
# The Interplay of Personality, Situations, Affect, and Behavior

**DISSERTATION**
zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

eingereicht an der
Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von

**M. Sc. Kai Tobias Horstmann**

Präsidentin der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin:
Prof. Dr. Bernhard Grimm

Gutachter:
1. Prof. Dr. Matthias Ziegler, Humboldt-Universität zu Berlin
2. Prof. Dr. Markus Bühner, Ludwig-Maximilians-Universität München
3. Prof. Dr. David Funder, University of California, Riverside

Tag der mündlichen Prüfung: 25.07.2018

# Acknowledgements

First and foremost, I am very thankful to my supervisor Matthias Ziegler. Matthias was, frankly, the best supervisor I could have imagined. Matthias gave me all the feedback, encouragement, support, and possibilities I could have wished for. Dear Matthias, thank you very much for all that you have done for me, and I look forward to continuing to work with you!

I am also very grateful to have met John Rauthmann at some point in the last three years. Someday, John must have decided that he – by all means – wanted to teach me how to write. All remaining errors are my own, but not making several others I owe to John. Thanks John, for all your feedback, the good food, hospitality, and fun that we had in the last years. Your help and patience were without comparison.

I am thankful to my very good friend and colleague Rasmus Bruckner for his encouragement ("I think you should do it") and helping me navigate through my Ph.D. I would like to thank Jens Nachtwei for his honest words and advice.

Of course, this work would also not have been possible without my colleagues from my workgroup, who supported me on a daily basis. Whether it was taking over my seminars, reading some rough drafts, listening to my talks, and giving me feedback – without them, my time here would not have been as good as it was. Thank you, Caro, for the proof-reads and the coffee-money, Max, for being the best Brokkoli, for the mediocre coffee, and sharing and discussing everything with me, Doreen, for always being helpful, Ulli, for giving advice, and Lena, for thinking about my problems. And special thanks to Sabine Schulz, who took care of everything. Thank you, Inga Meyhöfer, for the long discussions. I would also like to thank the Open Science crowd. Thanks for sharing ideas, materials, code, data, platforms, inspiration, enthusiasm, and giving advice for free. Finally, I would like to thank all student assistants and interns who helped completing these projects. Thank you, Max Knaut and Selina Reinhard, for proof-reading.

Of course, my well-being benefited from my great friends and my supportive family. I would like to thank my Lacrosse team, which provided a great outlet and distraction. Thanks for letting me still play with you, while getting older and slower[1]. I will make sure not to invite you to my defense, but we can meet afterwards in the Anna Hirsch bar?

Very important throughout my life and especially for my academic career are my grand-parents, Doris and Otto Wendt. They have been the backbone of this endeavor, and without their support, I might have slept less and worried more. I am blessed to have such wonderful, open-minded, and curious grand-parents. Thank you for showing me the world!

[1] Ok, let's get serious. I still got most assists this season, so you're lucky to have me.

Equally important were my mother and father and my two sisters. Thank you for being there whenever I needed you, thank you for giving me your unconditional support. I apologize for not having visited as much in the last years, and I am all the more thankful to know that you support me no matter what. Thank you, Christiane, Sophie, Hanna, and Markus for having my back especially in the last six months of my doctoral studies. I owe you.

Finally, I would like to thank my best friend and partner Lena Fliedner for everything she has done for me. Thank you for your support, for enduring my moods, for lifting me up, for sharing some of the best memories in the last years, and simply for being there. More than anything else, I am looking forward to spending time with you.

# Table of Contents

**Acknowledgements** ..... **2**

**Foreword** ..... **6**

A Word on Open Science ..... 7

**Summary** ..... **8**

**Zusammenfassung** ..... **9**

**Introduction** ..... **10**

**1 Part 1: Personality Traits and States** ..... **13**

1.1 The Person-Situation Debate ..... 13

1.2 The Consistency-Problem ..... 14

*1.2.1 Article Summary: Distinguishing Simple and Residual Consistency* ..... *16*

*1.2.2 Article: Distinguishing Simple and Residual Consistency* ..... *17*

1.3 Personality Theories after the Person-Situation Debate ..... 47

1.4 Conclusion Part 1 ..... 49

**2 Part 2: The Description and Measurement of Situations** ..... **50**

2.1 Background: Situation Research ..... 50

2.2 Assessing Situational Information ..... 52

*2.2.1 Book Chapter Summary: Assessment of Situational Perceptions* ..... *53*

*2.2.2 Book Chapter: Assessment of Situational Perceptions* ..... *54*

2.3 The Situation Five ..... 79

*2.3.1 Article Summary: Personality in Situations* ..... *80*

*2.3.2 Article: Personality in Situations* ..... *81*

2.4 Replicable Dimensions of Situation Perception ..... 107

*2.4.1 Book Chapter Summary: Measurement of Situational Influences* ..... *108*

*2.4.2 Book Chapter: Measurement of Situational Influences.* ..... *109*

2.5 Conclusion Part 2 ..... 129

**3 Part 3: Affect and Situation Perception** ..... **130**

3.1 Situation Perception and Appraisal Theories of Emotion ..... 130

*3.1.1 Article Summary: Situational Perception and Affect* ..... *132*

*3.1.2 Article: Situational Perception and Affect* ..... *133*

3.2 Affect and Situation Perception – Independent Predictors of Behavior? ..... 153

*3.2.1 Article Summary: Unveiling an Exclusive Link* ..... *154*

*3.2.2 Article: Unveiling an Exclusive Link* ..... *156*

3.3 Conclusion Part 3 ..... 215

**4 Part 4: Implications and Conclusion ..... 216**

4.1 Status Quo of Situation Research ..... 216
4.2 Assessing Situations via Situation Perceptions ..... 217
4.3 Application of Situational Taxonomies ..... 218
4.4 Process Model of Situation Perception ..... 219
*4.4.1 Overlap of Affect and Situation Perception ..... 224*
*4.4.2 Explaining idiosyncratic and consensual variance in situation perception ..... 225*
*4.4.3 The Role of Time ..... 226*
4.5 Relevance for Personality Theories ..... 227
4.6 Consequences for Consistency ..... 228

