit motives, and perceived abilities: The compensatory model of work motivation and volition. *Academy of Management*

Review, *29*, 479–499. https://doi.org/10.5465/AMR.2004.13670963

Koestner, R., & McClelland, D. C. (1992). The affiliation motive. In C. P. Smith (Ed.), *Motivation and Personality: Handbook of Thematic Content Analysis* (pp. 205–210). Cambridge, MA: Cambridge University Press.

Kouzes, J. M., & Posner, B. Z. (2002). *The Leadership Challenge*. San Francisco: JosseyBass.

Kuhl, J., & Scheffer, D. (2001). Der operante multi-motiv-test (OMT): Manual [The operant multi-motive-test (OMT): Manual]. Osnabrück, Germany: University of Osnabrück.

Lang, J. W. B., Zettler, I., Ewen, C., & Hülsheger, U. R. (2012). Implicit motives, explicit traits, and task and contextual performance at work. *Journal of Applied Psychology*, *97*, 1201–1217. https://doi.org/10.1037/a0029556

Langner, C. A., & Winter, D. G. (2001). The motivational basis of concessions and compromise: archival and laboratory studies. *Journal of Personality and Social Psychology*, *81*, 711–27. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11642356

Lee, K., & Ashton, M. C. (2018). Psychometric properties of the HEXACO-100. *Assessment*, *25*, 543–556. https://doi.org/10.1177/1073191116659134

Lord, R. G., Foti, R. J., & De Vader, C. L. (1984). A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. *Organizational Behavior and Human Performance*, *34*, 343–378. https://doi.org/10.1016/0030-5073(84)90043-6

Marcus, B., Lee, K., & Ashton, M. C. (2007). Personality dimensions explaining relationships between integrity tests and counterproductive behavior: Big Five, or one in addition? *Personnel Psychology*, *60*, 1–34.

Marcus, B., Taylor, O. A., Hastings, S. E., Sturm, A., & Weigelt, O. (2016). The structure of counterproductive work behavior. *Journal of Management*, *42*, 203–233. https://doi.org/10.1177/0149206313503019

McClelland, D. C. (1970). The two faces of power. *Journal of International Affairs*, *24*, 29 – 47.

McClelland, D. C. (1975). *Power: The Inner Experience.* New York: Wiley.

McClelland, D. C. (1984). *Motives, Personality, and Society: Selected Papers.* New York, NY, NY: Praeger Publishers.

McClelland, D. C. (1985). *Human Motivation*. Glenview, IL: Scott Foresman.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). The measuring instrument. In *The Achievement Motive* (pp. 185–217). East Norwalk, CT, US: Appleton-Century-Crofts.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*, 690–702. https://doi.org/10.1037/0033-295X.96.4.690

McCrae, R. R., & John, O. P. (1992). An introduction to the Five-Factor Model and its applications. *Journal of Personality*, *60*, 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

Morgan, C. D., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology & Psychiatry*, *34*, 289–306. https://doi.org/10.1001/archneurpsyc.1935.02250200049005

Murray, H. A. (1938). *Explorations in Personality; A clinical and Experimental Study of Fifty Men of College Age.* New York: Oxfort University Press. https://doi.org/10.2307/2084329

Pellegrini, A. D., Bartini, M., & Brooks, F. (1999). School bullies, victims, and aggressive victims: Factors relating to group affiliation and victimization in early adolescence. *Journal of Educational Psychology*, *91*, 216–224. https://doi.org/10.1037/0022-0663.91.2.216

Rauch, A., & Frese, M. (2007). Let's put the person back into entrepreneurship research: A meta-analysis on the relationship between business owners' personality traits, business creation, and success. *European Journal of Work and Organizational Psychology*, *16*, 353–385. https://doi.org/10.1080/13594320701595438

Robinson, S. L., & Bennett, R. J. (1995). A typology of deviant workplace behaviors: A multidimensional scaling study. *Academy of Management Journal*, *38*, 555–572. https://doi.org/10.2307/256693

Robinson, S. L., & Greenberg, J. (1999). Employees behaving badly: Dimensions, determinants and dilemmas in the study of workplace deviance. In D. M. Rousseau & C. Cooper (Eds.), *Trends in organizational behavior* (pp. 1–23). New York: Wiley.

Retrieved from http://search.proquest.com/openview/42ae61e2b571ca6dc61e81c936dea153/1?pq-origsite=gscholar&cbl=47893

Runge, J. M., Lang, J. W. B., Chasiotis, A., & Hofer, J. (2018). Improving the assessment of implicit motives using IRT: cultural differences and differential item functioning. *Journal of Personality Assessment*, 1–11. https://doi.org/10.1080/00223891.2017.1418748

Runge, J. M., Lang, J. W. B., Engeser, S., Schüler, J., den Hartog, S. C. S. C., & Zettler, I. (2016). Modeling motive activation in the operant motives test: A psychometric analysis using dynamic Thurstonian item response theory. *Motivation Science, 2*, 268–286. https://doi.org/10.1037/mot0000041

Salgado, J. F. (2002). The Big Five personality dimensions and counterproductive behaviors. *International Journal of Selection and Assessment*, *10*, 117–125. https://doi.org/10.1111/1468-2389.00198

Salgado, J. F., Moscoso, S., & Anderson, N. R. (2013). Personality and Counterproductive Work Behavior. In N. D. Christiansen & R. P. Tett (Eds.), *Handbook of Personality at Work* (pp. 606–632). New York: Routledge.

Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin*, *112*, 140–154. https://doi.org/10.1037/0033-2909.112.1.140

Spector, P. E. (2011). The relationship of personality to counterproductive work behavior (CWB): An integration of perspectives. *Human Resource Management Review*, *21*, 342–352. https://doi.org/10.1016/j.hrmr.2010.10.002

Spector, P. E., & Fox, S. (2005). A model of counterproductive work behavior. In S. Fox & P. E. Spector (Eds.), *Counterproductive workplace behavior: Investigations of actors and targets* (pp. 151–174). Washington, DC: APA.

Stanton, S. J., Hall, J. L., & Schultheiss, O. C. (2010). Properties of motive-specific incentives. In O. C. Schultheiss & J. C. Brunstein (Eds.), *Implicit Motives* (pp. 245–278). New York, NY: Oxford University Press.

Wainer, H. A., & Rubin, I. M. (1969). Motivation of research and development entrepreneurs:

Determinants of company success. *Journal of Applied Psychology*, *53*, 178–184.

Walker, E. L., & Heyns, R. N. (1962). *An Anatomy for Conformity*. Eaglewood Cliffs, NJ: Prentice-Hall.

Wayne, S. J., & Liden, R. C. (1995). Effects of impression management on performance ratings: A longitudinal study. *Academy of Management Journal*, *38*, 232–260. https://doi.org/10.5465/256734

Winter, D. G. (1973). *The power motive*. New York, NY, US: Free Press.

Winter, D. G. (1994). *Manual for scoring motive imagery in running text*. Ann Arbor: University of Michigan.

Winter, D. G. (2016). Taming power: Generative historical consciousness. *American Psychologist*, *71*, 161–174. https://doi.org/10.1037/a0039312

Winter, D. G., John, O. P., Stewart, A. J., Klohnen, E. C., & Duncan, L. E. (1998). Traits and motives: Toward an integration of two traditions in personality research. *Psychological Review*, *105*, 230–250. https://doi.org/10.1037/0033-295X.105.2.230

Winter, D. G., McClelland, D. C., & Stewart, A. J. (1981). *A New Case for the Liberal Arts. Assessing Institutional Goals and Student Development*. Jossey-Bass. Retrieved from http://search.proquest.com/docview/63540143?accountid=13042 NS -

Wrzesniewski, A., Dutton, J. E., & Debebe, G. (2003). Interpersonal sensemaking and the meaning of work. *Research in Organizational Behavior*, *25*, 93–135. https://doi.org/10.1016/S0191-3085(03)25003-6

Young, A. M. (2009). Frustration-instigated career decisions: A theoretical exploration of the role of frustration in career decisions. *Human Resource Development Review*, *8*, 281–299. https://doi.org/10.1177/1534484309337209

Zettler, I. (n.d.). Personality and Counterproductive Work Behavior. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences*. Cham: Springer International Publishing. https://doi.org/https://doi.org/10.1007/978-3-319-28099-8_789-1

Zettler, I., & Hilbig, B. E. (2010). Honesty-humility and a person-situation interaction at work. *European Journal of Personality*, *24*, 569–582. https://doi.org/10.1002/per.757

Tables

Table 1

Means, Standard Deviations, and Intercorrelations of the Variables in Study 1

Measure	*M*	*SD*	1	2	3	4	5	6	7	8
1. Extraversion	3.65	0.65								
2. Conscientiousness	3.85	0.56	.21**							
3. Agreeableness	3.63	0.48	-.09	.06						
4. Neuroticism	2.78	0.64	-.13*	-.23**	-.26**					
5. Openness	3.51	0.53	.38**	.07	.04	-.11				
6. Affiliation	1.87	1.15	.02	.02	.09	-.01	-.03			
7. Power	5.75	2.13	.01	.05	.06	-.04	-.05	-.43**		
8. Achievement	2.03	1.07	.08	.10	.03	-.14*	.18**	.04	-.14*	
9. CWB	1.80	0.57	.03	-.29**	-.25**	.05	.10	-.20**	-.00	.07

Note. Study 1: $N = 263$. *M* and *SD* are used to represent mean and standard deviation, respectively. CWB = Counterproductive Work Behavior.

* $p < .05$; ** $p < .01$

Table 2

Means, Standard Deviations, and Intercorrelations of the Variables in Study 1 and Study 2

Measure	*M*	*SD*	1	2	3	4	5	6	7	8	9
1. Honesty-Humility	3.66	0.53									
2. Emotionality	3.22	0.56	-.05								
3. Extraversion	3.74	0.46	-.12	-.06							
4. Agreeableness	2.95	0.50	.37**	-.07	.11						
5. Conscientiousness	3.70	0.45	.05	.17	.16	.12					
6. Openness	3.35	0.55	.27**	-.07	.16	.12	.06				
7. Affiliation	1.78	1.30	-.00	.00	.02	.07	-.06	-.00			
8. Power	4.09	1.85	.06	-.00	.05	.05	.15	.04	-.17		
9. Achievement	2.03	1.28	-.02	-.01	-.04	.04	-.05	.11	.16	-.24**	
10. CWB	1.33	0.42	-.00	.04	.02	.12	.02	.04	-.18*	-.14	-.13

Note. Study 2: $N = 121$. *M* and *SD* are used to represent mean and standard deviation, respectively. CWB = Counterproductive Work Behavior.

* $p < .05$; ** $p < .01$

Table 3

Regression Analyses Predicting Counterproductive Work Behavior With Implicit Motives and Personality Traits

Parameter	Model 1			Model 2			Model 3		
	Est.	*t*	ß	Est.	*t*	ß	Est.	*t*	ß
				Study 1					
Coefficients (*b*)									
Intercept	3.74	7.67*		2.11	12.63*		3.84	7.80*	
Extraversion	0.02	0.36	.02				0.03	0.56	.04
Conscientiousness	-0.31	-5.15*	-.31				-0.31	-5.24*	-.31
Agreeableness	-0.29	-4.16*	-.25				-0.27	-3.80*	-.23
Neuroticism	-0.06	-1.14	-.07				-0.05	-0.95	-.06
Openness	0.12	1.81	-.11				0.09	1.39	.09
Affiliation				-0.12	-3.67*	-.25	-0.10	-3.12*	-.20
Power				-0.03	-1.48	-.10	-0.01	-0.75	-.05
Achievement				0.03	1.06	-.06	0.04	1.38	-.08
R^2	.16			.05			.20		
F	9.73*			4.92*			7.81*		
Δ R^2 vs. Model 1							.04		
Δ F^2 vs. Model 1							4.02*		
				Study 2					
Coefficients (*b*)									
Intercept	0.95	1.68		1.72	13.01*		1.23	2.22*	
Honesty-Humility	-0.05	-0.63	-.07				-0.06	-0.71	-.07
Emotionality	0.04	0.55	.05				0.04	0.57	.05
Extraversion	-0.01	-0.08	-.01				-0.01	-0.09	-.01
Agreeableness	0.12	1.46	.14				0.15	1.79	.18
Conscientiousness	0.00	-0.01	.00				0.01	0.09	.01
Openness	0.03	0.45	.04				0.05	0.69	.07
Affiliation				-0.06	-2.05*	-.19	-0.06	-2.16*	-.20
Power				-0.05	-2.21*	-.20	-0.05	-2.29*	-.22
Achievement				-0.05	-1.60	-.15	-0.05	-1.74	-.16
R^2	.02			.08			.11		
F	0.43			3.37*			1.56		
Δ R^2 vs. Model 1							.09		
Δ F^2 vs. Model 1							3.75*		

Note. Study 1: $N = 263$. Study 2: $N = 121$. Est. = Estimate. To get standard errors for the *b* and ß coefficients, divide the *b* and ß coefficients by the *t* values. dfs Study 1: Model 1 = 257, Model 2 = 259, and Model 3 = 254; dfs Study 2: Model 1 = 114, Model 2 = 117, and Model 3 = 111. *$p < .05$.

Appendix

Job description	Study 1	Study 2
Physicians and nurses	29	13
Clerks	45	25
Human resource management		5
Social workers	5	8
Engineers or technicians	20	13
Managers	7	3
Sales	18	4
Business consultants	7	5
Personnel training	6	2
College teachers or instructors	9	2
Banking and financial	7	2
Dentist / support personnel	4	1
Pharmacist	5	4
Optician	21	1
School or kindergarten teacher	37	1
Lawyers	2	
Carpenter		1
Public administration	15	
Tax accountants	4	
Level in the organizational hierarchy	13	21
Not filled out		10

CHAPTER 5

CAN PEOPLE RECOGNIZE THEIR IMPLICIT THOUGHTS? THE MOTIVE SELF-CATEGORIZATION TEST[12]

Psychologists have long been interested in studying individual differences in implicit motives. Implicit motives are typically measured asking respondents to write fantasy-stories based on a series of pictures showing one or several persons. The stories are then coded for implicit motivational content by trained experts because researchers have long assumed that respondents have no conscious access to the motivational themes in the stories they write. However, empirical research on self-evaluation of implicit motives is scarce. In this article, we provide new insights into this topic with a new measurement procedure—the motive self-categorization (MSC) test. In the MSC, respondents first fill out an implicit motive measure and then self-code their stories using lists of picture-specific statements that are typical concrete manifestations of implicit motives in the specific picture. We studied the MSC in a sample of N = 247 respondents by analyzing convergence with expert codings using a latent multitrait-multimethod (MTMM) IRT model. Results showed respondents could evaluate the motivational content of their stories (latent motive-motive correlations: r = .37 to 62), IRT latent motive scores based on self-categorization showed evidence of reliability (r = .42 to .67), and we found small method effects. The discussion focuses on implications for theory on measuring implicit motives and the possibility that self-insight occasionally goes beyond expert insight.

---

[12] This chapter is based on: Runge, J. M. & Lang, J. W. B (2019). Can people recognize their implicit thoughts? The motive self-categorization test. *Psychological Assessment*. http://dx.doi.org/10.1037/pas0000720

Implicit motives are described as associative networks connecting situational cues with basic affective reactions that lead to behavioral tendencies (McClelland, Atkinson, Clark, & Lowell, 1953). Due to challenges in measuring implicit motives and the dependency on expert raters, they are rarely used in daily clinical practice (Weinberger, Cotler, & Fishmen, 2010), even though clinical researchers and practitioners typically show high interest in individual differences in implicit and nonconscious aspects of personality (Baumann, Kaschel, & Kuhl, 2005; Mihura, Meyer, Dumitrascu, & Bombel, 2013; Murray, 1938; Westen, Lohr, Silk, Gold, & Kerber, 1990).

Researchers measure implicit motives with a series of ambiguous pictures to which respondents write short stories or explanations. These stories are then coded for affiliation, achievement, or power motivation themes by expert raters. The common justification for the use of expert coders is that motivational content is not immediately apparent to the person producing the story. The main argument is that respondents actually do not have a direct route to self-knowledge of their implicit motives and is in line with evidence suggesting implicit motive measures do not correlate with questionnaires that directly ask respondents for their motives (Köllner & Schultheiss, 2014; McClelland, Koestner, & Weinberger, 1989; Sokolowski, Schmalt, Langens, & Puca, 2000; Spangler, 1992). However, several researchers have argued that the non-convergence between traditional implicit motive tests and motive questionnaires does not necessarily mean that people cannot evaluate the implicit themes in the stories they write and have found some initial evidence that respondents are able to recognize and classify some of the implicit themes in their stories (Engeser, 2005; Sherwood, 1966; Thrash, Cassidy, Maruskin, & Elliot, 2010). Engeser (2005) let respondents code their story with a simplified coding scheme, Sherwood (1966) showed that participants can have some insight after a comprehensive motive education, and Thrash et al. (2010) discussed methodological and theoretical factors that may be responsible for previous findings showing that people cannot report their own implicit motives.

Building on this earlier work, the goal of the present investigation was to study whether people can self-evaluate their implicit motives with a new measurement procedure that we named *Motive Self-Categorization* (MSC). In the MSC, we let respondents categorize their stories with empirically based picture specific statements representing typical

manifestations of implicit motives in the specific picture. This is, to our knowledge, the first measure using tailored and empirically based items to self-categorize stories for implicit motives. We analyze the reliability of the MSC and expert coding methods, convergent validity of the MSC with expert codings and discriminant validity of the measured motives with a latent multitrait-multimethod (MTMM) model, and potential overlap of the MSC with self-ratings of personality to study whether people can self-evaluate their implicit motives.

Implicit Motives

Implicit motives are described as peoples wishes, desires, and as classes of goals that people like to pursue (Winter, John, Stewart, Klohnen, & Duncan, 1998). The classes of goals that researchers typically study are the "big three" motives: affiliation, achievement, and power. Affiliation motivated people seek to build and maintain relationships with others (Heyns, Veroff, & Atkinson, 1958). Achievement motivated people seek experiences related to improving their performance and are concerned to do well at challenging tasks with a moderate probability of success (Atkinson, 1957; McClelland, Koestner, & Weinberger, 1989). Power motivated people seek to have impact on others and maintain reputation and prestige (Winter, 1973).

Researchers have long argued that one cannot ask people directly for their implicit motives because they are, by definition, nonconscious and not accessible through introspection (e.g. McClelland et al., 1989). Researchers have therefore developed other methods to measure individual differences in implicit motives. A prominent measurement approach with a long history are picture- and story-based measures building on the basic principles of the Thematic Apperception Test (TAT; Morgan & Murray, 1935). In picture-based tests, respondents write stories in an open format in response to pictures showing ambiguous social scenes. The story then gets evaluated by expert coders into a motivational category (e.g., affiliation, achievement, and power). This open-ended answer format has been called operant, to distinguish it from the respondent immediate choice behavior in questionnaires (McClelland et al., 1989). Researchers have developed various implicit motive tests based on this core principle (e.g. Heckhausen, 1963; Kuhl & Scheffer, 2001; McClelland et al., 1953; Pang, 2006; Pang & Schultheiss, 2005; Schultheiss & Brunstein, 2001; Smith, 1992; Winter, 1994). Here, we focus on a popular response format in which people imagine a story and provide

brief responses based on several questions (Hagemeyer & Neyer, 2011; Kuhl & Scheffer, 2001).

Researchers have also adapted implicit measures that were originally developed to measure other constructs to also measure implicit motives. For instance, researchers have used the measurement procedure of the Implicit Association Tests that was originally developed to measure implicit cognitions (Greenwald, McGhee, & Schwartz, 1998), to measure implicit motivation (Brunstein & Schmitt, 2004; Ziegler, Schmukle, Egloff, & Bühner, 2010). A different approach was suggested by James (1998) who measured implicit achievement with a Conditional Reasoning Test in which respondents are instructed to read a problem and then choose the correct solution. The test gives—among incorrect solutions—two correct solutions, one of which indicates achievement striving. In this study, however, we contribute to research on picture and story-based procedures as measure of implicit motives.

Self-Evaluation of Implicit Motives

The question of how well people evaluate their personality has been a prominent theme in psychological research (McAbee & Connelly, 2016; Vazire & Carlson, 2010; Wilson & Dunn, 2004). Researchers have also studied whether people can evaluate implicit motives themselves (e.g. Sherwood, 1966). A core issue is how can people evaluate their implicit motives if they are defined as nonconscious and inaccessible to introspection. One approach to study whether people can evaluate their implicit motives has directly asked respondents for their implicit motives with Likert type questionnaires like the Personality Research Form (Jackson, 1984). McClelland et al. (1989) argued and provided evidence that operant implicit motive tests and respondent explicit tests measure different constructs and predict different classes of behavior. The general empirical conclusion has been that both type of measures do not correlate (Köllner & Schultheiss, 2014; Spangler, 1992).

A second approach for studying self-evaluation of implicit motives showed respondents pictures as stimuli and let them then indicate a motivational statement—based on scientific definitions of the motives—that fits best to the picture (Schmalt, 1999; Sokolowski et al., 2000). Researchers have called this approach semi-projective because of the respondent response format. Empirical research showed no correlations with implicit motive measures (Schüler, Brandstätter, Wegner, & Baumann, 2015).