**5 Part 5: Outlook ..... 230**

**6 References ..... 232**

**7 Statement on Independence ..... 258**

# Foreword

This work was conducted between April 2015 and Mai 2018. Just prior to my dissertation, situation research had gained new momentum. In 2014, John Rauthmann published the DIAMONDS taxonomy for the description of situations (Rauthmann et al., 2014), and my supervisor Matthias Ziegler had already collected and analyzed all data for the Situation Five taxonomy and the accompanying measurement tool, the Big Five Inventory of Personality in Occupational Situations (B5PS; Ziegler, 2014a). I analyzed the norm data from the B5PS in my master thesis (Horstmann, 2015), and gained a first understanding of the person-situation debate and its impact on personality psychology. After the completion of my master thesis, I wondered how the perception of a situation and a person's current affect in a situation are related. At the same time, in 2015, Rauthmann, Sherman, and Funder (2015b) published a model suggesting a possible effect of affect on situation perception. It was then my goal to test elements of this model during my dissertation.

Since 2015, others have also suggested links between situation perception and affect (e.g., Sherman, Rauthmann, Brown, Serfass, & Jones, 2015; Wilson, Thompson, & Vazire, 2017; Wilt & Revelle, 2017). However, those publications have only partially influenced the work presented here, as some of these articles were not published when I planned and assembled my studies. Furthermore, additional research on situation taxonomies was published in the last three years (N. A. Brown, Neel, & Sherman, 2015; Gerpott, Balliet, Columbus, Molho, & de Vries, 2017; Parrigon, Woo, Tay, & Wang, 2017). The relation of the DIAMONDS and Situation Five dimensions with other dimensions of situation perception then became a second element of my dissertation.

Although this dissertation is divided into several parts, it would be incorrect to suggest that this work was conducted and completed strictly sequentially. Some of the work in the later stages was started before some of the work in the earlier stages, and then informed by later publications or ideas. Furthermore, all research articles or book chapters included in this dissertation were written as individual items. Hence, the chapters and publications overlap, and some content is repeated, sometimes in more or less detail. Furthermore, I changed and adapted some of the wording I used throughout the last three years. For example, situational perception was changed to situation perception, which is a more accurate description of the phenomenon. However, as all book chapters and articles have already been published or are currently under review, I have not changed the wording for this dissertation. The reader is kindly asked to forgive repetitions and inconsistencies that are due to these circumstances. If an article or book chapter has been published, the original publication is referenced at the beginning of each book chapter or article.

## A Word on Open Science

Right at the beginning of my dissertation, the replication crisis gained renewed attention. The term replication crisis refers to the observation that many published research findings may not replicate when conducted a second time. Although this had been suspected previously (e.g., Greenwald, 1976; Ioannidis, 2005), and failed replications can be seen as part of the scientific progress, the magnitude of the problem became apparent for psychology in August 2015, when the Open Science Collaboration published replication attempts from 100 previously published studies (Open Science Collaboration, 2015). The authors concluded that only 38% replications were successful. Although some disagree with this interpretation and argue that replicability is in fact high (e.g., Gilbert, King, Pettigrew, & Wilson, 2016; Stroebe & Strack, 2014), this finding has influenced the field of psychology substantially. For example, it is now more common to publish data and materials with a study, and studies can be pre-registered before data collection has commenced. This affects the way research is conducted as well as how (and which) results are presented in journal articles. The research progress takes longer, and it will be necessary to accept "imperfection in the data" (Cooper, 2016, p. 433), and thus in the results presented. As Cooper puts it, this should not be seen as a "lessening of standards, but rather a broadening of the vision of what constitutes good science" (Cooper, 2016, p. 433).

I personally belief that psychological science needs to be more open and transparent, and I am genuinely convinced that replicability is not as high as we would want it to be. There may be several reasons for it, from outright fraud to sloppiness or simply lack of knowledge. To counter this, research must be, first and foremost, as transparent as possible. This conviction is reflected in my work. From the four empirical articles presented here, two were pre-registered, three feature open data, open materials, and open code, and only one article has openly available code only, due to proprietary reasons.

Of course, it would be immature to claim that this work here is the "End of History." Replication and extension of the results presented here is always welcome, and future studies will surely help shaping a better understanding of the current findings. Having published my research as transparently as possible, I hope that I can contribute to these future endeavors. As is it the case with all scientific claims, one should not belief what has not been backed up by empirical support. I would therefore like to invite the reader to my work with the motto of the Royal Society – *Nullius in verba*, take no one's word.

# Summary

The person-situation debate in psychology was concerned with the following question: Can behavior be seen as a function of the person, and can patterns in behavior thus be explained with personality, or should behavior be understood primarily as a function of the situation, which would render personality traits meaningless (Funder, 2001)? As it is often the case, the truth lied in-between, and to establish consistency in behavior, both the person and the situation must be considered. This, however, meant that descriptive systems for situations had to be developed, and measures for situations had to be devised, validated, and tested empirically.

Part 1 of the current work briefly reviews the person situation debate and addresses the definition of personality traits and states. Central to the person situation debate was the definition of consistency. In the first empirical study presented here, we examined consistency in more detail. Specifically, we distinguished between person and situation effects in consistency research. To this end, we introduced the concepts of simple and residual consistency, and showed how functionally equivalent situations may influence behavior, even after controlling for effects of personality.

Part 2 is concerned with the definition and measurement of situations. As situations are crucial to understanding consistency, they must be defined and measured. In two book chapters, we reviewed the past development of situation research and specific challenges that can be faced during the construction of measures for situation dimensions. We then present a new taxonomy for the description of situations, the Situation Five, as well as a measurement tool to assess situation perception, a person's interpretation and perception of situations.

Part 3 addresses the validation of existing situation measures and, more specifically, the potential overlap of affect and situation perception. I the first empirical study, we examined this overlap, which turned out to be substantial. In the second study, we investigated if this overlap threatened the validity of situation measures: Would measures of situation perception predict behavior in daily life after controlling for affect? As it turned out, they did – and more importantly, controlling for affect unveiled specific, logically coherent links between situation perception and behavior.

Part 4 then discusses implications of the current work. A special focus is placed on explaining why affect and situation perception were correlated and yet contributed uniquely to the explanation of variance in behavior. To summarize, appraisal theories of emotion may serve well as a general framework for understanding the processes involved in situation perception.

Part 5 then briefly discusses implications of the current work for future research.

# Zusammenfassung

Die Person-Situation Debatte befasste sich vorrangig mit der Frage, ob Verhalten eher als Funktion der Person gesehen werden kann, und dieses somit durch Persönlichkeit aufzuklären ist, oder ob Verhalten vor allem von der Situation bestimmt ist, was wiederum Persönlichkeit nichtig werden ließe (Funder, 2001). Wie so oft der Fall, lag die Wahrheit dazwischen: Um Verhalten zu beschreiben, müssen sowohl die Person als auch die Situation berücksichtigt werden. Dies bedeutete jedoch, dass Systeme zur Beschreibung von Situationen und Methoden zu deren Erfassung entwickelt, validiert und getestet werden mussten.

Teil 1 dieser Dissertationsschrift fasst die Person-Situation Debatte zusammen und definiert stabile Persönlichkeitsmerkmale (Traits) sowie deren temporäre Expressionen (States). Zentrales Element der Debatte war unter anderem das Konsistenzproblem, welches in der ersten empirischen Studie weiter beleuchtet wurde. Hierzu unterschieden wir Person- und Situation-Effekte auf Verhalten. Um diese Effekte zu differenzieren schlugen wir die Konzepte simple Konsistenz und residuale Konsistenz vor und zeigten, wie Verhalten in funktional äquivalenten Situationen auch nach Kontrolle des Einflusses von Persönlichkeit konsistent ist.

Teil 2 befasst sich dann mit der Definition und Erfassung von Situationen, da diese zentral sind um Konsistenz im Verhalten zu untersuchen. Im Rahmen von zwei Buchkapiteln beschrieben wir die bisherige Entwicklung der Situationsforschung und zeigten Herausforderungen auf, welche während der Entwicklung von Situationsmaßen zu meistern sind. Anschließend stellten wir sowohl eine neue Taxonomie zur Beschreibung von Situationen vor, die Situation Five, als auch ein Messinstrument zur Erfassung von Situationswahrnehmung vor.

Teil 3 adressiert die Validierung bestehender Situationsmaße und im Besonderen die mögliche Überschneidung von Situationswahrnehmung und Affekt. In zwei Studien wurde untersucht, ob beide Phänomene überlappen und dennoch jeweils Verhalten vorhersagen können. Hierbei konnten wir zeigen, dass dies nicht nur der Fall ist, sondern dass spezifische Verbindungen zwischen Verhalten und Situationswahrnehmung erst nach Berücksichtigung von Affekt sichtbar waren.

Teil 4 diskutiert die Implikationen der vorliegenden Arbeit und zeigt auf, wie insbesondere die Überlappung von Affekt und Situationswahrnehmung durch Einschätzungs-Theorien der Emotionsforschung erklärt werden können.

Teil 5 schlussendlich gibt einen Ausblick auf zukünftige Forschung und zeigt die Bedeutung der vorliegen Arbeit auf.

# Introduction

Exactly half a century ago, in 1968, the person-situation-debate gained new momentum and changed the field of personality psychology (Donnellan, Lucas, & Fleeson, 2009; Mischel, 1968, 2009). At the core of the debate was the question which of the two – persons or situations – exerts a greater or relevant influence on behavior in any given moment. Can behavior best be explained by stable personality traits of a person or with momentary aspects of the environment? Although the distinction between the person and the situation was made as early as 1936 by Kurt Lewin (1936) and others (e.g., Murray, 1938), the importance of the situation to understand and explain behavior became apparent. Several questions arose, one of which was the question of consistency: Behavior can only be predicted or explained by stable characteristics of a person (i.e., personality) if a person acts in a systematic and thus consistent way (Fleeson & Noftle, 2008; Schmitt, 1990b; Sherman, Nave, & Funder, 2010). The first section of this thesis therefore addresses the topic of consistency and behavioral variability and lays the foundation of the remaining work.

To unify intra-individual variability in behavior with the call for consistency, it was important to define and investigate the situations in which behavior occurs (Rauthmann & Sherman, 2018b; Rauthmann, Sherman, & Funder, 2015b). Recognizing the importance of situations as an additional explanatory variable might allow describing how a person acts in a certain way in a given moment. A simple example could be "every Monday morning, this person acts very clumsily". Even though this person usually would not act clumsily, they do so on Monday mornings, thereby acting consistently. It was thus essential to find ways of describing and measuring situations (Horstmann, Rauthmann, & Sherman, 2018). The description and assessment of situations that allow accounting for variability in behavior is therefore at the focus of the second section of this work.

Situations are usually defined via situation perceptions and measured on several independent dimensions (Horstmann, Rauthmann, et al., 2018). Perceptions of situations allow explaining differences in behavior above and beyond the effects of personality traits (e.g., Sherman et al., 2015). At the same time, momentary states of a person (e.g., affect) are also linked to situations and behavior (Cattell, 1963; Kuppens, 2009; Rauthmann, Sherman, & Funder, 2015b) and allow explaining variance in behavior (Lench, Flores, & Bench, 2011; Wilson et al., 2017; Wilt, Bleidorn, & Revelle, 2017). Given the theoretical (Kuppens, 2009; Rauthmann, Sherman, & Funder, 2015b) and empirical (e.g., Edwards & Templeton, 2005; Parrigon et al., 2017; Sherman et al., 2015) overlap between affect and situation perception, it was important to examine and

establish discriminant validity of situational measures with measures of affect. This constitutes the third section of the current work.