In a third study, Engeser (2005) extended the semi projective technique described above and added story writing to the technique. He let 238 respondents categorize stories they wrote in response to pictures with a set of 21 general motivational statements directly reflecting a simplified version of the coding rules that experts use based on scientific motive definitions. An example statement for power motivation is "The person in the story wants to have influence". Respondents could indicate as many statements as they wanted. The statements were then added to a total score for each motive. The correlations were .18, .19, and .20 for achievement, power, and affiliation, respectively. A possible explanation for the low correlations may be that respondents have struggled to identify the motive in their story because scientific ideas on affiliation, achievement, and power motivation are complex and frequently counter to lay person theories. For instance, motivational researchers typically categorize some form of helping as a power theme while lay people typically categorize it as an affiliation theme (Hofer, Busch, Chasiotis, Kärtner, & Campos, 2008). Likewise, motivation because of a potential monetary reward is typically interpreted as achievement motivation by laypersons but not included in the scientific view of achievement motivation themes. Another explanation may be that stories in implicit motive tests are typically unique and picture specific and it may have been difficult to find a motivational statement fitting the individual story.

Motive Self-Categorization

In this study, we build on Engeser (2005) and investigate whether people can self-evaluate their implicit motives using a fourth strategy, the MSC. We developed picture specific statements (items) resembling typical responses to a picture that are directly linked to a motive, instead of broad and general motivational categories based on scientific definitions. We thereby avoid confounding the understanding of scientific definitions with the actual understanding. We asked respondents to write stories for all pictures and then to self-categorize each of their stories with a forced choice between the picture specific items. An example of the MSC procedure is given in Figure 1. Panel A shows the first part of the procedure where respondents imagine a story and then answer questions to summarize it. Panel B shows the self-categorization procedure.

Self-evaluation of implicit motives with the MSC has two potential advantages. First, the use of an operant response behavior format like traditional implicit motive measures ensured the same construct and not explicit or semi-implicit motives are measured (McClelland et al., 1989; Sokolowski et al., 2000). Second, the context specific items do not rely on abstract motive definitions and are closer to the actual stories than to broad and abstract motive definitions that were used by Engeser. Differences in how experts and lay persons interpret a motive should therefore not influence whether respondents can understand their own motives from their stories.

To comprehensively study our overarching research question whether people can effectively evaluate their implicit motives with the MSC, we focused on three more specific research objectives. First, we studied whether the MSC procedure actually leads to reliable self-categorizations. To study this objective, we used a simulation approach as an improved reliability estimate for implicit motive measures that is based on a recent IRT model (Lang, 2014). Our second objective was to study convergent validity of the MSC with expert codings of implicit motives. We investigated convergent validity by analyzing a MTMM matrix (Campbell & Fiske, 1959) and latent MTMM models. If the MSC measures implicit motives, there should be a reasonable overlap with expert codings and evidence for convergent validity. Our third objective was to study discriminant validity of the three motives and discriminant validity between the MSC and self-rated personality. We studied discriminant validity of the three motives by evaluating correlation between the motives in latent MTMM models. Implicit motives are typically seen as separate constructs (e.g. Schultheiss & Brunstein, 2001), so we expected to find low correlations. We studied discriminant validity of the MSC and a self-rated personality traits by analyzing correlations with the Big Five Inventory 2 (Soto & John, 2017). Possible overlap may occur because respondents score both tests themselves, thus, they share common measurement features. However, because traditional personality traits differ substantially from implicit motives, we expect little overlap.

Methods

Participants

We base our analysis on recently introduced IRT methodologies that allow to model multiple latent traits (Lang, 2014). Research suggests that the types of IRT models we used for

this present study require at least about 200 observations (Lang, Zettler, Ewen, & Hülsheger, 2012; Maydeu-Olivares & Brown, 2010). To ensure a sufficient sample size in case of a lower response rate, we decided to contact more than 200 participants. More specifically, four research assistants contacted 281 people in Belgium that participated in an earlier study with a community sample that was recruited through personal acquaintances. We received 247 complete questionnaires back. The 34 people who did not respond did not differ from the ones who responded. To reduce situational variance, we developed pictures that reflect typical situations in the daily life in an adult population. The pictures depict family, leisure, and work situations. We therefore did not contact children or adolescents. Participation was voluntarily, and no rewards were offered. Data were collected in accordance with the ethical guidelines for research involving human subjects by Ghent University, Belgium. The measures reported in this study were part of a data collection that also includes information on work relevant variables (job stress, burnout, work-life balance, job satisfaction, career satisfaction, and well-being).

Measures

Motive Self-Categorization. The first part of the MSC consists of an operant response format. We presented 15 drawings with ambiguous social scenes to the respondents. Respondents are instructed to freely imagine a story based on the picture and indicate a character depicted in the picture as the main character of their story. Then, they answer the following three questions to their story: (a) “What is important for the person in this situation and what is the person doing?” (b) “How does the person feel?” and (c) “Why does the person feel this way?”. Figure 1 panel A illustrates this procedure based on picture 6 in our study. The left panel shows a picture and the right panel shows questions and answers to an imaginative story written in response to this picture.

The second part consists of the actual self-categorization**.** After respondents had filled out the first part, we presented the picture specific self-categorization items and asked respondents to indicate a single item fitting best to the story they wrote for each picture in the first part. We presented between four and six different items for each picture and an option that none of the items fitted their story. To develop the items, we first collected a sample of 194 people responding to the 15 pictures with the same response process we later used in the

first part of the MSC. We then coded the stories into affiliation, achievement, and power content. For each picture, we looked at all responses that were coded for a specific motive and identified recurrent themes. We formulated one or more items for each story covering content typically written in response to these pictures. The items are also unambiguously linked to one of the three motives or to an *other* category to cover for not motive relevant recurrent content.

In research on implicit motives it is common to distinguish between approach and avoidance components for motives (Schultheiss, 2010). For instance, researchers differ between hope for success and fear of failure for achievement motivation (McClelland et al., 1953), hope for closeness and fear of rejection (Boyatzis, 1973), and hope for power and fear of weakness (Veroff & Veroff, 1972). IRT research has shown reliabilities are higher when only approach factors are modelled (Runge, Lang, Chasiotis, & Hofer, 2018). Building on these findings, we only use positive approach motivation and do not score affiliation when fear of rejection is the motive of the main actor of the story, achievement when anticipated failure is the motivation of a person in a story, and power when the person in a story is weak and submissive.

In Figure 1 Panel B, we give an example of the self-categorization part. The first option is linked to the affiliation motive, the second option is linked to the achievement motive, and the third and fourth options are linked to the power motive. The fifth and sixth answer options are avoidance motivation for affiliation and therefore not scored as a motive. The complete MSC test is provided as online supplemental material.

Expert Codings. To study self-evaluation of implicit motives, we compared the MSC with implicit motive codings by experts. Experts coded the stories written in step one of the self-categorization procedure with the standard coding procedure from the operant motive test coding manual (Kuhl & Scheffer, 2001). Each story got one coding, similar to the self-categorization. The coding includes the motivational category (affiliation, achievement, and power) and one of five modes of enactment. The modes of enactment indicate different approach and avoidance components. To match expert codings with the categories of the MSC, we integrated modes one, two, and three for affiliation and achievement, and mode one to four for power. Four Dutch native speakers coded the stories using the detailed information provided in the English translation of the operant motive test manual. The coders also received

an extensive two-day coding training where a high agreement with training material was reached. To ensure coding quality and evaluate interrater reliability, the first author of this study coded 20 tests for each of the four coders. To determine interrater agreement, we calculated Cohen's Kappa. Rater one coded 57 tests and showed good interrater agreement (.89, .88, and .86 for affiliation, achievement, and power), rater two coded 76 tests (.86, .89, and .88), rater three coded 56 tests (.91, .79, and .88), and rater four coded 58 tests (.82, .79, and .90). The interrater agreements were in the typical range for implicit motive measures (Köllner & Schultheiss, 2014).

Personality. Personality traits were measured with a Dutch translation of the BFI 2 (Soto & John, 2017), a brief measure (60 items) of the Big Five personality traits. Participants responded on a scale ranging from 1 (*very inaccurate*) to 5 (*very accurate*). Cronbach's *alpha* was .84, .75, .82, .88, and .84 for extraversion, agreeableness, conscientiousness, negative emotionality, and open-mindedness, respectively.

Analysis

Model Fitting. We analyze our three objectives based on a recently proposed IRT model for implicit motive measures (Lang, 2014). Modelling responses in implicit motive tests is not straight forward because researchers assume that motives compete for enactment which leads to a multidimensional response process (Atkinson & Birch, 1970). Researchers also found that people are less likely to respond with a specific motive when they have previously responded with that motive (Lang, 2014; Runge et al., 2016). Lang (2014) has developed the dynamic Thurstonian IRT model that accounts for the multidimensional response process and for the reduction in motive strength during assessment. We give a short overview of this dynamic Thurstonian IRT model and model estimation in a generalized linear mixed-effect model framework in the Appendix of this chapter. For some pictures in the MSC, people rarely write stories for a certain motive so that we did not make a self-categorization item. Researchers typically exclude those categories from model estimation (Debeer & Janssen, 2013; Runge et al., 2016). In this study, we excluded the affiliation category from the estimations of pictures 5, 11, and 13, as well as the achievement category from the estimation of pictures 2 and 7. As a first step of our analysis, we fitted basic dynamic Thurstonian IRT models for the MSC (Model 1) and for the expert codings (Model 2).

Objective 1: Reliability. In our first objective, we wanted to study the reliabilities of the MSC and expert codings. Researchers have shown that measures of internal consistency underestimate reliabilities in implicit motive measures because sum score approaches cannot model the response processes (Lang, 2014). The dynamic Thurstonian IRT model allows for advanced simulation-based measures of reliability. The goal of this simulation is to study how well traits are recovered in 100 simulated and bias corrected simulation runs. This procedure consists of three steps. First, one generates simulated data sets based on the model parameters in the Thurstonian IRT model from the empirical sample. Then, one fits the Thurstonian IRT model to these simulated data. Finally, one correlates the IRT estimates from the simulated and the true population to study how well traits are recovered. Those correlations are called fidelity correlations (Weiss, 1982), and the squared fidelity correlations equal the squared correlation reliability which is a common definition of reliability (Lord, Novick, & Birnbaum, 1968). We analyzed the IRT reliabilities based on the basic dynamic Thurstonian IRT models (Model 1 and 2). The procedure is described by Lang (2014) and an R tutorial is provided by Runge and colleagues (2016).

Objective 2: Convergent Validity with Expert Codings. In our second and main objective, we studied convergent validity of the MSC with expert codings in three ways. First, we analyzed latent correlations for the three motives between MSC and expert codings in a MTMM matrix. To obtain the correlations, we fitted a dynamic Thurstonian IRT model that includes MSC and expert codings with separate trait factors (Model 3).

Second, we tested convergent validity of the MSC and expert codings analyzing latent correlated traits, correlated methods (CT-CM) models. Latent CT-CM modelling is a flexible framework allowing researchers to explicitly test convergent validity (Eid, Geiser, & Koch, 2016) that has frequently been used to study different rating sources (Conway, 1996; Eid et al., 2016; Kenny, 1995; Kenny & Kashy, 1992; Kolk, Born, & Van Der Flier, 2002; Lance, Noble, & Scullen, 2002; Lievens & Conway, 2001). We used the dynamic Thurstonian IRT model as basis for the latent CT-CM models. Using a latent modelling approach for implicit motive measures has the additional advantage to account for the multidimensional and dynamic response process. To test convergent validity, we evaluated the amount of trait and method variance. Convergent validity is supported when trait variance is higher than method variance.

We fitted three different CT-CM models because they can be difficult to interpret. First, we fitted a standard CT-CM model (Model 4) described in Figure 2 (Widaman, 1985). Second, we also estimated the model using a new approach that uses conventional maximum likelihood estimation but adds a Bayesian prior to stabilize estimates (Chung, Rabe-Hesketh, Dorie, Gelman, & Liu, 2013). We estimated the model using the blme package in R with the default wishart prior (Model 5). Third, we fitted the CT-C(M-1) model (Model 6) by Eid (2000). The CT-C(M-1) model parametrizes only M-1 method factors and is frequently applied when parameters are difficult to interpret (Eid et al., 2016). This model typically uses minimum three traits and three methods but can also be modeled with only two methods. However, because only one of the two methods are parameterized, no correlations between the methods are estimated.

Third, CT-CM modelling also allows to study convergent validity with a formal test (Widaman, 1985). This test is based on the comparisons of the goodness of fit of two nested models. The first model is restricted to only include correlated methods but not traits, a no traits, correlated method model (Model 7). The second model is the full CT-CM model (Model 4). When the CT-CM model fits better than the no traits, correlated method model, it indicates that the latent motive factors explain important proportions of covariation (common trait variance). Thus, traits measured with both methods are related and convergent validity is supported.

Objective 3: Discriminant Validity between Motives and with Self-Ratings of Personality

We studied discriminant validity between the motives with correlations between the latent motives in the three CT-CM models (Models 4, 5, and 6). We studied discriminant validity between latent MSC scores estimated from Model 1 and self-ratings of the big five personality traits measured with the BFI 2 in a MTMM matrix. Low correlations would support discriminant validity in both analyses.

Results

We started our analysis by fitting the basic dynamic Thurstonian IRT models for the MSC (Model 1) and for the expert codings (Model 2). Both models converged, and Table 1 shows the fit statistics for them. Table 2 shows the parameter estimates for the fixed effects for both models. The estimates for each picture and motive show how likely a response with the

respective type of motive is. All three dynamic effects in each model are negative indicating that a previous story with content for this motive lowers the probability of following responses.

Objective 1: Reliability

Table 3 provides reliability estimates for the MSC (based on Model 1) and expert codings (based on Model 2). The IRT based reliability estimates were higher compared to Cronbach's alpha. This is in line with findings in the literature (Lang, 2014; Runge et al., 2016, 2018). The reliabilities for power and achievement were similar for the MSC and expert codings, for affiliation the MSC showed a higher reliability. These results suggest scores on the MSC show evidence of reliability.

Objective 2: Convergent Validity with Expert Codings

Table 4 shows the fit statistics for all models that we used to study convergent validity. All models converged. In the first analysis for convergent validity, we expected to find substantial correlations between implicit motives measured with the MSC and with expert codings. This would imply that respondents can categorize their own implicit motives. Table 5 shows a MTMM correlation table including latent motives from MSC and expert codings (Model 3). MSC and expert codings correlated r = .60, r =.37, and r =.62 for affiliation, achievement, and power, respectively. These findings indicate that the MSC and expert codings are related, stronger for affiliation and power and weaker for achievement.

Second, we analyzed trait and method variances in three CT-CM models (Models 4, 5, and 6). Table 6 provides the standard deviations for trait and method effects for all three models. The standard deviations for the traits are considerably higher than for the methods in all three models[13], thus, all models supported convergent validity of the MSC and expert codings. The correlation between MSC and expert codings in the standard CT-CM (Model 4) were r = -1.0 (note that they are non-directional). This phenomenon is common in mixed effects models when the variance of random effects is low and, thus, the information for determining the correlation is limited (Bates, Mächler, Bolker, & Walker, 2015; Chung et al.,

[13] For Model 6, we fitted a CT-C(M-1) model that parameterized only the MSC. We have also fitted a CT-C(M-1) model that parameterized only the expert codings (AIC = 22,029; BIC = 22,762; logLik = -10,922). Correlations and standard deviations were highly similar.

2013). In the present model, this situation occurred because the traits/motives were responsible for most of the variance leaving only a limited amount of variance to the methods.

Third, we performed a formal test to study convergent validity. Table 4 provides the fit statistics and the model comparison for the no traits, correlated method model (Model 7) and the standard CT-CM model (Model 4). A lower log-likelihood and lower information criterions (AIC and BIC) indicate a better model fit. The standard CT-CM model (AIC = 21,965; BIC = 22,714) provided a better model fit than the no traits, correlated method model (AIC = 22,968; BIC = 23,670), giving evidence that the latent motive factors explain important proportions of covariation beyond what is explained by method factors. Thus, all three analyses show that MSC and expert codings are related and that convergent validity is supported.

Objective 3: Discriminant Validity between Motives and with Self-Ratings of Personality

In our third objective, we analyzed discriminant validity between the latent motives by interpreting the latent correlations in all three CT-CM models (Models 4, 5, and 6). Table 6 shows that the motives correlate low in all three models, as we expected. These findings are also in line with earlier IRT research that shows that correlations between motives are low (Lang, 2014; Lang et al., 2012; Runge et al., 2016, 2018). Additionally, we analyzed discriminant validity between the MSC and self-ratings of personality measured with the BFI 2. We found that correlations between the MSC and personality traits were generally low (Table 7). This finding indicates that there is little common method variance between MSC and self-rated personality. Thus, the MSC does not measure personality traits. These results provide support for the discriminant validity of scores on the MSC.

Discussion

The aim of this study was to investigate whether people can self-evaluate their implicit motives with a new measurement approach using empirical based and tailored items to self-categorize stories for motivational content. The main finding of our study is that MSC and expert codings of implicit motives are substantially related. We conclude that people can evaluate their implicit motives. The correlations are substantial but not extremely high indicating a discrepancy in how respondents and how experts evaluate their motivational

themes. Additional analysis showed that the MSC has a similar reliability than expert codings. The three motives are distinct from each other as well as from self-ratings of personality traits.

Theoretical Implications

The main theoretical implication of this study is that motivational researchers may need to scrutinize the assumption that people cannot evaluate their implicit motives. Although our study did not show a perfect overlap between the MSC and expert codings of implicit motives, we could show that they are substantially related. Research on personality has also shown that self and other-rated personality do not correlate very high (Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2004; Vazire & Carlson, 2010). One reason is that other raters do not have perfect insight into the person's lives. For instance, in typical personality measures, people rate average behavior but other raters do not see the full range of the behaviors of the rated person (Vazire & Carlson, 2010). For implicit motive measures the basis—the written stories—is the same for other and self-categorization. However, the stories are frequently ambiguous and not very detailed so expert raters have less insight into the motives of the characters in the story compared to the person who wrote the story. A more differentiated explanation of why self and observer ratings typically do not correlate highly has recently be suggested by McAbee and Connelly (2016). They separate variances into agreement between self and other ratings (trait variance), self-only variance (identity), and observer only variance (reputation). In this model, non-agreement is not considered as error variance but as unique contribution of different rating sources to form a more detailed model of personality.

Our findings support the few studies suggesting that people may be able to evaluate their implicit motives themselves. Sherwood (1966) has shown that people can evaluate their implicit motives after a comprehensive education on motives. Researchers have also argued that one can ask people for their motives with questionnaires and that low correlations between implicit and explicit motives are due to methodological factors like low reliability (Thrash et al., 2010; Thrash, Maruskin, & Martin, 2012). Generally, however, few studies support the idea that one can directly ask people for their implicit motives. Theoretical arguments (McClelland et al., 1989) and previous research has shown that respondent

response formats do not measure implicit motives (Köllner & Schultheiss, 2014; Schüler et al., 2015).

In this study, we suggest another route to self-evaluation of implicit motives and show that it can be a useful approach to study motivation. We found implicit motives can be accessed indirectly with the same operant response format that researchers use with expert codings. In contrast to Engeser (2005) who used a similar procedure, the MSC uses items specifically tailored to each picture. We showed that respondents can better categorize their stories when the items are contextualized and specific for each motive instead of broad and abstract categories, which gives evidence that differences in scientific definitions and lay people understanding of implicit motives may have biased research on self-evaluation of implicit motives in the past.

Researchers have long debated whether implicit motives and implicit measures in general show sufficient criterion validity (Greenwald, Poehlman, Uhlmann, & Banaji, 2009; Ziegler et al., 2010). There is a large body of evidence that implicit motive measures can predict outside criterions such as career success over many years (e.g. Winter et al., 1998), hormonal correlates for power and affiliation motivation (Schultheiss, Wirth, & Stanton, 2004), functioning of romantic partnerships (Hagemeyer, Schönbrodt, Neyer, Neberich, & Asendorpf, 2015), and work performance (Lang et al., 2012). However, researchers have also not always found consistent correlations with outside measures and correlations can be low (Spangler, 1992; Ziegler et al., 2010). One obstacle in systematic research on predicting outcomes for implicit motives has been that coding stories is relatively time intensive and, thus, less economic. With the MSC, we introduce a tool to study implicit motives more economically. Future research may use the MSC or develop similar tests with this technique to study predictive validity.

Practical Implications

It is widely acknowledged that researchers in psychology should use multiple methods to measure constructs in empirical research (Campbell & Fiske, 1959). For implicit motives, most research has relied on expert codings of stories written by respondents. The primary goal of our study was to investigate whether respondents can evaluate the motivational content of the stories they write with a new measurement approach. We found that the MSC is measuring

implicit motives equally reliable than expert codings and distinct from self-ratings of personality traits. Thus, this study opens the possibility to broaden the perspectives for researchers and practitioners that are interested in implicit motives. One can use the MSC to study differences in self-evaluated and expert coded implicit motives, but also as an alternative way to measure implicit motives.

Compared to expert codings, the MSC has several advantages. A first advantage is that people who rate their own stories have better insight in the motives of the character in their story which is important when stories are ambiguous or very short. A second advantage is that the scoring objectivity is given because all items are unambiguously linked to a motive category. However, expert codings typically also have a sufficient scoring objectivity (e.g. Pang & Schultheiss, 2005). A third advantage is that the MSC is economical because coding stories is labor intensive and expert knowledge is required.