Finally, the fourth section discusses the implications of this thesis for personality psychology and the examination of behavioral consistency. Based on appraisal theories of emotion, I suggest how affect and situation perception may be separated conceptually. To summarize, the information gathered during the appraisal process could form situation perception, and the simultaneous evaluation of this information may lead to affect or emotions.

The last section then gives an outlook of future research that might benefit from and could be influenced by the current findings, as an enhanced understanding of behavior in situations and their relation to personality traits and states will allow revising and extending personality theories.

Table 1 gives an overview of some research articles or book chapters that I (co-)authored in the last four years (including my master thesis) and that are relevant for the current dissertation. It also lists which articles are included in this dissertation and which articles may be relevant to the current work, but are not formally included here.

Table 1

*Research Articles and Book Chapters written in the last Four Years*

| Research/Publication | Main Finding/Conclusion |
|---|---|
| **Part 1: Personality Traits and States** | |
| **Horstmann, K. T.** (2015). Putting Lewin's Equation to the Test: Assessing the Person-Situation Interaction with the B5PS. Master thesis, thesiscommons.org/z32ru. | Situation perception predicted behavior in hypothetical situations and is therefore a relevant construct to understand behavioral consistency. |
| **Horstmann, K. T.**, Rauthmann, J. F., & Ziegler, M. (in preparation). Distinguishing simple and residual consistency in functionally equivalent situations: Evidence from variable- and person-centered analyses in longitudinal data. * | People behaved consistently in functionally equivalent situations due to their personality but also due to stable influences of the situation. |
| Rauthmann, J. F., **Horstmann, K. T.**, & Sherman, R. A. (2018). Do self-reported traits and aggregated states capture the same thing? A nomological perspective on trait-state homomorphy. *Social Psychological and Personality Science*. | Personality traits and states are related and share a nomological net. The degree of nomological homomorphy however depends on the trait and state examined. |
| **Part 2: The Description and Measurement of Situations** | |
| Ziegler, M., & **Horstmann, K. T.** (2015). Discovering the Second Side of the Coin. *European Journal of Psychological Assessment*, *31*, 69–74. | Situation perception has a stable, trait-like component. |
| **Horstmann, K. T.**, & Ziegler, M. (2016). Situational Perception: Its Theoretical Foundation, Assessment, and Links to Personality. In U. Kumar (Ed.), *The Wiley Handbook of Personality Assessment* (1st ed., pp. 31–43). Oxford: Wiley Blackwell.* | Situation perception as a personality trait can be measured and assessed, and specific challenges for test-construction arise. |
| **Horstmann, K. T.**, Ziegler, J., & Ziegler, M. (2018). Assessment of Situational Perceptions: Measurement issues and a joint taxonomization of persons and situations. In D. C. Funder, J. F. Rauthmann, & R. A. Sherman (Eds.), *The Oxford Handbook of Psychological Situations*. Oxford University Press. * | The measurement of situation perception, both at trait and state level, requires different approaches, depending on the use of the measure, the assessed population, and the exact construct assessed. |
| Ziegler, M., **Horstmann, K. T.**, & Ziegler, J. (submitted after review). Personality in Situations: Going Beyond the OCEAN and Introducing the Situation Five. ✝,* | Five dimensions of situation perception, the Situation Five, were developed and used to predict relevant outcomes. |
| **Horstmann, K. T.**, Rauthmann, J. F., & Sherman, R. A. (2018). Measurement of situational influences. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *The SAGE Handbook of Personality and Individual Differences*. SAGE Publications. * | Theoretical considerations and empirical evidence support the assumption that six robust and replicable, overarching dimensions of situation perception exist. |
| **Part 3: Affect and Situation Perception** | |
| **Horstmann, K. T.**, & Ziegler, M. (2018). Situational perception and affect: Barking up the wrong tree? *Personality and Individual Differences*.* | Affect and situational perception overlapped substantially, threatening discriminant validity of situation perception scores. |
| **Horstmann, K. T.**, Rauthmann, J. F., Sherman, R. A., & Ziegler, M. (in preparation). Unveiling an Exclusive Link: Predicting Behavior with Personality, Situation Perception, and Affect in a Pre-Registered Experience Sampling Study. * | Affect, situation perception, and personality interpedently predicted behavior in daily life, supporting the discriminant predictive validity of situation perception scores. |

*Note.* * = work formally included in the dissertation; ✝ = shared first authorship

# 1 Part 1: Personality Traits and States

A personality trait is a stable characteristic of a person (Funder, 2001). Examples of personality traits are intelligence, extraversion, or sportiness. A trait can manifest itself in concrete behavior, for example, solving a difficult task, dancing on a party, or physical exercise. The degree to which a trait is related to this manifestation is called trait-expression (Horstmann, Rauthmann, & Ziegler, in preparation). The momentary manifestation itself is a personality state. A personality state has the "same affective, behavioral, and cognitive content as a corresponding trait (…), but as applying for a shorter duration" (Fleeson & Jayawickreme, 2015, p. 84). Traits and states are related: the higher a person scores on a trait, the higher he or she scores – on average – on corresponding states. Yet, a person can act in stark contrast to their general trait level. For example, a person may score highly in likeability (and should therefore, usually, behave likeable), yet there may be occasions during which this person exhibits completely unlikeable behavior. This observation – as trivial as it may look at first sight – is consequential for the conceptualization of personality traits and has influenced personality research over the last 50 years in the context of the person-situation debate.

## 1.1 The Person-Situation Debate

As Funder (2001) put it, the person-situation debate was "concern[ed] whether consistencies in individuals' behavior are pervasive or broad enough to be meaningfully described in terms of personality traits" ( p. 199). In other words, personality traits could only be meaningful and useful if persons acted consistently (Allport, 1936). The view that personality can in fact predict behavior was challenged most prominently (but not exclusively, see Bem & Allen, 1974) by Walter Mischel in 1968. Mischel argued that the observed correlation between instances of behavior and personality traits was $r = .30$ and therefore indicative of behavioral variability rather than behavioral consistency (Mischel, 1968, 1973).

There were several different approaches to deal with this fundamental criticism (Fleeson & Noftle, 2009; Steyer, Schmitt, & Eid, 1999). Some argued that this correlation might in fact be $r = .40$ and thus meaningful (Funder, 2001; Funder & Colvin, 1991). As Funder (2001) reviews, a correlation of .40 means that a binary outcome could be predicted in about 70% of the cases. Such an effect size is comparable to effect sizes from social psychology (F. D. Richard, Bond, & Stokes-Zoota, 2003).

Another approach was to recognize that psychology is not and should not be concerned with predictions of single instances of behavior (Epstein, 1980). The more instances of behavior

are aggregated, the more reliable a measure becomes (W. Brown, 1910; Spearman, 1910). The more reliable a measure is, the higher a correlation with another variable can be (Orom & Cervone, 2009). However, this approach treats intra-individual variability essentially as measurement error and thus discards possibly valuable information on intra-individual variability (Geukes, Nestler, Hutteman, Küfner, & Back, 2017).

Other authors, however, called for the integration of stable personality traits and variable states (Bem & Allen, 1974; Campus, 1974). This is referred to as the moderator approach (Steyer et al., 1999). Bem and Allen (Bem & Allen, 1974) argued that some people may be more consistent than others, and that consistency may thus be a personality characteristic in itself (Bem & Allen, 1972, 1974; Fleeson, 2001), moderating the effects of personality traits on personality states. The more consistently a person behaves, the better can this behavior be predicted with personality traits. Although Bem and Allen's study could not successfully be replicated (Chaplin & Goldberg, 1984), this idea persisted. Schmitt examined extensively why the moderator approach has not been successful (Schmitt, 1990a, 1990b, 1992). He came to the conclusion that measures of self-reported consistency, such as the one used by Bem and Allen (1974), were not very reliable. Asking participants to rate their own behavioral variability from one situation to another would require each participant to form their own judgment about the definition of variability and of situations. As Schmitt (1990b) concludes, the debate about consistency might not even have occurred if these and other methodological and theoretical issues had been considered in the first place. It was therefore necessary to develop a precise definition of consistency.

## 1.2 The Consistency-Problem

As the previous section highlighted, establishing consistency in behavior was a key factor to resolving the person situation debate (Fleeson & Noftle, 2009). However, consistency is difficult to define. Imagine a person was asked the same question that Bem and Allen (1974) asked their participants: "How much do you vary from one situation to another in how friendly and outgoing you are?" This simple question requires a person to make a judgement about the term "varying," and several interpretations are possible: It could either mean "change compared to yourself", simply in terms of more or less friendly and outgoing. However, it could also mean "change compared to others", for example, change more or less in friendliness than others. Furthermore, it also requires making a judgment about the term "situation". For example, it could mean "in similar situations, but at different times", to which we would refer as within-context variability, or it could mean "in different situations", which would nowadays be referred to as across-context

variability (Geukes et al., 2017). This simple example already highlights some of the complexities that surround the term consistency.

Fleeson and Noftle (2008) have examined the concept of consistency in great detail. They defined at least 36 different types of consistency, distinguishing, for example, "changing compared to yourself" as ipsative consistency and "changing compared to others" as rank-order consistency. Although Fleeson and Noftle (2008) defined different types of consistency that could be examined (see section 1.2.2), they did not distinguish *why* a person acts consistently across two situations. As was elaborated during the person-situation debate, this may either be explained in terms of the person or in terms of the current situation. For example, a person may exhibit sportiness in two situations (e.g., run) because he or she generally likes to run. On the other hand, a person may run in two situations, because, at both times, he or she tries to get away from something, even though this person would usually not run on their own account.

These two effects – of the person and of the situation – can be separated, namely by controlling for the general tendency of a person to run, or in other words, their personality trait of sportiness. We have coined two terms to distinguish between these two forms of consistency. *Simple consistency* refers to the stability of behavior due to the situation *and* personality traits. *Residual consistency* refers to the stability of behavior after controlling for effects of personality traits. These two forms of consistency were examined in the article *Distinguishing simple and residual consistency in functionally equivalent situations: Evidence from variable- and person-centered analyses in longitudinal data* (Horstmann et al., in preparation).

This research, the person-situation debate and the ongoing search for consistency, highlights the dire need to include situations in formalizations of personality theories. As no general factor of consistency could be established (Chaplin & Goldberg, 1984; Fleeson & Noftle, 2008; Horstmann et al., in preparation; Schmitt, 1990a), personality theories that were developed after the person-situation debate sought to reconcile the request for stable person parameters with the realization that persons vary, that a person's variability is systematic, and that the interaction of persons *and* situations play and independent role in understanding human behavior. This final approach to resolving the person-situation debate was called "modern interactionism" (Steyer et al., 1999, p. 390). I review the relevant personality theories and frameworks in the next part, section 1.3.

### *1.2.1 Article Summary: Distinguishing Simple and Residual Consistency*

**Background.** The quest for consistency has been a central element of the person-situation debate (Fleeson & Noftle, 2009; Funder, 2001), both at an intra-individual level as well as an inter-individual level. Fleeson and Noftle (2008) examined the concept of consistency and suggested a supermatrix of at least 36 different types of consistency. People behave consistently if they act similarly (for example, show rank-order stability) in different situations. This reflects so called *if…then* contingencies (see below, Mischel & Shoda, 1995): If a person is in situation A, they will express behavior X, and if the same person is again in situation A, they will again express behavior X. However, this definition confounds two effects, namely those of the situation and the person. It is thus not clear why a person acts consistently: It could either be due to their unchanged personality or the unchanged situation.