Limitations

Intuitively, one concern with the MSC could be that the items do not cover all possible stories. It could happen that respondents write a story with motivational content that is not covered by the self-categorization items. Respondents would then indicate that no item fitted their story. Expert coders may still be able to code those stories. However, in this actual data set this is rarely the case. Respondents indicated in only 5.60% of the stories that no item was fitting when experts coded a motive. Studies using this procedure in the future could add more self-categorization items to cover more content, however, to use more items also increases the complexity and length of the test.

A second limitation of the MSC is that the time to take the test increases. In our experience, the additional time is on average around five minutes. Researchers who want to include self-categorizations of implicit motives should therefore account for the increased test length.

A third limitation of the study is that the MSC shows an IRT reliability that may be too low for single case assessment especially for power. The IRT reliabilities are, however, similar compared to other IRT reliabilities for picture based implicit motive measures reported in the literature (Lang, 2014; Lang et al., 2012; Runge et al., 2016, 2018) and may be useful for research purposes (Ellis, 2013). Still, higher reliabilities are certainly desirable. Researchers

have suggested to study test length as a possibility to increase reliability. For the picture story exercise researchers have recommended to use at least four or five pictures (Pang, 2010; Schultheiss & Pang, 2007). The MSC uses more pictures but responses are shorter, and respondents give one coding for each picture so that the recommendations in the literature are not directly applicable. Reliability does also not only depend on test length, additional factors like motive variance, cue ambiguity, and dynamic effects may also influence reliability (Reitman & Atkinson, 1958; Runge et al., 2018). Future studies should systematically study possible effects of cue strength (item difficulty), cue ambiguity, the number of traits measured with a picture, and test length on the reliability of the MSC.

Future Research

A challenging and unanswered question in research on implicit motives has been that correlations between different implicit motive measures are typically low (Schüler et al., 2015; Ziegler et al., 2010). Possible theoretical explanations include that implicit motive tests measure different aspects of an implicit motive, that moderators influence the relationship between the different implicit measures (Ziegler et al., 2010), or that the measures may be domain specific (Schultheiss, Liening, & Schad, 2008). A possible methodological explanation is that implicit measures typically show a lower reliability which may also lower correlations among those measures (Thrash et al., 2010, 2012). This study does not contribute to the understanding of this issue. Moreover, the MSC and the possibility to self-categorize implicit motives introduces a new measurement procedure and with that the question whether it correlates with other implicit motive tests like the PSE. We encourage researchers to study more systematically under which circumstances implicit measures, including the MSC, correlate.

In this paper, we developed a self-categorization procedure with a brief response format. However, there are other popular approaches and there is some evidence that implicit motives from these approaches do not always capture the same motive as discussed above. Another interesting question for future research is therefore, whether the self-categorization procedure can also be applied to other picture-based procedures with a long response format like the PSE (Pang & Schultheiss, 2005; Schultheiss & Brunstein, 2001). The PSE asks for longer stories and allows to code more than one motive as well as to code motives more than

once in a story. It may be interesting to investigate whether the findings in this study generalize to this response format. One could allow respondents to choose more than one option to cover different motives that may show up in longer stories. Future research should address the question whether the response process we suggested in this study can also be implemented in other response formats of implicit motive tests.

Conclusion

We studied whether people can evaluate their implicit motives with the MSC—a test that builds on the idea of picture specific self-categorization items—and found a substantial overlap of latent motive scores with expert codings and no overlap with self-ratings of big five personality traits. These findings suggest that people can have some insight into their implicit motives. The MSC may be used in future research as an alternative measure of implicit motives with a different rating source than expert raters.

References

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64*. http://doi.org/10.1037/h0043445

Atkinson, J. W., & Birch, D. (1970). *Dynamics of action*. New York: Wiley.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 1*. http://doi.org/10.18637/jss.v067.i01

Bates, D., Maechler, M., & Bolker., B. (2011). lme4: linear mixed-effects models using S4 classes. R package version 0.999375-41. *R*. Retrieved from http://lme4.r-forge.r-project.org/

Baumann, N., Kaschel, R., & Kuhl, J. (2005). Striving for unwanted goals: Stress-dependent discrepancies between explicit and implicit achievement motives reduce subjective well-being and increase psychosomatic symptoms. *Journal of Personality and Social Psychology, 89*, 781–799. http://doi.org/10.1037/0022-3514.89.5.781

Böckenholt, U. (2001). Hierarchical modeling of paired comparison data. *Psychological Methods, 6*, 49–66. http://doi.org/10.1037//1082-989X.6.1.49

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology, 86*, 599–614. http://doi.org/10.1037/0022-3514.86.4.599

Boyatzis, R. E. (1973). Affiliation motivation. In D. C. McClelland & R. S. Steele (Eds.), *Human motivation -- a book of readings* (pp. 252–276). Morristown, NJ: General Learning Corporation.

Brunstein, J. C., & Schmitt, C. H. (2004). Assessing individual differences in achievement motivation with the implicit association test. *Journal of Research in Personality, 38*, 536–555. http://doi.org/10.1016/j.jrp.2004.01.003

Burke, B. L., Arkowitz, H., & Menchola, M. (2003). The efficacy of motivational interviewing: A meta-analysis of controlled clinical trials. *Journal of Consulting and Clinical Psychology, 71*, 843–861. http://doi.org/10.1037/0022-006X.71.5.843

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105.

http://doi.org/10.1037/h0046016

Chung, Y., Rabe-Hesketh, S., Dorie, V., Gelman, A., & Liu, J. (2013). A nondegenerate penalized likelihood estimator for variance parameters in multilevel models. *Psychometrika*, *78*, 685–709. http://doi.org/10.1007/s11336-013-9328-2

Conway, J. M. (1996). Analysis and design of multitrait-multirater performance appraisal studies. *Journal of Management*, *22*, 139–162. http://doi.org/10.1177/014920639602200106

De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., Tuerlinckx, F., & Partchev, I. (2011). The estimation of item response models with the lmer function from the lme4 package in R. *Journal Of Statistical Software*, *39*, 1–28. http://doi.org/10.18637/jss.v039.i12

Debeer, D., & Janssen, R. (2013). Modeling item-position effects within an IRT framework. *Journal of Educational Measurement*, *50*, 164–185.

Eid, M. (2000). A multitrait-multimethod model with minimal assumptions. *Psychometrika*, *65*, 241–261. http://doi.org/10.1007/BF02294377

Eid, M., Geiser, C., & Koch, T. (2016). Measuring method effects. *Current Directions in Psychological Science*, *25*, 275–280. http://doi.org/10.1177/0963721416649624

Ellis, J. L. (2013). A standard for test reliability in group research. *Behavior Research Methods*, *45*, 16–24. http://doi.org/10.3758/s13428-012-0223-z

Engeser, S. (2005). *Lernmotivation und volitionale Handlungssteuerung: Eine Längsschnittsuntersuchung beim Statistik Lernen im Psychologiestudium [Motivation to learn and volitional action control: A longitudinal study of learning statistics in a psychology course]*. University of Potsdam. Retrieved from https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/deliver/index/docId/217/file/ENGESER.PDF

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–80. http://doi.org/10.1037/0022-3514.74.6.1464

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *1*, 17–41. http://doi.org/10.1037/a0015575

Hagemeyer, B., & Neyer, F. J. (2011). Assessing implicit motivational orientations in couple relationships: The Partner-Related Agency and Communion Test (PACT). *Psychological Assessment, 24*, 114–128. http://doi.org/10.1037/a0024822

Hagemeyer, B., Schönbrodt, F. D., Neyer, F. J., Neberich, W., & Asendorpf, J. B. (2015). When 'together' means 'too close': agency motives and relationship functioning in coresident and living-apart-together couples. *Journal of Personality and Social Psychology*, *109*, 813–835. http://doi.org/10.1037/pspi0000031

Heckhausen, H. (1963). *Hoffnung und Furcht in der Leistungsmotivation [Hope and fear components of achievement motivation]*. Meisenheim am Glan: Anton Hain.

Heyns, R. W., Veroff, J., & Atkinson, J. W. (1958). A scoring manual for the affiliation motive. In J. W. Atkinson (Ed.), *Motives in Fantasy, Action, and Society* (pp. 208–218). Princeton, New Jersey: Van Nostrand.

Hofer, J., Busch, H., Chasiotis, A., Kärtner, J., & Campos, D. (2008). Concern for generativity and its relation to implicit pro-social power motivation, generative goals, and satisfaction with life: A cross-cultural investigation. *Journal of Personality*, *76*, 1–30. http://doi.org/10.1111/j.1467-6494.2007.00478.x

Jackson, D. N. (1984). *Manual for the personality research form (3rd edition).* Goshen: Research Psychologists Press.

James, L. R. (1998). Measurement of personality via conditional reasoning. *Organizational Research Methods*, *1*, 131–163. http://doi.org/10.1177/109442819812001

Kenny, D. A. (1995). The multitrait-multimethod matrix: design, analysis, and conceptual issues. In P. E. Shrout & S. T. Fiske (Eds.), *Personality research, methods, and theory: A festschrift honoring Donald W. Fiske* (pp. 111–124). Hillsdale, NJ: Erlbaum.

Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin*, *112*, 165–172. http://doi.org/10.1037/0033-2909.112.1.165

Kolk, N. J., Born, M. P., & Van Der Flier, H. (2002). Impact of common rater variance on construct validity of assessment center dimension judgments. *Human Performance*, *15*, 325–337.

Köllner, M. G., & Schultheiss, O. C. (2014). Meta-analytic evidence of low convergence

between implicit and explicit measures of the needs for achievement, affiliation, and power. *Frontiers in Psychology*, *5*, 826. http://doi.org/10.3389/fpsyg.2014.00826

Kuhl, J., & Scheffer, D. (2001). Der operante multi-motiv-test (OMT): Manual [The operant multi-motive-test (OMT): Manual]. Osnabrück, Germany: University of Osnabrück.

Lance, C. E., Noble, C. L., & Scullen, S. E. (2002). A critique of the correlated trait-correlated method and correlated uniqueness models for multitrait-multimethod data. *Psychological Methods*, *7*, 228–244. http://doi.org/10.1037/1082-989X.7.2.228

Lang, J. W. B. (2014). A dynamic Thurstonian item response theory of motive expression in the picture story exercise: Solving the internal consistency paradox of the PSE. *Psychological Review*, *121*, 481–500. http://doi.org/10.1037/a0037011

Lang, J. W. B., Zettler, I., Ewen, C., & Hülsheger, U. R. (2012). Implicit motives, explicit traits, and task and contextual performance at work. *Journal of Applied Psychology*, *97*, 1201–1217. http://doi.org/10.1037/a0029556

Lievens, F., & Conway, J. M. (2001). Dimension and exercise variance in assessment center scores: a large-scale evaluation of multitrait-multimethod studies. *Journal of Applied Psychology*, *86*, 1202–22.

Lord, F. M., Novick, M. R., & Birnbaum, A. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Maydeu-Olivares, A., & Brown, A. (2010). Item response modeling of paired comparison and ranking data. *Multivariate Behavioral Research*, *45*, 935–974. http://doi.org/10.1080/00273171.2010.531231org/10.1080/00273171.2010.531231

McAbee, S. T., & Connelly, B. S. (2016). A multi-rater framework for studying personality: The trait-reputation-identity model. *Psychological Review*, *123*, 569–591. http://doi.org/10.1037/rev0000035

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). The measuring instrument. In *The Achievement Motive* (pp. 185–217). East Norwalk, CT, US: Appleton-Century-Crofts.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*, 690–702. http://doi.org/10.1037/0033-295X.96.4.690

Mihura, J. L., Meyer, G. J., Dumitrascu, N., & Bombel, G. (2013). The validity of individual Rorschach variables: Systematic reviews and meta-analyses of the comprehensive system. *Psychological Bulletin, 139*, 548–605. http://doi.org/10.1037/a0029406

Morgan, C. D., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology & Psychiatry, 34*, 289–306. http://doi.org/10.1001/archneurpsyc.1935.02250200049005

Murray, H. A. (1938). *Explorations in Personality; A clinical and Experimental Study of Fifty Men of College Age.* New York: Oxfort University Press. http://doi.org/10.2307/2084329

Pang, J. S. (2006). *A revised content-coding measure for hope of success(HS) and fear of failure (FF). Unpublished Dissertation.* University of Michigan, Ann Arbor.

Pang, J. S. (2010). Content coding methods in implicit motive assessment: Standards of measurement and best practices for the picture story exercise. In O. C. Schultheiss & J. C. Brunstein (Eds.), *Implicit Motives* (pp. 119–1550). New York: Oxford University Press.

Pang, J. S., & Schultheiss, O. C. (2005). Assessing implicit motives in U.S. college students: Effects of picture type and position, gender and ethnicity, and cross-cultural comparisons. *Journal of Personality Assessment, 85*, 280–294. http://doi.org/10.1207/s15327752jpa8503_04

R Core Team. (2018). R: A language and environment for computing. Vienna, Austria: R Foundation for Statistical Computing.

Reitman, W. R., & Atkinson, J. W. (1958). Some methodological problems in the use of thematic apperceptive measures of human motives. In J. W. Atkinson (Ed.), *Motives in fantasy, action, and society* (pp. 664–683). Princeton N.J.: Van Nostrand.

Runge, J. M., Lang, J. W. B., Chasiotis, A., & Hofer, J. (2018). Improving the assessment of implicit motives using IRT: cultural differences and differential item functioning. *Journal of Personality Assessment*, 1–11. http://doi.org/10.1080/00223891.2017.1418748

Runge, J. M., Lang, J. W. B., Engeser, S., Schüler, J., den Hartog, S. C. S. C., & Zettler, I. (2016). Modeling motive activation in the operant motives test: A psychometric analysis using dynamic Thurstonian item response theory. *Motivation Science, 2*, 268–286. http://doi.org/10.1037/mot0000041

Schmalt, H. D. (1999). Assessing the achievement motive using the grid technique. *Journal of*

Research in Personality, *33*, 109–130. http://doi.org/10.1006/jrpe.1999.2245

Schüler, J., Brandstätter, V., Wegner, M., & Baumann, N. (2015). Testing the convergent and discriminant validity of three implicit motive measures: PSE, OMT, and MMG. *Motivation and Emotion*, *39*, 839–857. http://doi.org/10.1007/s11031-015-9502-1

Schultheiss, O. C. (2010). Implicit motives. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of Personality: Theory and Research* (pp. 603–633). New York: Guilford Press.

Schultheiss, O. C., & Brunstein, J. C. (2001). Assessment of implicit motives with a research version of the TAT: picture profiles, gender differences, and relations to other personality measures. *Journal of Personality Assessment*, *77*, 71–86. http://doi.org/10.1207/S15327752JPA7701_05

Schultheiss, O. C., Liening, S. H., & Schad, D. (2008). The reliability of a picture story exercise measure of implicit motives: Estimates of internal consistency, retest reliability, and ipsative stability. *Journal of Research in Personality*, *42*, 1560–1571. http://doi.org/10.1016/j.jrp.2008.07.008

Schultheiss, O. C., & Pang, J. S. (2007). Measuring implicit motives. In R. W. Robins, R. C. Fraley, & R. Krueger (Eds.), *Handbook of Research Methods in Personality Psychology* (pp. 322–344). New York: Guilford.

Schultheiss, O. C., Wirth, M. M., & Stanton, S. J. (2004). Effects of affiliation and power motivation arousal on salivary progesterone and testosterone. *Hormones and Behaviorehavior*, *46*, 592–599.

Sherwood, J. J. (1966). Self-report and projective measures of achievement and affiliation. *Journal of Consulting Psychology*, *30*, 329–337. http://doi.org/10.1037/h0023574

Smith, C. P. (1992). *Motivation and personality: Handbook of thematic content analysis*. New York: Cambridge University Press.

Sokolowski, K., Schmalt, H. D., Langens, T. A., & Puca, R. M. (2000). Assessing achievement, affiliation, and power motives all at once: The multi-motive grid (MMG). *Journal of Personality Assessment*, *74*, 126–145. http://doi.org/10.1207/S15327752JPA740109

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and

assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology, 113*, 117–143. http://doi.org/10.1037/pspp0000096

Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin, 112,* 140–154. http://doi.org/10.1037/0033-2909.112.1.140

Thrash, T. M., Cassidy, S. E., Maruskin, L. A., & Elliot, A. J. (2010). Factors that influence the relation between implicit and explicit motives : A general implicit – explicit congruence framework. In O. C. Schultheiss & J. Brunstein (Eds.), *Implicit Motives* (pp. 308–346). New York: Oxford University Press. http://doi.org/10.1093/acprof

Thrash, T. M., Maruskin, L. A., & Martin, C. C. (2012). Implicit–explicit motive congruence. In R. M. Ryan (Ed.), *The Oxford Handbook of Human Motivation*. New York: Oxford University Press.

Vazire, S., & Carlson, E. N. (2010). Self-knowledge of personality: Do people know themselves? *Social and Personality Psychology Compass, 4*, 605–620. http://doi.org/10.1111/j.1751-9004.2010.00280.x

Veroff, J., & Veroff, J. B. (1972). Reconsideration of a measure of power motivation. *Psychological Bulletin, 78,* 279–291. http://doi.org/10.1037/h0033184

Weinberger, J., Cotler, T., & Fishmen, D. (2010). Clinical implications of implicit motives. In O. C. Schultheiss & J. C. Brunstein (Eds.), *Implicit Motives* (p. 468487). New York: Oxford University Press.

Weiss, D. J. (1982). Improving measurement quality and efficiency with adaptive testing. *Applied Psychological Measurement, 6,* 473–492. http://doi.org/10.1177/014662168200600408

Westen, D., Lohr, N., Silk, K. R., Gold, L., & Kerber, K. (1990). Object relations and social cognition in borderlines, major depressives, and normals: A Thematic Apperception Test analysis. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*. http://doi.org/10.1037/1040-3590.2.4.355

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement, 9,* 1–26.

http://doi.org/10.1177/014662168500900101

Wilson, T. D., & Dunn, E. W. (2004). Self-knowledge: Its limits, value, and potential for improvement. *Annual Review of Psychology, 55*, 493–518. http://doi.org/10.1146/annurev.psych.55.090902.141954

Winter, D. G. (1973). *The power motive*. New York, NY, US: Free Press.

Winter, D. G. (1994). *Manual for scoring motive imagery in running text*. Ann Arbor: University of Michigan.

Winter, D. G., John, O. P., Stewart, A. J., Klohnen, E. C., & Duncan, L. E. (1998). Traits and motives: Toward an integration of two traditions in personality research. *Psychological Review, 105*, 230–250. http://doi.org/10.1037/0033-295X.105.2.230

Ziegler, M., Schmukle, S., Egloff, B., & Bühner, M. (2010). Investigating measures of achievement motivation(s). *Journal of Individual Differences, 31*, 15–21. http://doi.org/10.1027/1614-0001/a000002

Tables

Table 1

Fit Statistics for the Basic Dynamic Thurstonian IRT Models

Model	*df*	*logLik*	Devianc	AIC	BIC
Model 1: MSC					
Basic Dynamic Thurstonian Model	43	-5,518	11,036	11,134	11,487
Model 2: Expert codings					
Basic Dynamic Thurstonian Model	43	-5,180	10,359	10,457	10,809

Note. MSC: k = 9,880 pair-wise comparisons nested in N = 247 persons and 15 pictures.

Expert coding: k = 9,724 pair-wise comparisons nested in N = 247 persons and 15 pictures.

AIC = Akaike information criterion; BIC = Schwartz information criterion.

Table 2

Parameter Estimates for the Basic Thurstonian IRT Models

	MSC			Expert codings		
Estimate	Aff	Ach	Pow	Aff	Ach	Pow
Fixed effects (β)						
Picture 1	-0.36	0.81	0.76	-1.43	0.51	0.58
Picture 2	-0.74		-0.58	-0.61		-0.07
Picture 3	-0.91	-0.05	0.91	-1.79	-0.55	0.11
Picture 4	-0.07	0.25	-0.46	0.09	0.20	0.01
Picture 5		-0.01	-0.63		0.23	-0.39
Picture 6	0.15	-0.03	0.26	0.60	-0.58	0.26
Picture 7	-0.65		0.66	-1.18		0.31
Picture 8	1.29	1.49	1.19	0.46	2.08	1.30
Picture 9	-0.01	0.59	0.81	0.71	0.94	1.10
Picture 10	0.80	0.66	-0.04	-0.28	0.03	0.49
Picture 11		1.26	0.19		0.77	0.12
Picture 12	1.37	1.33	1.88	0.39	1.32	1.84
Picture 13		1.37	1.06		1.13	1.31
Picture 14	1.52	1.90	0.84	1.09	1.64	0.86
Picture 15	0.86	1.19	0.21	0.05	1.56	0.30
Dynamic Effect	-0.67	-0.66	-0.17	-0.60	-0.44	-0.15

Note. MSC: k = 9,880 pair-wise comparisons nested in N = 247 persons and 15 pictures. Expert coding: k = 9,724 pair-wise comparisons nested in N = 247 persons and 15 pictures. MSC = Motive self-categorization. Aff = affiliation vs other. Ach = achievement vs other. Pow = power vs other.