To distinguish these types of consistency, we introduced the concepts of residual and simple consistency. Simple consistency does not differentiate between consistency due to the person or due to the situation. Residual consistency, however, can be computed by controlling for person-effects in behavioral expression. If two behaviors in two situations have been ridded of person effects and still correlate, then this correlation may be attributed to stable situational influences. Furthermore, both types of consistency can be computed at a between-person level or at a within-person level. If consistency on a person-level could be established, this could be a first step to the search for consistency as the moderator variable that allows predicting some people's behavior better than others (Campus, 1974; Chaplin & Goldberg, 1984; Schmitt, 1990b).

**Study Design.** We aimed to abandon self-reports to assess consistency as they are flawed in many ways (Schmitt, 1990a). Instead, we used a longitudinal design: Participants first reported their trait scores on the Big Five Aspect Scales (DeYoung, Quilty, & Peterson, 2007), that measure the Big Five traits with two aspects each. After two weeks, participants were requested to imagine a hypothetical situation (A) and report how they acted in this situation, again on the BFAS. Two weeks later, participants received the same request, but the situation was slightly altered to avoid memory effects (A'). Situations A and A' were kept as similar as possible.

**Analyses.** We computed simple consistency by correlating the Big Five Aspect scores at A and A'. To examine simple consistency at person level, we computed profile correlations between the item-profiles at A and A', following a recommendation by Asendorpf (1990). To compute residual consistency, we regressed the scores at A and A' on the BFAS trait scores and correlated the regression residuals with each other. For the person-level analyses, we subtracted the item scores at trait-assessment from the item scores at A and A' and correlated the residual profiles with each other.

**Results.** Both simple and residual consistency could be established. Both forms of consistency occurred at within- as well as between-person level. Forms of simple consistency were higher than forms of residual consistency, reflecting trait-expression. Yet, residual consistency was substantial. As individual differences in within-person consistencies were only weakly correlated, no g-factor of consistency could be established.

**Conclusion.** Our findings support the assumption that both personality traits and situational factors influence behavior in a given situation. Without controlling for trait-influences, estimates of consistency will confound trait- and situation-influences. Situations thus play indeed a very important role in understanding and examining consistency (Mischel, Shoda, & Mendoza-Denton, 2002). However, the current study did not investigate why the two situations were functionally equivalent, that is, which characteristics made them functionally equivalent. Finally, no general factor of consistency could be established, thus supporting the idea that persons might act consistently in one domain and less so in another domain.

### *1.2.2 Article: Distinguishing Simple and Residual Consistency*

**Distinguishing simple and residual consistency in functionally equivalent situations: Evidence from variable- and person-centered analyses in longitudinal data**

Kai T. Horstmann[1], John F. Rauthmann[2], & Matthias Ziegler[1]

[1]Humboldt-Universität zu Berlin (Germany)

[2]Wake Forest University (USA)

**Author Note**

We thank Katja Witte for her support with data collection and study design. Parts of these data were reported in her Bachelor Thesis. This study was pre-registered on the Open Science Framework (see osf.io/8u7ka). Additional materials that allow a complete reproduction of the analyses and a replication of the study can be found at osf.io/xfhdu. Both links are anonymous view-only links and will be replaced with permanent links in the final version of the manuscript.

**Abstract**

The current study examines variable- and person-oriented consistencies of personality states across functionally equivalent situations. We argue that that a *simple consistency* needs to be distinguished from a *residual consistency*. The former correlates simple state scores not taking people's traits into account, while the latter correlates residual state scores that have been corrected for trait scores. Residual consistency means that state residuals are systematic. We examine the level and individual differences in all of these forms of consistency. In a pre-registered longitudinal study, 99 participants provided first trait-ratings and then 3 weeks later two times state-ratings in response to two functionally equivalent situation vignettes (each being 3 weeks apart). For traits and states, we chose the Big Five Aspect Scales (BFAS). Variable- and person-oriented analyses yielded that both simple and residual consistencies were substantial, the latter being only slightly smaller than the former. Further, individual differences in variable-oriented consistencies were only weakly correlated, suggesting no underlying g-factor but aspect-specific consistencies. Participants also varied in their person-oriented profile consistencies. Residual consistency was substantial, and state residuals ridded from trait variance are reliable and not just noise in the case of functionally equivalent situations.

## Distinguishing simple and residual consistency in functionally equivalent situations: Evidence from variable- and person-centered analyses in longitudinal data

*You cannot step twice into the same river.*

– Attributed to Heraclitus

This quote describes the notion that no person can ever be in the same situation more than once as both the situation as well as the person will have changed. Even if it may not be possible to be in the exact same situation twice, it is still possible to be in situations that are at least *functionally equivalent*, that is, they might elicit the same or highly similar thoughts, feelings, desires, or behaviors. Thus, the river may have changed in some ways, such as the amount of water flowing or the precise position of pebbles in the river-bed, but its essential quality – cold water – may not have changed. A person stepping in this functionally equivalent river may therefore show consistent behavior such as shrieking each time they touch the water. Such shrieking may be traced back in parts to the situation consistently affording shrieking and the person having a tendency to shriek. However, it could even be possible that the relation between functional equivalence of a situation and behavior might, at least in part, be independent of the trait (e.g., tendency to shriek) under focus. If this were the case, and such consistency of behaviors independent from a corresponding trait existed, it would alert us to systematic portions of variance in states that cannot be attributed to trait levels alone. However, previous studies have not quantified such a form of consistency. Nonetheless, it is important to demonstrate such consistency because people regularly inhabit functionally equivalent situations in their daily lives that repeat themselves and provide for routine (e.g., playing with one's children, commuting, working with colleagues, meeting friends for a coffee), yet perhaps sometimes even happening without strong correspondence to their personalities (Ickes, Snyder, & Garcia, 1997). This study therefore examines variable- and person-oriented forms of state consistencies across different, yet functionally equivalent, situations. To this end, we will introduce the novel concept of "residual consistency" alongside the traditional "simple consistency" which has already been used in the literature.