Table 3

Squared Correlation Reliability ($r^2_{\theta\hat{\theta}}$) and Cronbach's Alpha in Bias-Corrected Simulated Data

Motive	MSC		Expert codings	
	$r^2_{\theta\hat{\theta}}$	Cronbach's α	$r^2_{\theta\hat{\theta}}$	Cronbach's α
Affiliation	.67	.26	.44	.40
Achievement	.61	.11	.65	.42
Power	.42	.42	.49	.49

Note. MSC = Motive self-categorization. $r^2_{\theta\hat{\theta}}$ = squared correlation reliability estimated by squaring the correlation between the true underlying latent motives θ and the estimated maximum a-posteriori estimates of the latent motives $\hat{\theta}$.

Table 4

Fit Statistics and Model Comparison for latent Multitrait-Multimethod Models

Model	*df*	*logLik*	Deviance	AIC	BIC	Δ*d*	χ^2
Model 3: MSC – Expert codings with separate trait factors	107	-10,597	21,193	21,408	22,201		
Model 4: Correlated traits, correlated methods	95	-10,888	21,775	21,965	22,714		
Model 5: Correlated traits, correlated methods with Bayesian Prior	95	-10,888	21,776	21,966	22,715		
Model 6: Correlated traits, correlated (methods -1)	93	-10,920	21,840	22,026	22,760		
Model 7: No trait, correlated methods 0T-CM	89	-11,395	22,790	22,968	23,670		
Comparison Model 4 vs. Model 7						6	1014.80*

Note. k = 19,604 pair-wise comparisons nested in N = 247 persons and 15 pictures. AIC = Akaike information criterion; BIC = Schwartz information criterion. * $p < .0001$.

Table 5

Latent Multitrait-Multimethod Correlations Between MSC and Expert Codings From Model 3

	MSC			Expert codings		
	1	2	3	4	5	6
MSC						
1.Affiliation	-					
2. Achievement	.07	-				
3. Power	.17	.13	-			
Expert codings						
4. Affiliation	**.60**	.00	.05	-		
5. Achievement	.14	**.37**	.00	.24	-	
6. Power	.03	.02	**.62**	.17	.21	-

Note. N = 247. MSC = Motive self-categorization. Bold values indicate correlations between self-categorization and expert codings for the three implicit motives (validity diagonal).

Table 6

Latent Correlations and Standard Deviations From Different Correlated Traits, Correlated Methods Models

	Traits				Method			
	Aff	Ach	Pow	*SD*		Expert	MSC	*SD*
Correlated traits, correlated methods (Model 4)								
Affiliation	-			0.93	Expert	-		0.22
Achievement	.34	-		0.56	MSC	-1.0	-	0.19
Power	.27	.36	-	0.52				
Correlated traits, correlated methods with Bayesian prior (Model 5)								
Affiliation	-			0.95	Expert	-		0.23
Achievement	.34	-		0.57	MSC	-0.95	-	0.20
Power	.27	.36	-	0.52				
Correlated traits, correlated (methods -1) MSC (Model 6)								
Affiliation	-			0.91	Expert	-		-
Achievement	.35	-		0.55	MSC		-	0.30
Power	.28	.37	-	0.52				

Note. N = 247. MSC = Motive Self Categorization. CT-CM = Correlated traits, correlated methods.

Table 7

Multitrait-Multimethod Correlations with MSC and Big Five Traits

	MSC			Big five traits			
	1	2	3	4	5	6	7
MSC							
1. Affiliation	-						
2. Achievement	.02	-					
3. Power	.10	.04	-				
Big Five Traits							
4. Extraversion	.04	.12	.05	-			
5. Conscientiousness	.16	-.09	.05	-.03	-		
6. Agreeableness	.06	-.07	.14	.14	.19	-	
7. Negative emotionality	.00	-.06	.03	-.26	-.11	-.24	-
8. Open-mindedness	.08	.00	-.17	.25	.11	.01	-.12

Note. $N = 245$. MSC = Motive self-categorization. MSC scores are latent scores from Model 1.

Figures

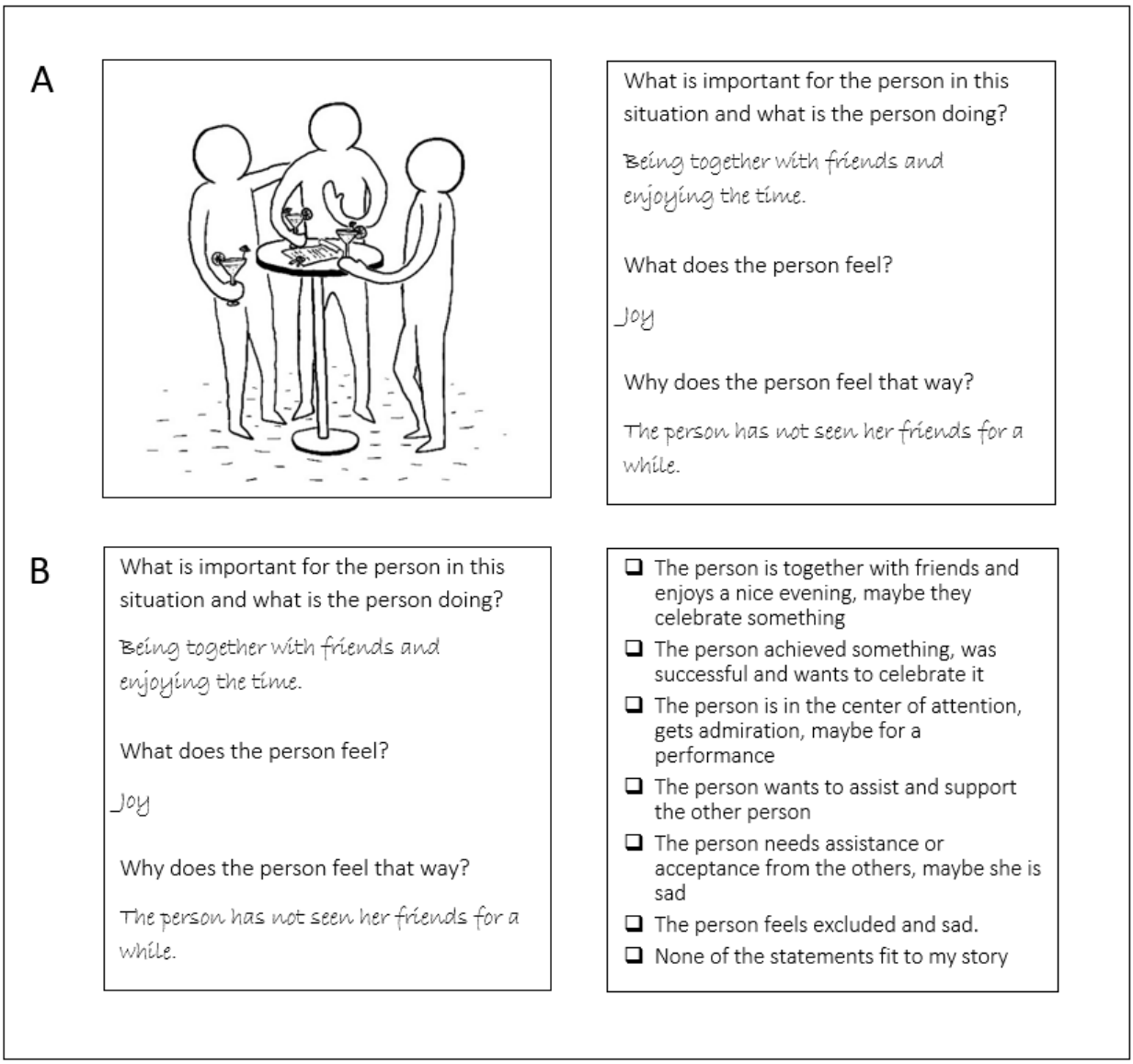

Figure 1. The upper two boxes (A) represent the standard procedure of implicit motives tests where people see a picture (upper left box) and write a story using questions (upper right box). The lower two boxes (B) represent the MSC procedure in which people go back to their original story (lower left box) and choose the best fitting item for their story (lower right box).

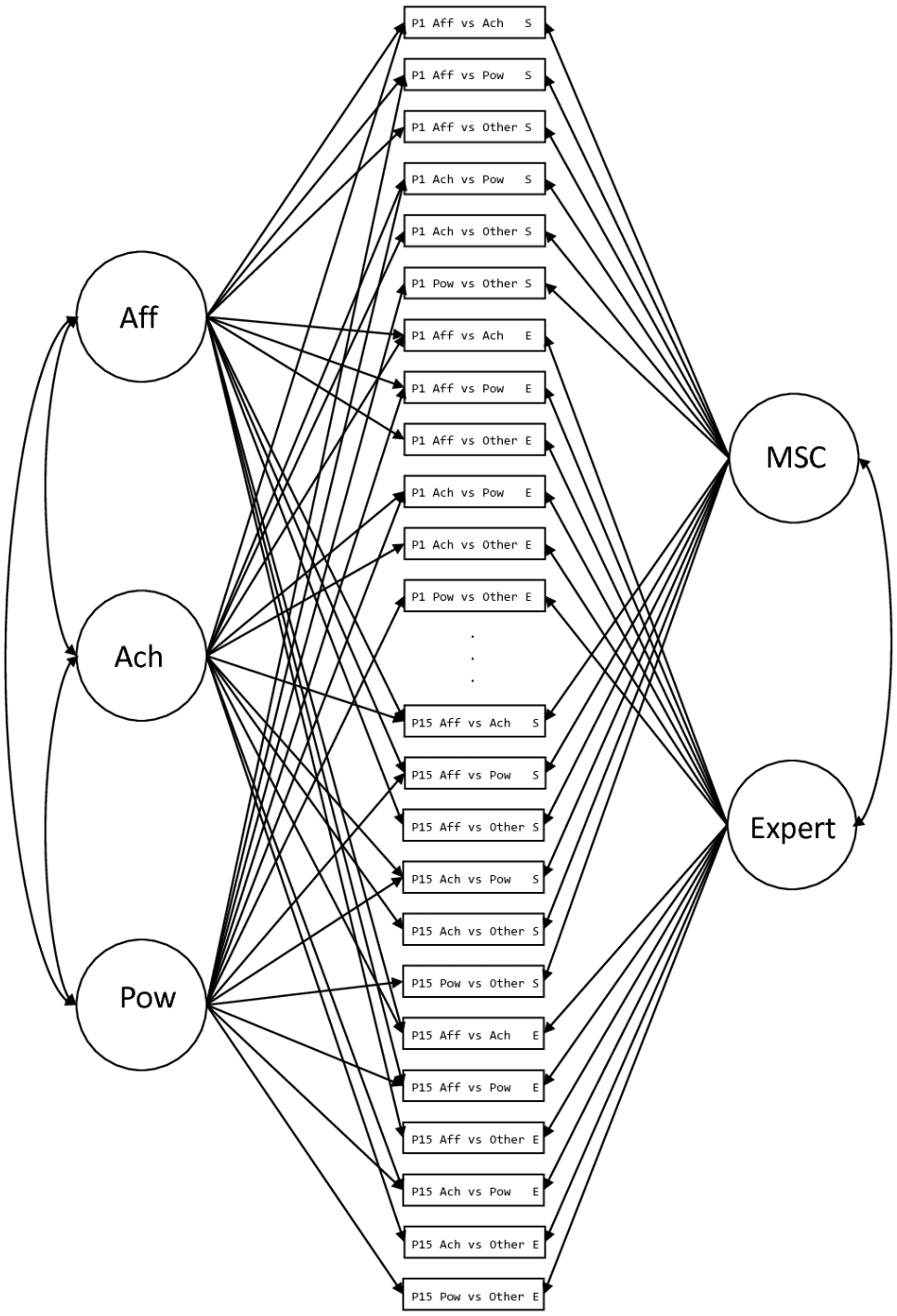

Figure 2. Correlated traits, correlated methods model. P = picture; Aff = affiliation; Ach = achievement; Pow = power; Other = other content; E = Expert coding; S = MSC.

Appendix

The dynamic Thurstonian IRT model for motive measures (Lang, 2014) can be modelled in a general linear mixed effect model. The model builds on the assumption that people compare latent motive utilities and that the strongest utility determines the response. Latent motive utilities (μ) are mean evaluation of a motive category in a specific picture. On level 1, motive competition is modelled as pairwise comparisons between latent utilities for each motive. The latent utilities are specified on level 2. The latent utility for each motive comprises of a fixed picture effect β_m and a picture covariate x_{ms} $(m = 1, \dots, M)$ with *M* being the total picture number, a fixed effect β and a covariate for the dynamic effect *DE* that sums up all previous responses for the motive, and a random person effect ν_s that captures the latent motive. The probability that person s prefers affiliation to achievement can be written as $\pi_{affachs} = \mu_{aff} - \mu_{ach} = (x_{affms}\beta_{affm} + DE_{aff}\beta_{aff} + \nu_{affs}) - (x_{achms}\beta_{achm} + DE_{ach}\beta_{ach} + \nu_{achs})$. A complete model description is provided by Lang (2014) and Runge et al. (2016).

We fitted all dynamic Thurstonian IRT models as general linear mixed effect models (GLMM) in the lme4 1.1-12 package (Bates, Maechler, & Bolker., 2011) in the R environment (R Core Team, 2015). In glmer, random effects are estimated with the maximum a posteriori method (De Boeck et al., 2011) and the glmer function uses the Laplace approximation. A detailed description of these procedures including an R code tutorial is provided by Runge et al. (2016).

CHAPTER 6

GENERAL DISCUSSION

This final chapter recapitulates and integrates the main findings of the four studies and documents our contribution to advancing the measurement of implicit motives. We furthermore outline the main theoretical implications of the psychometric and procedural advancements that were proposed and tested in this dissertation. Then, we discuss practical implications for the development and use of implicit motive tests in research and personality assessment in organizational contexts. At the end of this chapter, we will discuss future research directions and limitations of this dissertation.

Research overview

In **Chapter 2**, we focused on the psychometric framework of the dissertation and the reliability of implicit motive measures. We applied the recently developed DTM (Lang, 2014) to a dataset (*N*=633) of the OMT—a modern implicit motive measure with more pictures and briefer answers than more traditional picture-story based measures. The DTM models two important theory-based psychometric response mechanisms; motive competition and dynamic motive activation. We successfully modeled those response mechanisms in the OMT with the DTM. A first important contribution of this study is the replication of previous findings showing that motive strength declines during measurement (Lang, 2014). In addition, we extended these findings by testing additional conceptualizations of the dynamic effect. A second important contribution is the reliability analysis for the OMT based on simulations using the DTM. Cronbach's Alpha was relatively low for the OMT, consistent with findings for other implicit motives measures (Lang, 2014). We showed that the IRT reliability—accounting for motive competition and the dynamic effect—of the standard picture set in the OMT is within a range that is sufficient for research purposes (Ellis, 2013), but perhaps not high enough for individual assessment. The final contribution of the first study is an R code tutorial that allows other researchers to test their own data.

In **Chapter 3**, we used the DTM to study differential item functioning (DIF) in the OMT for respondents from three different cultural backgrounds (*N*=369). For a fair comparison between groups, it is necessary that the items do not have different parameters for the groups. Researchers have previously tested DIF to study cultural differences in implicit motives (Hofer, Chasiotis, Friedlmeier, Busch, & Campos, 2005). Differently to previous research, our IRT approach allowed for model-based separation of latent motives and item effects. An advantage of the approach using latent motives is that the response processes as described in Chapter 2 are included. The first contribution of this study is methodological. We extended the DTM to test pictures for DIF. To follow up on the R tutorial in Chapter 2, we have also provided the R code for researchers that are interested to compare implicit motives between groups accounting for possible DIF items. Our second contribution is a DIF analysis for a specific OMT picture set that was developed for research with respondents for three

different cultural backgrounds. We found that in three out of twelve pictures, one of the three response options showed DIF. The DIF effects were, however, relatively small. We conclude that implicit motive measures can be developed to study cultural differences with unbiased items. The third contribution of the study is the application of a partial invariance model to study latent mean differences specifying different item parameters for the three DIF items, and, therefore accounting for biased items. Because the DIF effects were small, conclusions from the group comparison were similar for the corrected (partial invariance) and the uncorrected model.

In **Chapter 4**, we focused on the predictive validity of implicit motives in an organizational context. More specifically, we wanted to predict counterproductive work behavior (CWB) with implicit motives. We investigated the effect of implicit motives on CWB in two studies (N = 263 and N = 121), controlling for trait personality using the BFI and the HEXACO. In the first study, respondents filled out self-reports of CWB and in the second study, supervisors rated CWB. We found in both studies that implicit motives show predictive validity over and above personality. More specifically, we found that the affiliation motive buffered CWB in both studies and the power motive in Study 2. In this Chapter, we contribute to a new generation on implicit motive research in OB. A strength of this Chapter is that we theorize and study main effects of implicit motives on organizational behavior—previous OB research typically focused on interaction effects (e.g. Kehr, 2004; Lang, Zettler, Ewen, & Hülsheger, 2012; Winter, 1987; Winter, John, Stewart, Klohnen, & Duncan, 1998).

In **Chapter 5**, we introduce a new procedure to measure implicit motives that does not rely on expert coders, the Motive Self-Categorization (MSC) test. The basic principle behind this new testing procedure is that respondents use items to self-categorize the stories they wrote. The items are picture specific, reflect frequently written stories, and are directly linked to an implicit motive. This procedure has several advantages. First, test scoring is more objective and does not rely on expert coders, who not always show perfect agreement. Second, the procedure does not need time intensive and trained expert coding procedures. Third, respondents have a detailed insight into ambiguous stories, whereas expert coders need to rely on what is written. Based on reliability simulations described in Chapter 2, we showed that the MSC reliability is comparable to expert codings, and in a series of latent multitrait-

multimethod models we found evidence for convergent validity between MSC and expert coding (N = 247). Furthermore, we showed that the MSC is different from personality traits and that the Big Three implicit motives show no substantial overlap with each other. The MSC is a substantial innovation and the first test that allows respondents to categorize implicit motives themselves.

Research implications

This dissertation has the goal to advance the measurement of implicit motives and thereby to promote their use in OB. In the following section, we discuss how our findings contribute to that goal and elaborate possible next research steps. First, we discuss the progress we made with this dissertation for the evaluation of implicit motives in different psychometric test criteria. Then we discuss various theoretical and practical implications of our findings.

Psychometric criteria

Reliability. In this dissertation, we applied the DTM to the OMT as well as the newly developed MSC. We found higher IRT reliabilities compared to internal consistencies using large datasets for both the OMT (Chapter 2 and 3) as well as the MSC (Chapter 5). The findings in this dissertation, thus, give strong evidence that the reliability of implicit motive tests is much higher than past criticism claimed. Note, however, that the reliabilities—although high enough for research purposes—are not as high as well-developed questionnaires measures. Nevertheless, this dissertation contributes to the discussion on the reliability of implicit motive measures by showing that the reliability can be in an acceptable range. We conclude that the DTM is an appropriate strategy to study reliability for implicit motives. The DTM substantially improves reliability estimation compared to traditional internal consistency approaches.

The next step is to identify possibilities to increase reliability in implicit motive measures. Based on our findings, it is difficult to give general advice on how to increase IRT reliability. Although the general rule that an increased test length leads to higher reliabilities may also apply to implicit motive tests, it comes with new challenges. Implicit motive tests are already relatively long, and further increasing test length may lead to test fatigue (Lundy, 1988; Smith, 1992). Because of the dynamic reduction in motive strength, it may also be possible that an increased test length will lead to relatively low motive strength for the pictures

at the end of a test. Adding additional pictures may therefore not be useful after a certain point (Lang, 2014). A possible way to study this question would be to let respondents work on 20 instead of 15 pictures. If the dynamic effects reduce motive strength substantially, one would expect that responses to the pictures presented at the end of the test mainly reflect picture effects. This would lead to a reduced variance between respondents in the last pictures. Additionally, to evaluate possible effects on reliability, one could fit the first 15 pictures in a basic DTM model and compare the simulated IRT reliability with the reliability of models that include 16, 17, 18, 19, and 20 pictures to see whether or not more pictures increase the IRT reliability.

Researchers have also suggested that ambiguous pictures measuring all three motives may influence reliability (Murstein, 1965; Pang, 2010; Smith, 1992). The basic idea behind this suggestion is that a picture with low variance where all respondents write stories for the same motive does not differentiate well between respondents. One could test this idea with picture sets differing in the amount of non-ambiguous pictures. For instance, one could develop picture sets that have zero, five, and ten non-ambiguous pictures and compare IRT reliabilities. To reduce sampling bias, it would be important to use large samples to receive meaningful comparisons.

Another possible way to increase reliability could be to focus on a single motive of interest. Researchers were interested in studying a single motive in the past and suggested picture sets based on the motive of interest (Smith, 1992). An advantage of this approach is, that the effective test length may be increased. When respondents would write on average 5 stories for affiliation, achievement, and power each in a standard OMT picture set with 15 pictures (note that this is unrealistic because there are also non-motive responses and the response frequency is typically not evenly distributed), test length may be considered as five pictures for a particular motive. One could alternatively develop a test with 15 pictures that each measures achievement vs power or achievement vs affiliation. Respondents would write more achievement stories and the test length for measuring achievement would, therefore, be longer compared to the standard set. Using the DTM, one could study whether or not this approach would increase reliability for a single motive. A possible caveat is that achievement

motive strength may decrease strongly due to the dynamic effect. A possible way to study whether this effect would limit reliability is described above.

Another possible way to increase reliability is to select items carefully and study additional item characteristics. The DTM is similar to a Rasch model in that it only models item difficulty (in case of the DTM it is item easiness). IRT models have also studied additional item parameters like item discrimination and guessing parameters. Guessing parameters are not of special interest for tests with forced choice like response processes without a single correct solution, like implicit motive measures. However, it may be of interest to study item discrimination parameters for implicit motive tests. Item discrimination parameters describe the slope of an item characteristic curve. When the slope is steep, the item discriminates well between respondents that have an ability (motive) which is higher and lower than the item difficulty for this item. Using both, item difficulty and item discrimination parameters, one can better select items for a specific test purpose. For instance, if one is interested in discriminating between participants high in implicit power motivation, one may focus on pictures that have higher item difficulties for power and also discriminate well for high power. Modeling item discrimination parameters is difficult in lme4 and almost doubles the number of fixed effects in the general linear mixed-effect model. Because the present version of the DTM already has a high number of fixed effects, it may also be difficult to fit and especially simulate such a model and a relatively high number of respondents may be needed. To develop a useful picture set for a specific research interest, one would need to develop a substantial picture pool as a basis to choose fitting items from. Although this procedure would be relatively complex it is certainly worthwhile to study possibilities to increase the reliability of implicit motive tests.

Differential item functioning. In Chapter 3, we extended the DTM to study differential item functioning (DIF) for implicit motive measures and investigated the usefulness of this approach in a dataset with participants from three different cultures. It is relatively straightforward to apply this procedure to other group characteristics such as gender or age. This makes our contribution especially valuable in a selection context where test fairness is a widespread concern (Robertson & Smith, 2001). In an applied context, it is important to not discriminate specific groups. When a test result is biased because of DIF, it

can lead to the rejection of a better candidate. Additionally, it is also of legal interest that selection instruments do not favor specific groups.

In this dissertation, we provided an R tutorial for the DTM and a description of how DIF analysis can be performed. Researchers and practitioners interested in developing implicit motive tests can use this procedure to remove DIF items in the test construction stage. Researchers have also been interested in comparing implicit motives in different groups like gender (Denzinger, Backes, Job, & Brandstätter, 2016; Drescher & Schultheiss, 2016), age (Denzinger et al., 2016), or different cultures (Hofer & Chasiotis, 2011). Although researchers addressed biased items in group comparisons of cultures (Hofer et al., 2005), it is not a standard procedure yet. Our study contributes to this type of research in two ways. First, researchers can develop implicit motive tests specifically for the comparison of specific groups based on DIF analysis. Second, it is also possible to correct for potential biases in group comparisons on a sample level with the partial invariance model we introduced for implicit motives.

Test economy. Although it is not a psychometric test criterion, it is also useful when a test is economical regarding test application and evaluation. Implicit motive measures have long relied on expert coders evaluating stories that respondents produce in response to pictures. Using expert coders is a relatively time-consuming procedure that needs extensive coder training, time to code stories (e.g. experienced coders evaluate four to six OMTs per hour), and story coding needs a lot of focus and therefore coders need frequent breaks. To ensure that motive scores are objective, researchers need to double code stories and calculate rater agreement. When raters have a low agreement, a third rater may be needed, and individual coding biases need to be identified and corrected. This inefficient procedure leads to a delay in the typical workflow of researchers. Normally, after data has been collected one can start structuring and analyzing the data. When implicit motives need to be coded first, this step can be delayed for a long time, depending on the sample size, resources, and rater agreement.

It is certainly necessary to invest time and resources into good measurement and research is generally not something that is easily done. Many researchers, however, may have avoided using implicit motives for practical reasons. The first three chapters of this

dissertation use expert coders to measure implicit motives. Considering the huge resource investment that is needed, it is not surprising that research on implicit motives is lagging behind personality research in many areas—measuring implicit motives is not convenient.

In Chapter 5, we developed and evaluated a new measurement procedure that is based on motive self-categorization (MSC) and does not rely on expert coders. We found evidence for convergent validity with expert codings. While testing time increases slightly with the MSC, time and resources to evaluate implicit motives reduces substantially. This innovation allows to study implicit motives more efficiently and facilitates access to measuring implicit motives for researchers new to the field. An economical measure that does not rely on extensive resources may facilitate future research to close the gap to the progress of trait personality research.

A next step in studying motive self-categorization could be to adapt the MSC procedure to other response formats for implicit motive measures. The MSC response format is similar to the OMT in that it uses brief responses to 15 pictures. Another popular response format that has emerged from the original TAT is the Picture Story Exercise (PSE; e.g. Pang & Schultheiss, 2005; Schultheiss & Pang, 2007). In the PSE, respondents typically respond to around six pictures (some use more, and some use less pictures). Responses are somewhat longer. A key difference in test scoring is that coders typically code based on sentences so that more than one coding per picture is possible. In the MSC procedure, respondents choose only one item per picture. For a self-categorization test based on the PSE response procedure, it may be useful to allow respondents to choose more than one item. Note, however, that this could lead to inflated correlations between motives when respondents frequently choose items from all motives. Researchers also typically correct for story length in the PSE, because a longer story may naturally contain more motive content which is not necessarily linked to a stronger motive, but to verbal abilities (Pang, 2010; Schultheiss & Pang, 2007; Smith, 1992). A natural caveat of correcting for story length may be that stories could also be long because a motive was strong instead of high verbal abilities. For motive self-categorization, it would also be possible to adjust the self-categorization scores for story length by dividing the score by the number of words in the story.

A different approach to reducing time and resources in measuring implicit motives is to develop software with a coding algorithm. Researchers have used a relatively simple approach based on the marker-word hypothesis (Schultheiss, 2013). The basic idea is that some motive categories are linked to the use of specific words in stories. This procedure correlated between .35 and .54 with expert codings using a weighted regression for the marker words (so no actual coding is given from the procedure). A natural issue for these types of approaches is the question whether models generalize to new samples, especially considering low sample sizes. Using models to predict behavior in future samples is generally difficult, especially when testing context changes. The approach may not work for different pictures and respondents that differ regarding language, verbal abilities, or social and educational backgrounds. For instance, researchers have shown that socio-economic background has a strong influence on used vocabulary (e.g. Hoff, 2003). It seems thus likely that the findings may greatly vary for different populations. Nevertheless, research in artificial intelligence is developing rapidly and future research may be successful in developing codings algorithms that reflect human coding behavior.

Validity

Test validity is a criterion that is less clear in how it should be evaluated. Researchers have focused on different types of validity. A validity measure that is relatively easy to study is predictive validity. The idea behind predictive validity is that a test measuring a certain construct should predict behavior that—based on theoretical arguments—should be predicted by the construct. In this dissertation, we contribute to studying the predictive validity of the OMT in Chapter 4. We found that the affiliation and the power motive buffer counterproductive work behavior. For studying the role of implicit motives in predicting organizational behavior, it is one of few studies focusing on main effects. In light of the recent replication crisis, researchers have found that interaction effects replicate less often than main effects (Open Science Collaboration, 2015). Within Chapter 4, we replicated two main hypotheses in our second study. It is also of note that predictor and criterion use different measurement procedures in these studies, and the second study used different rating sources for the criterion. Those precautions lower the chance that our findings were influenced by common method variance, a frequent critique on same source, same method (e.g. Likert type

questionnaires), and cross-sectional studies (Kozlowski, 2009), and strengthen our contribution in research on the predictive validity of implicit motive measures.

Another way to study the validity of a measure is to study whether a test converges with other tests that measure the same construct and whether a test score discriminates from test scores that measure a different construct. Researchers typically study convergent and discriminant validity using a multitrait-multimethod matrix (Campbell & Fiske, 1959). In this dissertation, we applied a latent modeling approach to study convergent validity of the MSC with expert coding and discriminant validity with personality trait ratings. A general problem in research on implicit motives is that different measures of implicit motives do not always converge consistently. Although researchers found some evidence for the convergent validity between the OMT and the PSE (Scheffer, 2001), and the motive grid (Schmalt, 1999) and implicit association tests (Slabbinck, De Houwer, & Van Kenhove, 2011), other researchers found that different implicit motive measures do not converge as expected (Schüler, Brandstätter, Wegner, & Baumann, 2015; Matthias Ziegler, Schmukle, Egloff, & Bühner, 2010). An explanation is, that they use different response formats that may measure different aspects of an implicit motive. The response procedure differs systematically between picture-story based procedures like the PSE or the OMT (Kuhl & Scheffer, 1999; Pang, 2010), reaction time based measures based on the implicit association test (Brunstein & Schmitt, 2004; Slabbinck et al., 2011; Slabbinck, De Houwer, & Van Kenhove, 2012; Matthias Ziegler et al., 2010), and explicit choice measures that also use pictures (Schmalt, 1999; Sokolowski, Schmalt, Langens, & Puca, 2000). Although some overlap was found for measures using different procedures (Slabbinck et al., 2011), research findings generally suggest that different procedures do not overlap. Future research may differentiate between those measures based on different implicit constructs, as suggested by Schultheiss (2007), who differentiated between implicit constructs based on memory systems.

For measures using the same response procedure as the OMT and the PSE, findings were also mixed. A possible explanation may be that when both tests are administered consecutively, the motive with more responses in the first test is lowered due to the dynamic effect and will, hence, be weaker in the second test. Consistent with this idea one study found no convergent validity between OMT and PSE when administered consecutively (Schüler et

al., 2015) and a different study found some evidence for convergent validity when OMT and PSE were measured at different time points (Scheffer, 2001). This explanation is supported by the evidence for dynamic motive reduction during measurement that we found in this dissertation. Future studies should carefully evaluate strategies to study convergent validity between, for instance, PSE and OMT. One strategy could be to alternate for each picture between the response procedures for the OMT and the PSE. Effects of dynamic motive decline would influence both measures simultaneously and the effect may be reduced. Alternatively, one could measure OMT and PSE at two different time points with a time lag and reverse the order of test administration for both measurements. If the motive strength declines as explained, one would expect that correlations between the same measures on each time point are low and correlations between the first measure on both measurement occasions are high. Additionally, it may be interesting to include the MSC in futures studies examining convergent validity between implicit measures. One would expect lower convergence compared to convergence between expert coding measures. Research has repeatedly shown that different rating sources show lower convergence (Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2004; McAbee & Connelly, 2016; Vazire & Carlson, 2010). A reason is, that different rating sources vary in the insight they have for the measured person and construct.

A strength of this dissertation is that we provide additional evidence for the DTM approach (Lang, 2014) as a psychometric theory of the response processes in implicit motive measures. Psychometricians have repeatedly argued that a latent variable model providing a theory of the response process in a test is a desirable type of evidence for its validity. Latent variable modeling requires a philosophical position that has been described as entity realism. In entity realism, a causal interpretation may be formulated for the relation between latent variables and their indicators (Borsboom, 2006; Borsboom, Mellenbergh, & van Heerden, 2003, 2004; Edwards & Bagozzi, 2000; Glymour, 2001; Pearl, 2000). This causal relation is commonly restricted to a between-subjects form for latent variable models and can under normal conditions not be generalized to processes at the individual level. Nevertheless, Borsboom et al. (2003) encouraged researchers to build integrative models that incorporate both individual differences and intraindividual variation in items into joint psychometric models. As a first step into this direction, the DTM includes individual dynamic processes and

demonstrates that accounting for these processes makes the extraction of meaningful latent between-person variables possible. Future research on implicit motives could seek to also include experimental picture manipulations to gain further insights. Efforts of this type would build models similar to the model that Embretson (1994) described for intelligence items. Borsboom, Mellenbergh, and van Heerden (2004, p. 1068) mention this model as an important example for a psychometric model that integrates both latent variables and intraindividual cognitive processes and thus bridges correlational and experimental research in the way they suggest. Research of this type would build on earlier experimental research in the motivation domain (e.g. McClelland, Atkinson, Clark, & Lowell, 1953; McClelland & Liberman, 1949). Testing interactions between motives and experimental conditions would also address Eysenck's and Revelle's call for experimental personality research (Eysenck & Eysenck, 1985; Revelle & Condon, 2015).

Theoretical Implications

A first theoretical contribution of this dissertation is modeling and studying the dynamic motive consumption using the DTM. In all analyses including the DTM, we found that a response for a specific motive would lower the response probability for this motive for all following pictures. This finding is consistent with previous research (Lang, 2014) and in line with the theoretical argument that motive strength reduces temporarily after enactment (Atkinson & Birch, 1970). In Chapter 2, we have also studied different conceptualizations of the dynamic motivational process. Atkinson and Birch suggested that the reduction in motive strength is only temporary and recovers after one picture, which would lead to a sawtooth pattern of responses. Again, consistent with the literature (Lang, 2014), we found that a model with a sustained dynamic effect fits better than the model with a quick recovery. We further contributed to the literature by studying whether motive strength recovers not after one but after two or three stories. Those models did not improve model fit so that a continuous decline in motive strength after each response describes the data best. To test whether the reduced response probability may also be an unspecific fatigue effect, we fitted a model in which the dynamic effect is not motive specific. However, this model also did not fit better as the motive specific consumption model.

A first implication is that motives do not recover during measurement. This implies that increasing the test length to increase reliability may have natural limits as described above. When motive strength is considerably reduced, responses may mainly reflect cue strength in pictures. While the DTM accounts for this in latent motive estimation, the information of additional pictures may not improve the estimation of latent motives.

Motive dynamics outside of measurement. Modern research on personality traits has focused on intraindividual variation and more detailed theories of when and how personality is relevant for the prediction of behavior (Baird, Le, & Lucas, 2006; Cervone, 2005). Researchers have also been studying stability and change of personality traits over the life-span (Specht et al., 2014), including effects on work-life (Wille & De Fruyt, 2014), and have called for process approaches to explain variability in personality and behavior.

The motive dynamics we showed in a measurement situation may build a framework for a process approach to explain within-person variability of motives and behavior that goes beyond just the measurement situation. Need for hunger, the first motive McClelland studied, is relatively intuitive to understand in a dynamic framework. When someone is hungry the need for hunger is relatively high, therefore the probability of the behavior *eating* increases. After eating, the need for hunger decreases. Although the need for hunger differs from the Big Three motives (this example did not include individual differences in hunger), the general within-person mechanism of dynamic motive decline that we observed in the measurement situation, may apply for the Big Three motives as well.

From a lifespan perspective, it may be interesting to study whether implicit motives reduce over a lifespan depending on continuous satisfaction in behavior. Research has found different results studying age differences in implicit motives. For instance, a study found evidence that implicit motives decrease with age (Denzinger et al., 2016), another study found they increase with age (Valero, Nikitin, & Freund, 2015), and a third one found no change (Thielgen, Krumm, Rauschenbach, & Hertel, 2015). A life-span perspective of motive reduction as a moderator may be useful to explain the contradictory findings. For instance, people that worked in a competitive work environment and reached several achievements over their work life may now have a lower achievement motivation compared to the beginning of their career. The basic idea would be that motives reduce not only temporarily but that a more

stable reduction occurs over a lifespan. To test this type of hypothesis one would need to expand present cross-sectional research with longitudinal designs. If this idea was true one would find that a decline in achievement motive could be explained by, for instance, career success. Conversely, someone who stayed home with their kids and has not engaged in a career may show higher achievement motivation in old age. A different and contrasting idea of motive change in old age is that work-related motives actually increase after retirement. When someone who is very achievement motivated working in an achievement environment retires, the implicit achievement motive may increase because it is not fulfilled any more. Recent progress in modeling longitudinal data allows to test the theory that an event can change an outcome by using the discontinuous growth model (Bliese & Lang, 2016; Lang & Bliese, 2009). Therefore, changes in implicit motives due to retirement may be studied in a longitudinal design using a discontinuous growth model. Research designs like these are, however, relatively difficult to implement because motives need to be measured at least at two to three time points before and after the event.

A motive perspective based on the findings in this dissertation may not only be useful to understand intraindividual changes of motive strength over one's life course but also—and possibly even more—to understand motive changes and possible effects on work behavior over weeks or on a daily level. The principles modeled in the DTM (motive competition and motive consumption) give a theoretical framework to study intraindividual dynamic changes in implicit motives and behavior. Motives are continually compared over time and the strongest motive determines behavior at any given moment. When a motive is temporarily reduced due to a previous action satisfying it, other motives may become stronger and win the motive competition to determine future behavior. For instance, after a long walk and talk with a close colleague, the affiliation motive may be lower, and an employee may be able to focus better on a complex work issue. The idea that leisure activities in between work days can have beneficial effects for recovery, and subsequently on both well-being (Sonnentag, 2001) and performance (Fritz & Sonnentag, 2005) has long been a focus in OB research (Sonnentag, 2003).

The dynamic effect provides a theoretical explanation based on motivational theory on what happens during a recovery period. Next to physical recovery, it is interesting to study (1)

whether the implicit achievement motive (this could be one of the other motives as well since all of them may be relevant in work contexts) declines during a work day when achievement related tasks are done and (2) whether motive strength restores during evening and weekend activities and (3) which factors facilitate motive recovery—for instance, whether activities that satisfy other motives play a role. To study these questions, employees could fill out an implicit motive measure before and after work in a diary design over several days. One would need to ask participants what work activities were performed and how they were related to the three motive themes. If for instance, the achievement motive reduces during work (1), the after-work measure should show lower achievement. If the achievement motive recovers overnight (2), measures on the following morning should be higher compared to the previous evening. Reports on evening activities could indicate whether employees did things that address other motives, like spending time with their families or friends (i.e. affiliation). If it were true that activities that are tied to other motives facilitate a quicker recovery (3), one would find that the motive relevance of an evening activity moderates motive recovery overnight. It may also be interesting to study motive recovery not on a daily level, but between weeks or over holidays. The idea is that longer recovery periods would allow for a better recovery. A caveat for studying these questions is, that implicit motives are defined as stable constructs. If one wants to measure temporal changes of motive strength, one would need to develop a state measure for motives. A possible alternative could be to model motive decline and recovery in an explanatory IRT model (De Boeck & Wilson, 2004) based on the DTM. A formalized theory on changes of behavior is explained in *dynamics of action* (Atkinson & Birch, 1970) and a simpler version is provided by Kuhl and Blankenship (1979). Dynamics of action describes a set of rules that go beyond motive competition and dynamic motive reduction and includes other mechanisms. Another very important driver of motive guided behaviors are situational factors that we describe in a future research section below.

Practical implications

Practical applications of implicit motive measures include, among others, clinical psychology and OB. In clinical psychology, implicit motives are rarely used in daily practice (Weinberger, Cotler, & Fishmen, 2010). Although there is a general interest in implicit motives in a clinical context (Baumann, Kaschel, & Kuhl, 2005; McAdams, 1997; Mihura,

Meyer, Dumitrascu, & Bombel, 2013; Westen, Lohr, Silk, Gold, & Kerber, 1990), researchers and practitioners may be deterred by the complex measurement procedure and previous reports of low reliability. This dissertation provides resources for both concerns. First, we showed that reliability is actually not as low as previous criticism suggested. Second, the MSC removes obstacles in testing. Moreover, it may be of special interest in a clinical context to study whether self-categorizations using the MSC differ systematically from expert codings. For instance, a person may repeatedly write stories about being rejected (avoidance motivation for affiliation) but categorizes those stories into healthy affiliative motives. This would indicate that it may be worthwhile to look into the person's perception of social relationships.

Next to research in OB, practitioners are frequently interested in the application of psychological measurement in personnel selection. The basic goal of psychological tests in personnel selection is to differentiate between candidates and predict which candidate suits best to a position. For implicit motives, previous research has stated that a hindrance for testing implicit motives in personnel selection is the use of expert coders (Lang et al., 2012). With the MSC, we address this concern and suggest a procedure that does not rely on expert coders. Another concern mentioned in the literature is that researchers argued that implicit motive measures are susceptible to situational factors of the presented pictures (Lang et al., 2012; Schultheiss & Brunstein, 2010; Schultheiss, Liening, & Schad, 2008). Following this argumentation, pictures showing scenes that are not relevant for a work context would show a low predictive validity for behavior in a work context. In the development of the MSC, we have considered these arguments and designed a new set of pictures depicting scenes in a work context. In this dissertation, we addressed the two main concerns mentioned in the literature. As such we hope to facilitate the applicability of implicit motive measures in a selection context.

Another intuitive concern for implicit motives as a selection tool is that a test needs to be reliable. In this dissertation, we showed that earlier concerns about the reliability of implicit motives were not justified. Although the reliability of implicit motive measures is not as high as the one of questionnaire-based personality measures, it is within an acceptable range. We have also discussed possible ways to increase reliability earlier in this discussion. Another possibility to increase the reliability of implicit motive tests in a selection context is to test

implicit motives at two time points on different days using different pictures. This strategy would increase the effective test length and avoid possible problems regarding test fatigue and motive reduction during the test. Although reliability is a very important criterion to evaluate the usefulness of a test in selection contexts, researchers have also argued that reliability does not necessarily need to be extremely high. Some procedures in personnel selection show relatively low reliability and are still used. For instance, situational judgment tests vary in internal consistency between .43 and .94, but researchers showed they are a valuable contribution in personnel selection (Lievens, Peeters, & Schollaert, 2008). The central criterion for the usefulness for a measure in personnel selection is criterion-related validity. A test is useful in personnel selection when it predicts work-related behavior. In this dissertation, we could show that implicit motives predict counterproductive work behavior. We have also presented a short review of past findings where implicit motives predict job-related behavior. Compared to personality trait research, there is, however, relatively little knowledge on the link between implicit motives and work-related behavior. This dissertation improves the measurement of implicit motives in several ways and we hope to thereby facilitate future research to increase the practical knowledge of criterion-related validity for implicit motive tests.