### Background

#### Traits and States

A personality *trait* is usually conceptualized as a stable characteristic of a person (Funder, 2001), and trait levels may differ between people. For example, some people can be described as *generally* more extraverted, whereas others can better be described as *generally* more introverted. Such general descriptions (across many time-points) notwithstanding, people regularly exhibit a range of personality *states*. The latter have been defined as "having the same affective, behavioral,

and cognitive content as a corresponding trait (…), but as applying for a shorter duration" (Fleeson & Jayawickreme, 2015, p. 84). Notably, states can form distributions within persons, and the parameters of these distributions (especially the local tendencies) are often substantially associated with self-reports of traits (Fleeson & Jayawickreme, 2015; Horstmann & Rauthmann, in preparation). Such parameters have additionally been shown to be stable, meaningful, and consequential (Fleeson, 2001; Jones, Brown, Serfass, & Sherman, 2017). A conceptualization of personality – with stable components (traits) and variable components (states) – reconciles structure- and process-oriented approaches (Baumert et al., 2017; Fleeson, 2001).

Nonetheless, the situations in which states occur have often been neglected (but see Fleeson, 2007; Sherman, Rauthmann, Brown, Serfass, & Jones, 2015). For example, in latent state-trait theory (Steyer, Mayer, Geiser, & Cole, 2015; Steyer et al., 1999), the latent state variable is a placeholder, but its characteristics remain largely unexplained. Conceptually, a state variable may be composed of systematic individual differences (= traits) as well as systematic situation-specific influences (= state residuals), whatever they may be. It remains to be empirically shown to what extent *state residuals* that have been ridded from trait variance are consistent across occasions. In other words, if the state residual in one situation differs from the state residual in another functionally equivalent situation, then these state residuals would not be consistent. Here, we argue that state residuals should be consistent *if* situations are indeed functionally equivalent.

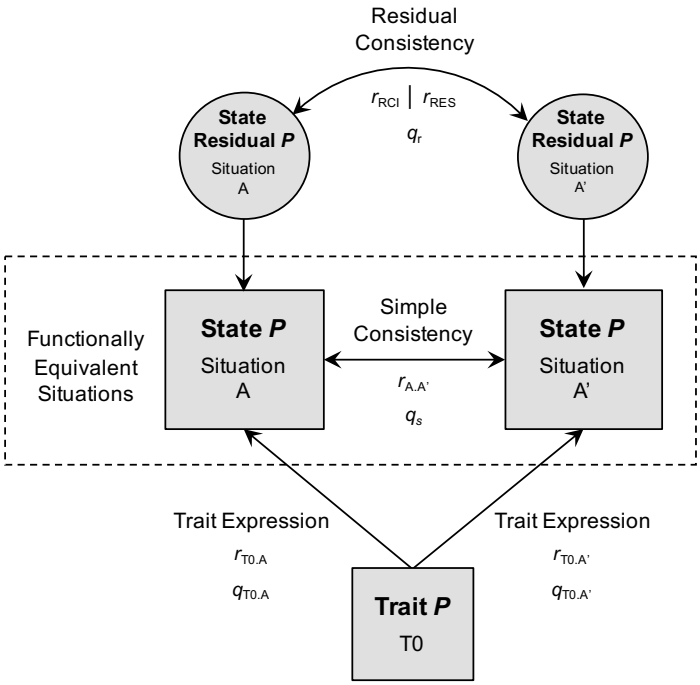
**Consistency**

**Forms of consistency.** Fleeson and Noftle (Fleeson & Noftle, 2008) have unpacked the consistency concept and came to the conclusion that "there is no one way to define consistency" (p. 1356). They proposed a supermatrix with 36 different types of consistency, depending on (a) the definition of computed similarity (absolute, rank-order, ipsative), (b) the competing determinant across which consistency is to be established (behavior content, situation content, time intervals), and (c) the definition of behavioral enactment (single, aggregated/averaged, contingent, patterned). In terms of similarity, we could be interested in the preservation of rank-orderings between people in one variable (*variable-oriented* approach) or the ipsative rank-orderings among different variables within one person (*person-oriented* approach). In terms of the competing determinant, we are interested in *lapses of time* as situation content would not vary when examining functionally equivalent situations. In terms of the definition of enactment, we are interested in *contingent enactment* (Mischel et al., 2002; Shoda & Mischel, 2000) – essentially an *if...then* pattern of a state (*then*) exhibited in a given situation (*if*).

To use the river example, we would be interested in the *if…then* pattern of "*If* in cold water, *then* shrieking." Here, the state "shrieking" is to some degree composed of at least a trait (general

tendency to shriek) and a state residual. Consequently, we can define consistency in two ways, depending on the state score used. If raw state scores (e.g., the average of several items assessing state shrieking) are correlated with each other, then *simple consistency* is computed. If state residual scores are correlated (i.e., state scores that have been somehow controlled for a trait score), then *residual consistency* is computed. The literature on consistency has so far only quantified simple forms only (Fleeson & Noftle, 2008), but not residual forms. However, a high degree of residual consistency suggests that state residuals are systematic and potentially important towards understanding persons and situations better.

*Figure 1.2.1.* Visualization of Residual State-Consistency




*Note.* One domain *P* for traits and states is given as an illustration. A double-headed arrow reflects a correlation. All *r*s are variable oriented, estimated for each aspect separately (see Table 1.2.2); all *q*s are person-oriented, estimated across the entire profile of items for each person separately (see Table 1.2.5).