Limitations and future research directions

Retest reliability

One aim of this dissertation was to study the reliability of implicit motive measures. The theoretical basis revolved around modeling response processes to show that reliabilities can be much higher compared to internal consistency. While we were successful with this aim, other ways to study reliability were left out. Most notably we did not study the test-retest reliability of the OMT or the MSC. Test-retest reliability is a stability coefficient that is calculated using the correlation between the same test measured at two different time points. A high test-retest correlation gives evidence that the measure is reliable and stable. Researchers have studied test-retest reliability of implicit motives before. Findings are overall mixed. The main criticism of implicit motive having low reliability included low test-retest reliability (Entwisle, 1972; Fineman, 1977; Lilienfeld, Wood, & Garb, 2000). Advocates of implicit motive measures have pointed out that these criticisms omitted several studies that had

acceptable levels of test-retest reliability (Smith, 1992). Another potential issue why lower test-retest reliabilities were found in some studies is that respondents may intentionally write stories that differ from their previous stories because respondents want to show their creativity. When researchers give the instruction that respondents can write the same story as the previous time because it is not a creativity test, test-retest reliability is typically between .50 and .60 (Koestner & McClelland, 1992; Scheffer, 2001; Schultheiss et al., 2008). However, this instruction may lead to inflated correlations when participants write the same story on purpose. For the MSC, future research should study whether test-retest reliability is within the same range as expert coded implicit motives.

Individual differences in the dynamic effect

Another interesting question both for measurement as well as applied contexts is whether individual differences exist in the reduction of motive strengths, as well as recovery of motives. Research suggested in the past that such effects may exist (Revelle, 1986). Lang (2014) tested a model that included individual differences in the reduction of motive strength and found that it improved model fit. However, latent motives correlated highly with the DTM model that did not account for those effects. He concluded that for latent motive estimation it may, thus, not bring additional advantages. An interesting theoretical question is, whether individual differences in motive reduction correlate with motive strength. This would reflect the intuitive idea that a strong motive is more difficult to satisfy. However, it also seems plausible that a strong motive reduces faster, as there is more to reduce overall. Considering motive strength recovery, it is interesting whether recovery of motive strength also varies between persons, and if so, if differences relate to individual differences in motive reduction. For applied research, it may be interesting whether individual differences in dynamic effects—both reduction and recovery—predict organizational behavior. For instance, a person in whom the achievement motive recovers quicker may also show a better task performance compared to a person with a slower recovery. This idea relates to research showing that physical and mental recovery relates to higher performance (Binnewies, Sonnentag, & Mojza, 2009). One could study whether motive specific recovery improves performance with a diary design controlling for physical recovery. However, it may also be interesting whether or not motive recovery is related to physical recovery.

Future research with the Motive Self Categorization test

In this dissertation, we developed and performed initial psychometric research on motive self-categorization. Our findings suggest that the MSC is a promising approach to measure implicit motives without expert coding. Although we found initial evidence for reliability and construct validity, there are several questions for future research. The answer processes for the self-categorization part are not fully understood yet. Although we could find the dynamic effect also in the MSC, there may be additional processes involved in the self-categorization. Correlations between self and expert coding were in the typical range for self-other ratings of personality. That means that there was a substantial overlap, but also that for some stories expert coders and respondents differed in motive categorization. The disagreement could be attributed to different reasons. Generally, there may be two types of systematic reasons behind disagreement.

The first type describes a case when respondent and coder disagree on what motive actually lies behind the story. This may happen for different reasons. For instance, sometimes stories can be relatively short (only a few words). For short stories, it can be difficult for expert coders to identify the underlying motive. In this case, it may still be possible for respondents to categorize their story. Another possible source of disagreement could be that multiple motives are present in a picture. For the OMT, only one motive can be coded. In cases of multiple strong motives in a story, coders typically follow the "primacy rule" for coding (Kuhl & Scheffer, 2001) stating that one should code the first motive that comes up in a story. For self-categorization, respondents are instructed to choose the option that fits best to their story when multiple items apply. It is possible, that respondents think that a motive that appears later in a story is more important compared to the first, which would then lead to a different coding compared to the expert.

The second general type of disagreement is intentional changes in self-categorization. Respondents may change their story in hindsight when they see other options. A possible motivation for respondents would be to present themselves in a better light and to manage the impression they make on others (Bolino, Varela, Bande, & Turnley, 2006). When stories with dominant actions or failures are written, respondents may want to reframe their story in a more positive light. When respondents are tested for a specific purpose like a job interview, they

may also change their answers to a motive that they believe is more important for a good evaluation (e.g. achievement instead of affiliation).

In addition to systematic disagreement, there may also be unsystematic disagreement between respondent and coder. This could be the case when they believe that a different motivational theme simply fits better to the story when a coder is influenced by other stories that were previously coded, a respondent accidentally crosses a different item, or a coder accidentally writes down a different coding as intended. Unsystematic disagreement is generally difficult to detect. A possibility to study systematic disagreement would be to let expert coders immediately code the stories after they have been filled out. Then, one could look at specific stories where coder and respondent disagree and allocate the disagreement to one of the above mention sources, based on "live" feedback from both the coder and the respondent.

In a selection context, it would be interesting to study whether respondents try to fake the test to improve the impression they make in the selection process. One can investigate this question using a classic faking study in which one group is instructed to fake the test for a specific criterion ("be as achievement motivated as possible") and a control group with no instruction. When the test is easy to fake, one would see that the experimental group has a significantly higher achievement score compared to the control group. However, this would only study whether it is possible to fake the test. Because it is relatively easy to change the answer option to what one believes is favorable, one would expect that the MSC is very fakeable. Whether respondents actually fake the test needs to be studied in a high-stakes testing situation where respondents apply for a real job. It is less clear whether respondents would fake the test in a real selection procedure because strong discrepancies between the written story and the self-categorization would be visible. Respondents who are aware that differences could be detected may not try to fake the test to not make the impression to be dishonest.

Another interesting future research direction is to study the criterion-related validity of the MSC. Because we found evidence for convergent validity, one could argue that expert coded and self-categorized implicit motives predict similar behavior. However, the overlap was not perfect, so it is also possible that expert coded and self-categorized implicit motives

predict different behavior. If different behavior is predicted, one should investigate the source of the difference—possible reasons for systematic disagreement are discussed above. For instance, in a selection context, it seems plausible that social desirability and faking could be responsible. When respondents fake the test, the relationship between implicit motives and criterion-related validity may be biased. However, research has also shown that social desirability is not much of a concern for personality tests in selection (Ones & Viswesvaran, 1998). Future studies should evaluate whether or not these findings apply to implicit motive self-categorizations as well.

Construct clarification

The terms *implicit* and *operant*. Another topic for future research—that is not directly linked to the findings in this dissertation, but highly relevant—is the clarification of terms in research on implicit motives. Some of the concepts and terms that are frequently used in research on implicit motives are ambiguous or used in other research contexts and therefore lead to confusion. This section is not meant to resolve this issue but rather to point it out.

The first term we discuss is *implicit*. As described in the introduction, implicit motives have not always been termed implicit. Early research studied them as *needs*. The term need is nowadays associated with self-determination theory, a motivational theory that focusses on basic motivational needs but does not focus on individual differences in needs (Ryan & Deci, 2000). Although researchers have pointed out several conceptual similarities between implicit motives and needs as in self-determination theory (Schüler, Baumann, Chasiotis, Bender, & Baum, 2019), it seems useful to not use similar words for the different concepts.

McClelland and colleagues (1989) suggested distinguishing between motive constructs measured with a questionnaire and TAT-based measures and introduced the terms implicit motives and explicit motives. The terminology has intuitive appeal, because it clearly described the differences between the constructs, and quickly became the standard for motive research. For McClelland, the main feature of an implicit motive was that "the person is not explicitly describing him or herself as having the motive" (McClelland et al., 1989, p. 691) and implicit motives are described as not accessible through conscious awareness.

Schultheiss and Brunstein (2010) describe that the term implicit was chosen borrowing from a recent paper on implicit memory (Schacter, 1987). The work by Schacter and previous

work on implicit memory did not only influence the naming of implicit motives but also of what is known as implicit social cognition (e.g. Greenwald & Banaji, 1995). Implicit social cognitions are typically measured using reaction-time based measures such as the Implicit Association Test (IAT; De Houwer, 2003; Greenwald, McGhee, & Schwartz, 1998). Research studying whether reaction time based and picture story based measures are measuring the same construct found mixed results (Brunstein & Schmitt, 2004; Slabbinck et al., 2011, 2012; Matthias Ziegler et al., 2010). Schultheiss (2007) suggested that implicit motives and individual differences measured with the IAT refer to different memory systems and measure different constructs. The dissimilarity between both approaches may be large enough to think about using a different term for them.

Another source of confusion is the term *operant* that McClelland used to describe implicit motive measures and the type of behavior that they predict, in distinction to *respondent* that describes explicit motive measures (McClelland, 1980, 1985; McClelland, Koestner, & Weinberger, 1989). Operant measures—as used by McClelland—refers to spontaneous behavioral trends and measurement procedures in which people respond freely and spontaneously. In respondent measures, people react to a fixed presented stimulus like a test item with a Likert scale that leaves little freedom for responses (Emmons & McAdams, 1991). The terms operant and respondent were adopted by the work of Skinner (1938). McClelland stated that "Skinner used these terms to distinguish between occasions in which the stimulus for a response could be identified (respondent behavior) or could not be identified (operant behavior)." (McClelland et al., 1989, p. 692). Today, the research lines of Skinner and Murrays are seen as two very different research traditions. Skinner's work is mainly associated with behaviorism and (operant) conditioning, whereas Murray's needs focus on individual differences in motivational needs. Because the term "operant" is closely tied to Skinner's concept of operant conditioning, it seems inappropriate to use it in the context of implicit motives. Interestingly, however, both research traditions may not be as distinct as commonly supposed. Paul E. Meehl, a student of Skinner, has argued that Skinner's approach of seeing drive as a state-variable agrees closely with Murray's concept of needs (Meehl, 1992). Nevertheless, to avoid confusion, it would be useful to discontinue the use of the terms operant and respondent in the literature of implicit motives. For the measurement side, it

seems plausible to stay close to the response process and describe implicit and explicit motive measures as free or spontaneous and closed or fixed responses. For the behavior part, it may be useful to define more clearly what type of behaviors are predicted by implicit and explicit motives, respectively. The term operant behavior is somewhat fuzzy and leaves much room for interpretation for what behaviors are meant.

What is implicit in implicit motives? Research on implicit social cognition has long discussed how to best define what is "implicit" in implicit measures (De Houwer, 2006). It may be useful to build on this research and clarify in what way implicit motive measures are implicit. As summarized by De Houwer (2006), implicit can refer to the fact (1) that respondents are not aware of what construct is measured (Brunel, Tietje, & Greenwald, 2004), (2) that respondents have no insight or conscious access to the construct (Asendorpf, Banse, & Mücke, 2002), or (3) that respondents cannot control the measurement outcome (Fazio & Olson, 2003). Later research has more rigorously scrutinized properties of implicit measures in the tradition of the IAT (De Houwer, 2006; De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009). An important note of this newer research is that a measure is not implicit *per se,* but that it is more useful to define more specifically in what way a measure is implicit. Implicit motives have been described as implicit, in the sense that (1) respondents are not aware what is measured and that (3) they have no conscious access to their implicit motives (McClelland et al., 1989). Contrary to some reaction time-based measures, respondents could control the measurement outcome (2) when they are instructed what is measured (writing a story for a specific motive) so that this property does not apply to implicit motives.

To conclude that a measure is really implicit, one has to empirically test whether it has the properties it claims to have (De Houwer, 2006). To study whether respondents are aware of what is measured in implicit motives tests, one could ask respondents after test completion what they believe was measured. We would expect that in most cases respondents are unaware of the measured construct. The question of whether or not respondents have conscious access to implicit motives is more difficult to study. Although researchers have argued that implicit motives are not accessible through introspection (McClelland et al., 1989; Schultheiss & Brunstein, 2010) and there is substantial evidence that implicit and explicit motive measures have low convergence (Köllner & Schultheiss, 2014; Spangler, 1992), other researchers have

argued that after extensive education what implicit motives are, people have some insight into their implicit motives (Sherwood, 1966). In this dissertation, we developed a motive self-categorization that is not based on actual insight, but respondents still use some insights on their stories to categorize them into motivational themes. Future research should more rigorously study whether people can have insights into their implicit motives. A useful starting point may be the findings that people need some education on what implicit motives are before they can report them. It may not be sufficient to conclude that respondents have no conscious access to their implicit motives based on research showing questionnaire based self-ratings do not converge with implicit motives. Especially because researchers have argued that for some explicit motive measures there is little overlap with how implicit motives are defined (Schultheiss, Yankova, Dirlikov, & Schad, 2009; Thrash, Cassidy, Maruskin, & Elliot, 2010).

Integrating the motive perspective with other concepts

Motives and traits. Researchers have suggested that both personality traditions—motives and personality traits—that we described in the introduction of this dissertation interact to predict behavior. Both traditions have "led separate lives" from their beginning at Harvard University in the 1930ies (Winter et al., 1998). But researchers started to integrate both concepts to develop a more fine-grained personality system at the end of the last century. The foundation for the interactional theory is the *channeling hypothesis* introduced by Winter and colleagues (1998). The channeling hypothesis says that personality traits channel implicit motives. The core idea is that implicit motives influence the choice of behavior (typically with the goal to satisfy a motivational need) and traits determine how and in what way the behavior is performed. For instance, a person that is high in the affiliation motive wants warm and positive relationships with others. Extraversion will influence how the person will pursue that goal. Extraverted people satisfy their affiliation motive at parties and meeting many different people, while introverted people satisfy their affiliation motive with a few close friends. To study this idea, Winter and colleagues (1998) found that affiliation motivation interacts with extraversion to predict volunteer work. In two samples, affiliation motivated people engaged more in volunteer work when extraversion was high. In I/O psychology, researchers suggested that some motive-trait combinations are useful, and others may lead to frustration (Barrick, Mount, & Li, 2013; Lang et al., 2012). For instance, Lang and colleagues (2012) found that

task performance is highest when implicit and trait-based achievement motivation is high, while a misfit between implicit and trait-based achievement motivation leads to need frustration and lower performance.

In chapter 4, we focused on testing main effect hypotheses for implicit motives, so we do not directly contribute to this line of research. However, before theorizing that traits and motives interact as a different personality construct to predict behavior, it is useful in a first step to search for evidence that they differ. Both constructs are measured with self-rating procedures which may lead to common method variance and higher motive trait correlations so that it is important to study discriminant validity. We found in Chapter 4 and 5 evidence for discriminant validity between traits and motives that were measured with expert ratings and self-categorized implicit motives. We hope that the evidence we found of discriminant validity for self-categorized implicit motives and traits facilitates future research on channeling effects. Our findings, thus, build a ground for future research on channeling with a less resource intensive measure.

Motives and situations. A specific interest of personality research has been the interplay between personality and situational factors. For instance, researchers have suggested in the trait activation theory that not every trait may be relevant in all situations (Tett & Guterman, 2000) and able to be applied to OB (Burnett & Tett, 2003; Lievens, Chasteen, Day, & Christiansen, 2006). For implicit motives, we found support for the importance of a situational factor in the assessment context. In the DTM, not only the motives themselves are compared, but a *latent utility*, which is a combination of motive strength and situational incentives. This idea is similar to trait activation in that not all motives are relevant in all situations, but that they can be activated similar to traits. In a testing situation, some pictures lead to responses that do not include a motive. A picture that shows two people hugging each other would typically not lead to achievement stories. It seems plausible that not all motives are relevant in all situations. Researchers studying implicit motives have also argued for the importance of the situational context in understanding how implicit motives work. For instance, McClelland and colleagues (1989) argued that it is important for implicit motives that a motive relevant incentive is present in a situation for the motive to predict behavior. For the achievement motive, a challenging task could be such an incentive. Behavior in a situation

with no challenging task is likely not influenced by the achievement motive. Other researchers have argued that specific situational characteristics may also influence whether a motive is active or not (Schultheiss et al., 2008). They argue that people learn for specific types of situations that a motive relevant incentive and subsequent behavior leads to a positive emotional reaction. For instance, people may have learned to express the achievement motive in a sports context but not in an academic context.

The importance of situational variables is certainly not limited to the measurement context. Similar to trait activation theory, situational characteristics may also play an important role in "motive activation". For instance, during a day, a person's high affiliation motivation may not influence behavior when a complex technical problem needs to be solved, but it may influence the decision later that day on whether or not to join colleagues for a drink. To study the role of situational characteristics, situations need to be measured. Recent research has made progress in measuring situations using ratings of different situational dimensions (Parrigon, Woo, Tay, & Wang, 2017; Rauthmann et al., 2014; Ziegler, Horstmann, & Ziegler, in press). Some of the situational dimensions map relatively well to the content of motives so that they may be a useful tool to study the importance of situational characteristics for implicit motives. For the DIAMONDS (Rauthmann et al., 2014), the dimensions mating and sociality may be connected to the affiliation motive and intellect may be related to the achievement motive. For the CAPTION (Parrigon et al., 2017), complexity may be related to the achievement motive. However, those measures do not specifically ask for motive relevant incentives. Asking whether a situation is complex or intellectual is not the same as asking whether a challenging task is present. A recent study (Brandstätter, Job, & Schulze, 2016) has asked more specifically whether motive specific incentives are present at work. Future research could build on this study and ask for the presence of motive specific incentives in a situation.

A possible extension would be to ask for motive incentives in a specific situation and not as a general job characteristic. Such a measure could allow studying motive activation on a situational basis using experience sampling methods. If it is true that motives are activated in a situation, one could observe on a within-person level that achievement-motivated people show a better job performance in a situation where there are achievement incentives, compared to a

situation with fewer incentives. On a between-person level, one should see performance differences when the situation has achievement incentives, but in a situation with no achievement incentives, one would also not expect to find performance differences between employees high and low in the achievement motive.

Conclusion

In this dissertation, we present different ways to improve the measurement of implicit motives in an organizational context. Based on theory, we used psychometric models to study and improve the properties of implicit motive measures. We learned that implicit motives can be measured reliably, found evidence that different theoretically motivated response mechanisms influence the response procedure, and showed the usefulness in an organizational context. Furthermore, we innovated the measurement procedure and introduced a test in which respondents categorize implicit motives themselves. Together with the provided code, we give useful tools to develop, evaluate, and study the measurement of implicit motives.

References

Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: the case of shy behavior. *Journal of Personality and Social Psychology*, *83*, 380–93. http://dx.doi.org/10.1037/0022-3514.83.2.380

Atkinson, J. W., & Birch, D. (1970). *Dynamics of Action*. New York: Wiley.

Baird, B. M., Le, K., & Lucas, R. E. (2006). On the nature of intraindividual personality variability: reliability, validity, and associations with well-being. *Journal of Personality and Social Psychology*, *90*, 512–27. https://doi.org/10.1037/0022-3514.90.3.512

Barrick, M. R., Mount, M. K., & Li, N. (2013). The theory of purposeful work behavior: The role of personality, job characteristics, and experienced meaningfulness. *Academy of Management Review*, *38*, 1–51. https://doi.org/10.5465/amr.2010.0479

Baumann, N., Kaschel, R., & Kuhl, J. (2005). Striving for unwanted goals: Stress-dependent discrepancies between explicit and implicit achievement motives reduce subjective well-being and increase psychosomatic symptoms. *Journal of Personality and Social Psychology*, *89*, 781–799. https://doi.org/10.1037/0022-3514.89.5.781

Binnewies, C., Sonnentag, S., & Mojza, E. J. (2009). Daily performance at work: feeling recovered in the morning as a predictor of day-level job performance. *Journal of Organizational Behavior*, *30*(1), 67–93. https://doi.org/10.1002/job.541

Bliese, P. D., & Lang, J. W. B. (2016). Understanding relative and absolute change in discontinuous growth models: Coding alternatives and implications for hypothesis testing. *Organizational Research Methods*, *19*, 562–592. https://doi.org/10.1177/1094428116633502

Bolino, M. C., Varela, J. A., Bande, B., & Turnley, W. H. (2006). The impact of impression-management tactics on supervisor ratings of organizational citizenship behavior. *Journal of Organizational Behavior*, *27*, 281–297. https://doi.org/10.1002/job.379

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology*, *86*, 599–614. https://doi.org/10.1037/0022-3514.86.4.599

Borsboom, D. (2006). The attack of the psychometricans. *Psychometrika*, *71*, 425–440.

https://doi.org/10.1007/s11336-006-1447-6

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*, 203–219. https://doi.org/10.1037/0033-295X.110.2.203

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The Concept of Validity. *Psychological Review*, *111*, 1061–1071. https://doi.org/10.1037/0033-295X.111.4.1061

Brandstätter, V., Job, V., & Schulze, B. (2016). Motivational Incongruence and Well-Being at the Workplace: Person-Job Fit, Job Burnout, and Physical Symptoms. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01153

Brunel, F. F., Tietje, B. C., & Greenwald, A. G. (2004). Is the implicit association test a valid and valuable measure of implicit consumer social cognition? *Journal of Consumer Psychology*, *14*, 385–404. https://doi.org/10.1207/s15327663jcp1404_8

Brunstein, J. C., & Schmitt, C. H. (2004). Assessing individual differences in achievement motivation with the implicit association test. *Journal of Research in Personality*, *38*, 536–555. https://doi.org/10.1016/j.jrp.2004.01.003

Burnett, D. D., & Tett, R. P. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, *88*, 500–517. https://doi.org/10.1037/0021-9010.88.3.500

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105. https://doi.org/10.1037/h0046016

Cervone, D. (2005). Personality architecture: Within-person structures and processes. *Annual Review of Psychology*, *56*, 423–452. https://doi.org/10.1146/annurev.psych.56.091103.070133

De Boeck, P., & Wilson, M. (2004). *Explanatory item response models : a generalized linear and nonlinear approach*. Springer New York. https://doi.org/10.1007/978-1-4757-3990-9

De Houwer, J. (2003). The Extrinsic Affective Simon Task. *Experimental Psychology*, *50*, 77–85. https://doi.org/10.1026//1618-3169.50.2.77

De Houwer, J. (2006). What are implicit measures and why are we using them. In R. W. Wiers & A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction* (pp. 11–28).

Thousand Oaks, CA: Sage Publishers.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368. https://doi.org/10.1037/a0014211

Denzinger, F., Backes, S., Job, V., & Brandstätter, V. (2016). Age and Gender Differences in Implicit Motives. *Journal of Research in Personality*. https://doi.org/10.1016/j.jrp.2016.09.003

Drescher, A., & Schultheiss, O. C. (2016). Meta-analytic evidence for higher implicit affiliation and intimacy motivation scores in women, compared to men. *Journal of Research in Personality*, *64*, 1–10. https://doi.org/http://dx.doi.org/10.1016/j.jrp.2016.06.019

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, *5*, 155–174. https://doi.org/10.1037/1082-989X.5.2.155

Ellis, J. L. (2013). A standard for test reliability in group research. *Behavior Research Methods*, *45*, 16–24. https://doi.org/10.3758/s13428-012-0223-z

Embretson, S. E. (1994). Applications of cognitive design systems to test development. In C. R. Reynolds (Ed.), *Cognitive assessment: A multi- disciplinary perspective.* (pp. 107–135). New York: Plenum Press.

Emmons, R. A., & McAdams, D. P. (1991). Personal strivings and motive dispositions: Exploring the links. *Personality and Social Psychology Bulletin*, *17*, 648–654. https://doi.org/10.1177/0146167291176007

Entwisle, D. R. (1972). To Dispel Fantasies About Fantasy-Based Measures of Achievement Motivation. *Psychological Bulletin*, *77*, 377–391. https://doi.org/10.1037/h0020021

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and Individual Differences: A Natural Science Approach*. New York: Plenum Press.

Fazio, R. H., & Olson, M. A. (2003). Implicit Measures in Social Cognition Research: Their Meaning and Use. *Annual Review of Psychology*, *54*, 297–327. https://doi.org/10.1146/annurev.psych.54.101601.145225

Fineman, S. (1977). The achievement motive construct and its measurement: Where are we

now? *British Journal of Psychology*, *68*, 1–22. https://doi.org/10.1111/j.2044-8295.1977.tb01554.x

Fritz, C., & Sonnentag, S. (2005). Recovery, health, and job performance: Effects of weekend experiences. *Journal of Occupational Health Psychology*, *10*, 187–199. https://doi.org/10.1037/1076-8998.10.3.187

Glymour, C. N. (2001). *The mind's arrows : Bayes nets and graphical causal models in psychology*. MIT Press.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7878162

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–80. https://doi.org/10.1037/0022-3514.74.6.1464

Hofer, J., & Chasiotis, A. (2011). Implicit motives across cultures. *Online Readings in Psychology and Culture*, *4*, 5.

Hofer, J., Chasiotis, A., Friedlmeier, W., Busch, H., & Campos, D. (2005). The measurement of implicit motives in three cultures power and affiliation in Cameroon, Costa Rica, and Germany. *Journal of Cross-Cultural Psychology*, *36*, 689–716. https://doi.org/10.1177/0022022105280510

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*, 1368–1378. https://doi.org/10.1111/1467-8624.00612

Kehr, H. M. (2004). Implicit/explicit motive discrepancies and volitional depletion among managers. *Personality and Social Psychology Bulletin*, *30*, 315–327. https://doi.org/10.1177/0146167203256967

Koestner, R., & McClelland, D. C. (1992). The affiliation motive. In C. P. Smith (Ed.), *Motivation and Personality: Handbook of Thematic Content Analysis* (pp. 205–210). Cambridge, MA: Cambridge University Press.

Köllner, M. G., & Schultheiss, O. C. (2014). Meta-analytic evidence of low convergence between implicit and explicit measures of the needs for achievement, affiliation, and

power. *Frontiers in Psychology*, *5*, 826. https://doi.org/10.3389/fpsyg.2014.00826

Kozlowski, S. W. J. (2009). Editorial. *Journal of Applied Psychology*, *94*, 1–4. https://doi.org/10.1037/a0014990

Kuhl, J., & Blankenship, V. (1979). *The Dynamic Theory of Achievement Motivation: From Episodic to Dynamic Thinking*. *Psychological Review* (Vol. 86).

Kuhl, J., & Scheffer, D. (1999). Auswertungsmanual für den Operanten Motiv-Test (OMT).

Kuhl, J., & Scheffer, D. (2001). Der operante multi-motiv-test (OMT): Manual [The operant multi-motive-test (OMT): Manual]. Osnabrück, Germany: University of Osnabrück.

Lang, J. W. B. (2014). A dynamic Thurstonian item response theory of motive expression in the picture story exercise: Solving the internal consistency paradox of the PSE. *Psychological Review*, *121*, 481–500. https://doi.org/10.1037/a0037011

Lang, J. W. B., & Bliese, P. D. (2009). General mental ability and two types of adaptation to unforeseen change: Applying discontinuous growth models to the task-change paradigm. *Journal of Applied Psychology*, *94*, 411–428. https://doi.org/10.1037/a0013803

Lang, J. W. B., Zettler, I., Ewen, C., & Hülsheger, U. R. (2012). Implicit motives, explicit traits, and task and contextual performance at work. *Journal of Applied Psychology*, *97*, 1201–1217. https://doi.org/10.1037/a0029556

Lievens, F., Chasteen, C. S., Day, E. A., & Christiansen, N. D. (2006). Large-scale investigation of the role of trait activation theory for understanding assessment center convergent and discriminant validity. *Journal of Applied Psychology*, *91*, 247–258. https://doi.org/10.1037/0021-9010.91.2.247

Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment tests: A review of recent research. *Personnel Review*. https://doi.org/10.1108/00483480810877598

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The Scientific Status of Projective Techniques. *Psychological Science in the Public Interest*, *1*, 27–66. https://doi.org/10.1111/1529-1006.002

Lundy, A. (1988). Instructional Set and Thematic Apperception Test Validity. *Journal of Personality Assessment*, *52*, 309–320. https://doi.org/10.1207/s15327752jpa5202_12

McAbee, S. T., & Connelly, B. S. (2016). A multi-rater framework for studying personality: The trait-reputation-identity model. *Psychological Review*, *123*, 569–591.

https://doi.org/10.1037/rev0000035

McAdams, D. P. (1997). A Conceptual History of Personality Psychology. *Handbook of Personality Psychology*, 3–39. https://doi.org/10.1016/B978-012134645-4/50002-0

McClelland, D. C. (1980). Motive dispositions: The merits of operant and respondent measures. In L. Wheller (Ed.), *Review of personality andsocialpsychology(* (pp. 10–41). Beverly Hills: CA: Sage.

McClelland, D. C. (1985). *Human Motivation*. Glenview, IL: Scott Foresman.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). The measuring instrument. In *The Achievement Motive* (pp. 185–217). East Norwalk, CT, US: Appleton-Century-Crofts.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*, 690–702. https://doi.org/10.1037/0033-295X.96.4.690

McClelland, D. C., & Liberman, A. M. (1949). The Effect of Need for Achievement on Need-Related Words. *Journal of Personality*, *18*, 236–251. https://doi.org/10.1111/j.1467-6494.1949.tb01243.x

Meehl, P. E. (1992). Needs (Murray, 1938) and State-Variables (Skinner, 1938). *Psychological Reports*, *70*, 407–450. https://doi.org/10.2466/pr0.1992.70.2.407

Mihura, J. L., Meyer, G. J., Dumitrascu, N., & Bombel, G. (2013). The validity of individual Rorschach variables: Systematic reviews and meta-analyses of the comprehensive system. *Psychological Bulletin*, *139*, 548–605. https://doi.org/10.1037/a0029406

Murstein, B. I. (1965). Projection of hostility on the TAT as a function of stimulus, background, and personality variables. *Journal of Consulting Psychology*, *29*, 43–48. https://doi.org/10.1037/h0021666

Ones, D., & Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance*, *11*, 245–269. https://doi.org/10.1207/s15327043hup1102&3_7

Open Science Collaboration, O. S. (2015). Estimating the reproducibility of psychological science. *Science*, *349*. https://doi.org/10.1126/science.aac4716

Pang, J. S. (2010). Content coding methods in implicit motive assessment: Standards of

measurement and best practices for the picture story exercise. In O. C. Schultheiss & J. C. Brunstein (Eds.), *Implicit Motives* (pp. 119–150). New York: Oxford University Press.

Pang, J. S., & Schultheiss, O. C. (2005). Assessing implicit motives in U.S. college students: Effects of picture type and position, gender and ethnicity, and cross-cultural comparisons. *Journal of Personality Assessment*, *85*, 280–294. https://doi.org/10.1207/s15327752jpa8503_04

Parrigon, S., Woo, S. E., Tay, L., & Wang, T. (2017). CAPTION-ing the situation: A lexically-derived taxonomy of psychological situation characteristics. *Journal of Personality and Social Psychology*, *112*, 642–681. https://doi.org/10.1037/pspp0000111

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge, England: Cambridge University Press.

Rauthmann, J. F., Gallardo-Pujol, D., Guillaume, E. M., Todd, E., Nave, C. S., Sherman, R. A., … Funder, D. C. (2014). The Situational Eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology*, *107*, 677–718. https://doi.org/10.1037/a0037250

Revelle, W. (1986). Motivation and Efficiency of Cognitive Performance BT - Frontiers of Motivational Psychology: Essays in Honor of John W. Atkinson. In D. R. Brown & J. Veroff (Eds.) (pp. 107–131). New York, NY: Springer US. https://doi.org/10.1007/978-1-4684-6341-5_7

Revelle, W., & Condon, D. M. (2015). A model for personality at three levels. *Journal of Research in Personality*, *56*, 70–81. https://doi.org/10.1016/j.jrp.2014.12.006

Robertson, I. T., & Smith, M. (2001). *Personnel selection*. *Journal of Occupational and Organizational Psychology*. Retrieved from https://onlinelibrary.wiley.com/doi/pdf/10.1348/096317901167479

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 501–518. https://doi.org/10.1037/0278-7393.13.3.501

Scheffer, D. (2001). *Entwicklungsbedingungen impliziter Motive: Bindung, Leistung & Macht*

[Developmental conditions of implicit motives: Attachment, achievement and Power]. University of Osnabrück. Retrieved from http://d-nb.info/962853496/34

Schmalt, H. D. (1999). Assessing the achievement motive using the grid technique. *Journal of Research in Personality*, *33*, 109–130. https://doi.org/10.1006/jrpe.1999.2245

Schüler, J., Baumann, N., Chasiotis, A., Bender, M., & Baum, I. (2019). Implicit motives and basic psychological needs. *Journal of Personality*, *87*, 37–55. https://doi.org/10.1111/jopy.12431

Schüler, J., Brandstätter, V., Wegner, M., & Baumann, N. (2015). Testing the convergent and discriminant validity of three implicit motive measures: PSE, OMT, and MMG. *Motivation and Emotion*, *39*, 839–857. https://doi.org/10.1007/s11031-015-9502-1

Schultheiss, O. C. (2007). A memory-systems approach to the classification of personality tests: Comment on meyer and kurtz (2006). *Journal of Personality Assessment*, *89*, 197–201. https://doi.org/10.1080/00223890701357431

Schultheiss, O. C. (2013). The hormonal correlates of implicit motives. *Social and Personality Psychology Compass*, *7*, 52–65. JOUR.

Schultheiss, O. C., & Brunstein, J. C. (2010). Introduction. In *Implicit Motives2*.

Schultheiss, O. C., Liening, S. H., & Schad, D. (2008). The reliability of a picture story exercise measure of implicit motives: Estimates of internal consistency, retest reliability, and ipsative stability. *Journal of Research in Personality*, *42*, 1560–1571. https://doi.org/10.1016/j.jrp.2008.07.008

Schultheiss, O. C., & Pang, J. S. (2007). Measuring implicit motives. In R. W. Robins, R. C. Fraley, & R. Krueger (Eds.), *Handbook of Research Methods in Personality Psychology* (pp. 322–344). New York: Guilford.

Schultheiss, O. C., Yankova, D., Dirlikov, B., & Schad, D. J. (2009). Are implicit and explicit motive measures statistically independent? A fair and balanced test using the picture story exercise and a cue- and response-matched questionnaire measure. *Journal of Personality Assessment*, *91*, 72–81. https://doi.org/10.1080/00223890802484456

Sherwood, J. J. (1966). Self-report and projective measures of achievement and affiliation. *Journal of Consulting Psychology*, *30*, 329–337. https://doi.org/10.1037/h0023574

Skinner, B. F. (1938). *The behavior of organisms: an experimental analysis.* New York:

Appleton-Century.

Slabbinck, H., De Houwer, J., & Van Kenhove, P. (2011). A pictorial attitude IAT as a measure of implicit motives. *European Journal of Personality*, *25*, 76–86. https://doi.org/10.1002/per.778

Slabbinck, H., De Houwer, J., & Van Kenhove, P. (2012). The pictorial attitude implicit association test for need for affiliation. *Personality and Individual Differences*, *53*, 838–842. https://doi.org/10.1016/J.PAID.2012.06.016

Smith, C. P. (1992). *Motivation and personality: Handbook of thematic content analysis*. New York: Cambridge University Press.

Sokolowski, K., Schmalt, H. D., Langens, T. A., & Puca, R. M. (2000). Assessing achievement, affiliation, and power motives all at once: The multi-motive grid (MMG). *Journal of Personality Assessment*, *74*, 126–145. https://doi.org/10.1207/S15327752JPA740109

Sonnentag, S. (2001). Work, recovery activities, and individual well-being: A diary study. *Journal of Occupational Health Psychology*, *6*, 196–210. https://doi.org/10.1037/1076-8998.6.3.196

Sonnentag, S. (2003). Recovery, work engagement, and proactive behavior: A new look at the interface between nonwork and work. *Journal of Applied Psychology*, *88*, 518–528. https://doi.org/10.1037/0021-9010.88.3.518

Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin*, *112*, 140–154. https://doi.org/10.1037/0033-2909.112.1.140

Specht, J., Bleidorn, W., Denissen, J. J. A., Hennecke, M., Hutteman, R., Kandler, C., … Zimmermann, J. (2014). What drives adult personality development? A comparison of theoretical perspectives and empirical evidence. *European Journal of Personality*, *28*, 216–230. https://doi.org/10.1002/per.1966

Tett, R. P., & Guterman, H. A. (2000). Situation trait relevance, trait expression, and cross-situational consistency: Testing a principle of trait activation. *Journal of Research in Personality*, *423*, 397–423. https://doi.org/10.1006/jrpe.2000.2292

Thielgen, M. M., Krumm, S., Rauschenbach, C., & Hertel, G. (2015). Older but wiser: Age

moderates congruency effects between implicit and explicit motives on job satisfaction. *Motivation and Emotion*, *39*, 182–200. https://doi.org/10.1007/s11031-014-9448-8

Thrash, T. M., Cassidy, S. E., Maruskin, L. A., & Elliot, A. J. (2010). Factors that influence the relation between implicit and explicit motives : A general implicit – explicit congruence framework. In O. C. Schultheiss & J. Brunstein (Eds.), *Implicit Motives* (pp. 308–346). New York: Oxford University Press. https://doi.org/10.1093/acprof

Valero, D., Nikitin, J., & Freund, A. M. (2015). The effect of age and time perspective on implicit motives. *Motivation and Emotion*, *39*, 175–181.

Vazire, S., & Carlson, E. N. (2010). Self-knowledge of personality: Do people know themselves? *Social and Personality Psychology Compass*, *4*, 605–620. https://doi.org/10.1111/j.1751-9004.2010.00280.x

Weinberger, J., Cotler, T., & Fishmen, D. (2010). Clinical implications of implicit motives. In O. C. Schultheiss & J. C. Brunstein (Eds.), *Implicit Motives* (p. 468487). New York: Oxford University Press.

Westen, D., Lohr, N., Silk, K. R., Gold, L., & Kerber, K. (1990). Object relations and social cognition in borderlines, major depressives, and normals: A Thematic Apperception Test analysis. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*. https://doi.org/10.1037/1040-3590.2.4.355

Wille, B., & De Fruyt, F. (2014). Vocations as a source of identity: Reciprocal relations between big five personality traits and RIASEC characteristics over 15 years. *Journal of Applied Psychology*, *99*, 262–281. https://doi.org/10.1037/a0034917

Winter, D. G. (1987). Leader Appeal, Leader Performance, and the Motive Profiles of Leaders and Followers: A Study of American Presidents and Elections. *Journal of Personality and Social Psychology*, *52*, 196–202. https://doi.org/10.1037/0022-3514.52.1.196

Winter, D. G., John, O. P., Stewart, A. J., Klohnen, E. C., & Duncan, L. E. (1998). Traits and motives: Toward an integration of two traditions in personality research. *Psychological Review*, *105*, 230–250. https://doi.org/10.1037/0033-295X.105.2.230

Ziegler, M., Horstmann, K. T., & Ziegler, J. (in press). Personality in Situations: Going Beyond the OCEAN and Introducing the Situation Five. *Psychological Assessment*.

Ziegler, M., Schmukle, S., Egloff, B., & Bühner, M. (2010). Investigating measures of

achievement motivation(s). *Journal of Individual Differences*, *31*, 15–21. https://doi.org/10.1027/1614-0001/a000002

ENGLISH SUMMARY

Industrial and organizational psychologists have long been interested in understanding the role of personality in organizational behavior and studied the usefulness of personality measures in personnel selection (Sackett, Lievens, Van Iddekinge, & Kuncel, 2017). The focus of this research has been on explicit personality constructs. Explicit personality constructs are typically measured using self-ratings describing personality with statements that reflect typical behavior. Most explicit measures are studied as trait concept (Allport, 1931; John & Strivastava, 1999) and describe individual differences in behavior that is (mostly) consistent over situations. Explicit personality research has been and still is very successful in contributing to the understanding of organizational behavior.

At the same time that trait research started in modern psychology (Allport, 1931), another research tradition was initiated, that explains individual differences with implicit personality constructs (Morgan & Murray, 1935; Murray, 1938). Implicit personality constructs describe nonconscious attributes such as motivational needs or implicit motives. Because people cannot access their implicit motives through introspection, they cannot be measured using questionnaires like explicit personality measures. To assess implicit motives, researchers present pictures with ambiguous social stimuli and respondents are instructed to imagine and write stories based on the pictures. The stories then get coded into different motivational categories by expert coders. Implicit motives have also been in the focus of organizational researchers in the 1970ies and 1980ies, but research waned due to substantial criticism of implicit motive measures. The main criticism focused on reports of low reliabilities using internal consistency measures (Entwisle, 1972; Fineman, 1977; Lilienfeld, Wood, & Garb, 2000). As a consequence of the criticism, researchers with interest in individual differences of motivational constructs such as achievement motivation focused on explicit measures of those constructs. Another reason for the discontinuation of this research line is that implicit motives are difficult to measure because the expert coding procedure is resource intensive.

Recent research has proposed that traditional measures of internal consistency do not adequately account for the response process of implicit motives measures. Using an archival dataset, Lang (2014) could show that reliability is substantially higher using an IRT model based on Thurstonian coding of forced choice responses accounting for the multidimensional nature and dynamic changes in motivation during measurement. Recent publications have begun to revive implicit motive research studying the relationship with job-related variables (Apers, Lang, & Derous, 2018; Lang, Zettler, Ewen, & Hülsheger, 2012).

Within this context, the present dissertation studies psychometric properties of modern implicit motive measures and innovates their measurement to promote the use in applied contexts. **Chapter 1** describes the historical development of implicit motives including the major criticism but also the achievements of researcher studying the relationship between implicit motives and organizational behavior. In **Chapter 2**, we adapt the dynamic Thurstonian IRT model (Lang, 2014) to model implicit motives measured with the Operant Motive Test (OMT; Kuhl & Scheffer, 2001) and study different forms of dynamic effects. Using this psychometric framework, we show that the OMT is a reliable measure of implicit motives. In **Chapter 3**, we expand the dynamic Thurstonian IRT model to study differential item functioning. This procedure identifies items with different parameters for people from different groups that have the same motive strength. Using this procedure, we show that an OMT picture set developed for cultural comparison shows little differential item functioning and build a partial invariance model that accounts for the small differences in some items we found. Based on the partial invariance model, we report mean differences between three cultures that are not affected by biased items. In **Chapter 4**, we study the relationship between implicit motives and counterproductive work behavior. More specifically, we theorize that people with a high affiliation and power motive show less counterproductive work behavior. We test this idea using two studies that measure counterproductive work behavior with self and supervisor ratings, respectively. We find support for our hypothesis and show that implicit motives predict counterproductive work behavior, when explicit personality measures are considered. In **Chapter 5**, we introduce a new procedure—the Motive Self Categorization (MSC) test—to measure implicit motives that allows respondents to identify motivational

themes in the stories they have written. The MSC does not rely on expert raters. The new measure shows a similar reliability as expert coded measures, shows convergent validity with expert coding, and discriminant validity with explicit personality measures as well as between measured motives. The MSC is a promising innovation that reduces resources necessary to measure implicit motives. In the final **Chapter 6**, we discuss the contribution of this dissertation in terms of measuring implicit motives better in an applied context, point to shortcomings, and suggest new directions for future research with implicit motives in organizational behavior.

References

Allport, G. W. (1931). What is a trait of personality? *Journal of Abnormal and Social Psychology*, *25*, 368–372.

Apers, C., Lang, J. W. B., & Derous, E. (2018). Who earns more? Explicit traits, implicit motives and income growth trajectories. *Journal of Vocational Behavior*. https://doi.org/10.1016/J.JVB.2018.12.004

Entwisle, D. R. (1972). To Dispel Fantasies About Fantasy-Based Measures of Achievement Motivation. *Psychological Bulletin*, *77*, 377–391. https://doi.org/10.1037/h0020021

Fineman, S. (1977). The achievement motive construct and its measurement: Where are we now? *British Journal of Psychology*, *68*, 1–22. https://doi.org/10.1111/j.2044-8295.1977.tb01554.x

John, O. P., & Strivastava, S. (1999). The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of Personality: Theory and Research* (2nd ed., p. 102–138). New York: Guilford Press.

Kuhl, J., & Scheffer, D. (2001). Der operante multi-motiv-test (OMT): Manual [The operant multi-motive-test (OMT): Manual]. Osnabrück, Germany: University of Osnabrück.

Lang, J. W. B. (2014). A dynamic Thurstonian item response theory of motive expression in the picture story exercise: Solving the internal consistency paradox of the PSE. *Psychological Review*, *121*, 481–500. https://doi.org/10.1037/a0037011

Lang, J. W. B., Zettler, I., Ewen, C., & Hülsheger, U. R. (2012). Implicit motives, explicit traits, and task and contextual performance at work. *Journal of Applied Psychology*, *97*, 1201–1217. https://doi.org/10.1037/a0029556

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The Scientific Status of Projective Techniques. *Psychological Science in the Public Interest*, *1*, 27–66. https://doi.org/10.1111/1529-1006.002

Morgan, C. D., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology & Psychiatry*, *34*, 289–306. https://doi.org/10.1001/archneurpsyc.1935.02250200049005

Murray, H. A. (1938). *Explorations in Personality; A clinical and Experimental Study of Fifty*

Men of College Age. New York: Oxfort University Press. https://doi.org/10.2307/2084329

Sackett, P. R., Lievens, F., Van Iddekinge, C. H., & Kuncel, N. R. (2017). Individual differences and their measurement: A review of 100 years of research. *Journal of Applied Psychology, 102*, 254–273. https://doi.org/10.1037/apl0000151

NEDERLANDSTALIGE SAMENVATTING

Arbeids- en organisatiepsychologen zijn reeds lange tijd geïnteresseerd in de precieze rol van persoonlijkheid in organisatiecontexten en in de waarde van persoonlijkheidsinstrumenten in de context van personeelsselectie (Sackett, Lievens, Van Iddekinge, & Kuncel, 2017). Binnen dit omvangrijke onderzoeksdomein lag de focus nagenoeg exclusief op expliciete persoonlijkheidsconstructen. Dergelijke expliciete persoonlijkheidsconstructen worden typisch in kaart gebracht aan de hand van zelfrapporteringen waarbij stellingen worden beoordeeld die verwijzen naar hoe men zich doorgaans gedraagt. De meeste expliciete persoonlijkheidsmaten worden bestudeerd als trek concept (Allport, 1931; John & Strivastava, 1999) en beschrijven individuele verschillen in gedragingen die (grotendeels) consistent zijn over situaties heen. Onderzoek naar expliciete persoonlijkheid blijft een belangrijke bijdrage leveren aan het verklaren en begrijpen van gedrag in organisatiecontexten.

Ongeveer gelijktijdig aan de start van het onderzoek naar (expliciete) persoonlijkheidstrekken (Allport, 1931) startte ook een tweede onderzoekstraditie waarin individuele verschillen worden verklaard aan de hand van impliciete persoonlijkheidsconstructen (Morgan & Murray, 1935; Murray, 1938). Impliciete persoonlijkheidsconstructen beschrijven onbewuste eigenschappen zoals motivationele behoeften of impliciete motieven. Omdat mensen via introspectie geen toegang hebben tot deze impliciete motieven, kunnen deze—in tegenstelling tot expliciete persoonlijkheidsconstructen—niet worden gemeten aan de hand van vragenlijsten. Het in kaart brengen van impliciete motieven gebeurt aan de hand van afbeeldingen van ambigue sociale stimuli, waarrond respondenten een verhaal dienen te bedenken en op te schrijven. Deze verhalen worden vervolgens door expert codeurs geclassificeerd in verschillende motivatie categorieën. Impliciete motieven werden reeds bestudeerd binnen een organisationele context in de periode tussen 1970 en 1990, waarna de onderzoeksinteresse opnieuw afnam omwille van een aantal hardnekkige kritieken op het gebruik van de gehanteerde meetinstrumenten. De voornaamste kritiek richtte zich op het gegeven dat maten van interne consistentie lage betrouwbaarheden voor deze instrumenten opleverden (Entwisle, 1972; Fineman, 1977; Lilienfeld, Wood, & Garb, 2000). Als gevolg

hiervan wendden onderzoekers met interesse in individuele verschillen rond motivatie constructen, bijvoorbeeld prestatiemotivatie, zich nagenoeg exclusief tot het gebruik van expliciete maten. Een tweede reden waarom onderzoek rond impliciete motieven werd afgebroken heeft te maken met het gegeven dat expert coderingen een eerder intensieve procedure vereisen.

Recent onderzoek wees erop dat traditionele maten van interne consistentie niet adequaat omgaan met de processen onderliggend aan impliciete motivatie maten. Lang (2014) toonde op grond van een bestaande (archief)dataset aan dat deze betrouwbaarheid substantieel hoger is wanneer gebruik gemaakt wordt van een IRT model gebaseerd op 'Thurstonian coding'. Deze manier van codering van gedwongen antwoordopties houdt rekening met het multidimensionale karakter van motieven en met dynamische veranderingen in motieven tijdens de meting. Recente publicaties leidden tot een heropleving van het onderzoek naar impliciete motieven en toonden de relatie aan met jobgerelateerde variabelen (Apers, Lang, & Derous, 2018; Lang, Zettler, Ewen, & Hülsheger, 2012).

Binnen deze context heeft voorliggend proefschrift als doel om de psychometrische eigenschappen verder in kaart te brengen van hedendaagse maten van impliciete motieven. Daarnaast wordt getracht om via innovaties in deze meetmethoden het gebruik van impliciete maten in toegepaste (organisatie) contexten te promoten. **Hoofdstuk 1** beschrijft de historische ontwikkeling van impliciete motieven met bijzondere aandacht voor zowel de belangrijkste kritieken op deze traditie als de belangrijkste verwezenlijkingen van onderzoek omtrent de relatie tussen deze motieven en gedrag binnen organisaties. In **Hoofdstuk 2** wordt een aanpassing voorgesteld van het 'dynamische Thurstonian IRT model' (Lang, 2014) om impliciete motieven te modelleren zoals gemeten aan de hand van de 'Operant Motive Test' (OMT; Kuhl & Scheffer, 2001) en worden bovendien verschillende vormen van dynamische effecten onderzocht. Op grond van dit psychometrisch raamwerk wordt aangetoond dat de OMT een betrouwbare maat is voor impliciete motieven. In **Hoofdstuk 3** wordt het dynamische Thurstonian IRT model uitgebreid om 'differential item functioning' in kaart te brengen. Deze procedure identificeert items met verschillende parameters voor mensen uit verschillende groepen die even hoog (of laag) scoren op een bepaald motief. Op basis van

deze procedure wordt aangetoond dat een set van OMT afbeeldingen, ontwikkeld voor culturele vergelijking, weinig 'differential item functioning' vertoont. Daarnaast wordt ook een 'partial invariance model' voorgesteld waarbij rekening wordt gehouden met de kleine verschillen die voor sommige items werden aangetoond. Op grond van dit 'partial invariance model' worden verschillen gerapporteerd tussen drie culturen waarbij voor item bias wordt gecorrigeerd. In **Hoofdstuk 4** wordt de relatie bestudeerd tussen impliciete motieven en contraproductief werkgedrag. De hypothese is dat mensen met hoge affiliatie en macht motieven minder contraproductief werkgedrag zouden vertonen. Dit idee wordt getoetst aan de hand van twee studies waarin contraproductief werkgedrag wordt gemeten aan de hand van respectievelijk zelfbeoordelingen en beoordelingen gemaakt door bovengeschikten. De resultaten bieden steun voor onze centrale hypothese en tonen dat impliciete motieven een voorspeller zijn van contraproductief werkgedrag, waarbij ook rekening wordt gehouden met expliciete persoonlijkheidsconstructen. In **Hoofdstuk 5** wordt een nieuwe procedure geïntroduceerd om impliciete motieven te meten –de 'Motive Self Categorization' (MCS) test– die respondenten toelaat om motivationele thema's te identificeren in de verhalen die ze hebben geschreven. Deze MSC maakt geen gebruik van expert beoordelaars. De nieuwe methode toont een niveau van betrouwbaarheid dat gelijkaardig is aan dat van expert coderingen, toont convergente validiteit met expert coderingen, en discriminante validiteit ten opzichte van zowel expliciete persoonlijkheidsmaten als ten opzichte van verschillende gemeten motieven. De MSC is een veelbelovende innovatie die minder veeleisend is in termen van benodigde hulpmiddelen om impliciete motieven te meten. In het concluderende **Hoofdstuk 6** wordt dieper ingegaan op de contributie van voorliggend proefschrift met betrekking tot het meten van impliciete motieven in een toegepaste context; worden beperkingen van het onderzoek besproken; en worden tenslotte een aantal richtingen voor toekomstig onderzoek rond impliciete motieven in een organisatiecontext voorgesteld.

References

Allport, G. W. (1931). What is a trait of personality? *Journal of Abnormal and Social Psychology*, *25*, 368–372.

Apers, C., Lang, J. W. B., & Derous, E. (2018). Who earns more? Explicit traits, implicit motives and income growth trajectories. *Journal of Vocational Behavior*. https://doi.org/10.1016/J.JVB.2018.12.004

Entwisle, D. R. (1972). To Dispel Fantasies About Fantasy-Based Measures of Achievement Motivation. *Psychological Bulletin*, *77*, 377–391. https://doi.org/10.1037/h0020021

Fineman, S. (1977). The achievement motive construct and its measurement: Where are we now? *British Journal of Psychology*, *68*, 1–22. https://doi.org/10.1111/j.2044-8295.1977.tb01554.x

John, O. P., & Strivastava, S. (1999). The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of Personality: Theory and Research* (2nd ed., p. 102–138). New York: Guilford Press.

Kuhl, J., & Scheffer, D. (2001). Der operante multi-motiv-test (OMT): Manual [The operant multi-motive-test (OMT): Manual]. Osnabrück, Germany: University of Osnabrück.

Lang, J. W. B. (2014). A dynamic Thurstonian item response theory of motive expression in the picture story exercise: Solving the internal consistency paradox of the PSE. *Psychological Review*, *121*, 481–500. https://doi.org/10.1037/a0037011

Lang, J. W. B., Zettler, I., Ewen, C., & Hülsheger, U. R. (2012). Implicit motives, explicit traits, and task and contextual performance at work. *Journal of Applied Psychology*, *97*, 1201–1217. https://doi.org/10.1037/a0029556

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The Scientific Status of Projective Techniques. *Psychological Science in the Public Interest*, *1*, 27–66. https://doi.org/10.1111/1529-1006.002

Morgan, C. D., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology & Psychiatry*, *34*, 289–306. https://doi.org/10.1001/archneurpsyc.1935.02250200049005

Murray, H. A. (1938). *Explorations in Personality; A clinical and Experimental Study of Fifty*

Men of College Age. New York: Oxfort University Press. https://doi.org/10.2307/2084329

Sackett, P. R., Lievens, F., Van Iddekinge, C. H., & Kuncel, N. R. (2017). Individual differences and their measurement: A review of 100 years of research. *Journal of Applied Psychology, 102*, 254–273. https://doi.org/10.1037/apl0000151

DATA STORAGE FACT SHEETS

% Data Storage Fact Sheet

% Name/identifier study: Chapter 2
% Author: Malte Runge
% Date: 16.5.2019

1. Contact details
==========

1a. Main researcher

- name: Malte Runge
- address: Henri Dunantlaan 2, B-9000 Gent
- e-mail: malte.runge@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Jonas Lang
- address: Henri Dunantlaan 2, B-9000 Gent
- e-mail: jonas.lang@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
==========
* Reference of the publication in which the datasets are reported:

Runge, J. M., Lang, J. W. B., Engeser, S., Schüler, J., den Hartog, S. C., & Zettler, I. (2016). Modeling motive activation in the Operant Motive Test: A psychometric analysis using

dynamic Thurstonian item response theory. *Motivation Science*, 2, 268-286. http://dx.doi.org/10.1037/mot0000041

* Which datasets in that publication does this sheet apply to?:

The sheet applies to all the data used in Chapter 2.

3. Information about the files that have been stored
==========

3a. Raw data
----

* Have the raw data been stored by the main researcher? [ ] YES / [X] NO
If NO, please justify:

The raw data were collected by and are owned by co-authors.

* On which platform are the raw data stored?
 - [ ] researcher PC
 - [ ] research group file server
 - [X] other (specify): Raw data are owned by co-authors

* Who has direct access to the raw data (i.e., without intervention of another person)?
 - [ ] main researcher
 - [ ] responsible ZAP
 - [ ] all members of the research group
 - [ ] all members of UGent
 - [X] other (specify): co-authors

3b. Other files
----

* Which other files have been stored?
 - [X] file(s) describing the transition from raw data to reported results. Specify: R-Script
 - [X] file(s) containing processed data. Specify: Coded OMT data.
 - [X] file(s) containing analyses. Specify: R-Script

- [ ] files(s) containing information about informed consent
- [ ] a file specifying legal and ethical provisions
- [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...
- [ ] other files. Specify: ...

* On which platform are these other files stored?
- [X] individual PC
- [X] research group file server
- [ ] other: ...

* Who has direct access to these other files (i.e., without intervention of another person)?
- [X] main researcher
- [X] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): ...

4. Reproduction
==========
* Have the results been reproduced independently?: [ ] YES / [X] NO

% Data Storage Fact Sheet

% Name/identifier study: Chapter 3
% Author: Malte Runge
% Date: 16.5.2019

1. Contact details
==========

1a. Main researcher

- name: Malte Runge
- address: Henri Dunantlaan 2, B-9000 Gent
- e-mail: malte.runge@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Jonas Lang
- address: Henri Dunantlaan 2, B-9000 Gent
- e-mail: jonas.lang@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
==========
* Reference of the publication in which the datasets are reported:

Runge, J. M., Lang, J.W.B., Chasiotis, A., & Hofer, J. (2018). Improving the assessment of implicit motives using IRT: Cultural differences and differential item functioning. *Journal of Personality Assessment*. http://dx.doi.org/10.1080/00223891.2017.1418748

* Which datasets in that publication does this sheet apply to?:

The sheet applies to all the data used in Chapter 3.

3. Information about the files that have been stored
==========

3a. Raw data
---

* Have the raw data been stored by the main researcher? [ ] YES / [X] NO
If NO, please justify:

The raw data were collected by and are owned by co-authors.

* On which platform are the raw data stored?
 - [ ] researcher PC
 - [ ] research group file server
 - [ ] other (specify): ...

* Who has direct access to the raw data (i.e., without intervention of another person)?
 - [ ] main researcher
 - [ ] responsible ZAP
 - [ ] all members of the research group
 - [ ] all members of UGent
 - [X] other (specify): co-author

3b. Other files
---

* Which other files have been stored?
 - [X] file(s) describing the transition from raw data to reported results. Specify: R-Script
 - [X] file(s) containing processed data. Specify: Coded OMT data
 - [X] file(s) containing analyses. Specify: R-Script
 - [ ] files(s) containing information about informed consent
 - [ ] a file specifying legal and ethical provisions
 - [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...
 - [ ] other files. Specify: ...

* On which platform are these other files stored?
  - [X] individual PC
  - [X] research group file server
  - [ ] other: ...

* Who has direct access to these other files (i.e., without intervention of another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [ ] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction
==========
* Have the results been reproduced independently?: [ ] YES / [X] NO

% Data Storage Fact Sheet

% Name/identifier study: Chapter 4
% Author: Malte Runge
% Date: 16.5.2019

1. Contact details
==========

1a. Main researcher

- name: Malte Runge
- address: Henri Dunantlaan 2, B-9000 Gent
- e-mail: malte.runge@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Jonas Lang
- address: Henri Dunantlaan 2, B-9000 Gent
- e-mail: jonas.lang@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
==========
* Reference of the publication in which the datasets are reported:

Runge, J. M. & Lang, J. W. B, Zettler, I., & Lievens, F. (under review). Predicting counterproductive work behavior: Do implicit motives have incremental validity beyond explicit traits?

* Which datasets in that publication does this sheet apply to?:

The sheet applies to all the data used in Chapter 4.

3. Information about the files that have been stored
==========

3a. Raw data
---

* Have the raw data been stored by the main researcher? [X] YES / [ ] NO
If NO, please justify:

* On which platform are the raw data stored?
 - [X] researcher PC
 - [X] research group file server
 - [ ] other (specify): ...

* Who has direct access to the raw data (i.e., without intervention of another person)?
 - [X] main researcher
 - [X] responsible ZAP
 - [ ] all members of the research group
 - [ ] all members of UGent
 - [ ] other (specify): ...

3b. Other files
---

* Which other files have been stored?
 - [X] file(s) describing the transition from raw data to reported results. Specify: R-Script
 - [ ] file(s) containing processed data. Specify: ...
 - [X] file(s) containing analyses. Specify: R-Script
 - [ ] files(s) containing information about informed consent
 - [ ] a file specifying legal and ethical provisions
 - [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...
 - [ ] other files. Specify: ...

* On which platform are these other files stored?
 - [X] individual PC

- [X] research group file server
- [ ] other: ...

* Who has direct access to these other files (i.e., without intervention of another person)?
- [X] main researcher
- [ ] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): ...

4. Reproduction
==========
* Have the results been reproduced independently?: [ ] YES / [X] NO

% Data Storage Fact Sheet

% Name/identifier study: Chapter 5
% Author: Malte Runge
% Date: 16.5.2019

1. Contact details
==========

1a. Main researcher

- name: Malte Runge
- address: Henri Dunantlaan 2, B-9000 Gent
- e-mail: malte.runge@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Jonas Lang
- address: Henri Dunantlaan 2, B-9000 Gent
- e-mail: jonas.lang@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
==========
* Reference of the publication in which the datasets are reported:

Runge, J. M. & Lang, J. W. B (2019). Can people recognize their implicit thoughts? The motive self-categorization test. *Psychological Assessment*. http://dx.doi.org/10.1037/pas0000720

* Which datasets in that publication does this sheet apply to?:

The sheet applies to all the data used in Chapter 5.

3. Information about the files that have been stored
==========

3a. Raw data
---

* Have the raw data been stored by the main researcher? [X] YES / [ ] NO
If NO, please justify:

* On which platform are the raw data stored?
 - [X] researcher PC
 - [X] research group file server
 - [ ] other (specify): ...

* Who has direct access to the raw data (i.e., without intervention of another person)?
 - [X] main researcher
 - [X] responsible ZAP
 - [ ] all members of the research group
 - [ ] all members of UGent
 - [ ] other (specify): ...

3b. Other files
---

* Which other files have been stored?
 - [X] file(s) describing the transition from raw data to reported results. Specify: R-Script
 - [ ] file(s) containing processed data. Specify: ...
 - [X] file(s) containing analyses. Specify: R-Script
 - [X] files(s) containing information about informed consent: Scanned informed consents
 - [ ] a file specifying legal and ethical provisions
 - [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...
 - [ ] other files. Specify: ...

* On which platform are these other files stored?
 - [X] individual PC

- [X] research group file server
- [X] other: Data and Analysis are stored on OSF: https://osf.io/rz4sy/

* Who has direct access to these other files (i.e., without intervention of another person)?
- [X] main researcher
- [X] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): ...

4. Reproduction
==========
* Have the results been reproduced independently?: [ ] YES / [X] NO