Figure 1.2.1 illustrates our explanations above, depicting (a) a single trait or trait profile (Trait *P*), (b) two regular states or profiles of states (State *P*) measured at two time-points with functionally equivalent situations (here: Situations A and A'), and (c) two state residuals or profiles of state residuals (State Residual *P*). First, traits predict states because they manifest or can be expressed in them. This *trait expression* corresponds to a correlation between a trait variable and a state variable across individuals (variable-oriented) or a correlation between a trait profile and a

corresponding state profile within a single individual (person-oriented). If there are two time-points with functionally equivalent situations (as in Figure 1.2.1), then trait expression can be estimated for both time points. Notably, if the situations are indeed functionally equivalent, then trait expressions should be highly similar. Second, *simple consistency* would refer to simply correlating two states. The variable-oriented form of simple consistency computes correlations *r* across individuals for one variable (e.g., Neuroticism), while the person-oriented form computes correlations *q* across variables within individuals (e.g., a Big Five profile). Third, state residuals – as states that are controlled for traits – can be correlated across time-points to obtain estimates of residual consistency. As with simple consistency, this can be done in a variable-oriented and a person-oriented fashion.

**Individual differences in consistency.** For the different forms of consistency outlined, we can ask how strong individual differences in them are. For variable-oriented consistency forms, we would want to know whether being consistent in one variable (e.g., Neuroticism) also entails being consistent in another one (e.g., Extraversion). Indeed, if people were simultaneously consistent in several variables, then this would point towards a general factor of consistency. Previous research has, however, provided mixed evidence for such general consistency (Bem & Allen, 1974; Chaplin & Goldberg, 1984; Schmitt, 1990a, 1990b), and it is thus an open question if general consistency exists across different traits like the Big Five domains and aspects (DeYoung et al., 2007). Measures of individual, variable oriented consistency can be obtained using a score proposed by Asendorpf (1990). To obtain measures of individual differences in person-oriented consistency, each person in a sample obtains a *q*-correlation, and differences in *q*-correlation estimates directly represent individual differences in ipsative consistency. The current study will thus seek to quantify individual differences (and their interrelations) for variable- and person-oriented consistencies.

## The Current Study

The general research question, methods, and parts of the data-analytical strategy of this study were pre-registered on the Open Science Framework (OSF)[2]. A synopsis can be found in Table 1.2.1 where pre-registered elements are contrasted with those that were not. Additionally, Table 1.2.1 lists all research questions we addressed in this work, along with the tables and figures that correspond to them.

The primary purpose of this study is the examination of different types of consistency (Figure 1.2.1), specifically illuminating the as of yet understudied residual consistencies. Using hypothetical but functionally equivalent situational vignettes, we examine to what extent (a) participants' states (simple consistency) as well as (b) their state residuals as deviations from trait scores (residual consistency) are correlated between two functionally equivalent Situations A and A'. As can be seen in Table 1.2.1, we report findings for both a variable- and person-oriented approach to consistencies.

[2] osf.io/8u7ka

Table 1.2.1
*Overview of Study Questions, Tables, and Figures*

| Point | Pre-registered? | Table | Figure |
|---|---|---|---|
| *Study Modalities* | | | |
| General Research Question | Yes | | 1 |
| Methods | | | |
| Sample size | Yes[a] | | |
| Procedures | Yes | | 2 |
| Materials | Yes | | 2 |
| Data pre-processing[b] | Yes | | |
| *Variable-oriented Analyses (per Aspect)* | | | |
| (No) Mean-level change in states between Situation A and A' | Yes | 2 | 3 |
| Trait expressions ($r_{T0.A}$ and $r_{T0.A'}$) | No | 2 | |
| Simple consistency ($r_{A.A'}$) | No | 2 | |
| Correlation of state scores across time and aspects | No | 3 | |
| Quantification and intercorrelation of individual differences in $r_{A.A'}$ | No | 4 | |
| Residual consistency ($r_{RCI}$, $r_{RES}$) | | | |
| Use as residual state scores: reliable change indices (RCI) | Yes | 2 | |
| Use as residual state scores: regression residuals (RES)[c] | No | 2 | |
| Correlation of residual state scores (RCI, RES) across time and aspects | No | 3 | 4[d] |
| Quantification and intercorrelation of individual differences in $r_{RCI}$ and $r_{RES}$ | No | 4 | |
| *Person-oriented Analyses* | | | |
| Trait expressions ($q_{T0.A}$ and $q_{T0.A'}$) | No | 5 | |
| Simple consistency ($q_s$) | No | 5 | |
| Quantification of individual differences in $q_s$ | No | 5 | 5 |
| Residual consistency ($q_r$) | No | 5 | |
| Quantification of individual differences in $q_r$ | No | 5 | 5 |

*Note.* [a] We pre-registered $N = 250$ (based on recommendations for cross-sectional studies from (Schönbrodt & Perugini, 2013), but were only able to gather full usable data from $N = 99$. [b] Recoding, scale score computations, handling of missing data, etc. [c] Done to estimate the robustness of findings across different ways of computing state residuals. [d] For simplicity sake, this is only done aspect-specific.

Within a *variable-oriented approach* that examines findings at the sample-level for each Big Five aspect separately (DeYoung et al., 2007), we report several findings. First, we examined mean-level change in states between Situations A to A'. As pre-registered, we expected no significant changes here because the two situations were supposed to be functionally equivalent. Second, we report trait expressions which are correlations of a trait assessed at an initial measurement occasion with a corresponding state at later situations Situation A and A'. If the situations are functionally equivalent, then trait expression correlations ($r_{T0.A}$ and $r_{T0.A'}$) should be similar. Third, we report estimates of simple consistency ($r_{A.A'}$) as the correlation between corresponding states at Situation A and A'. For example, state Intellect at Situation A should be substantially correlated with state Intellect at Situation A'. The evidence for such simple consistency will be strongest when convergent correlations (e.g., Intellect at A and A') exceed the discriminant correlations in the off-diagonal (e.g., Intellect at A with Volatility at A and A'). Additionally, we examine the correlations of inter-individual differences in these simple consistencies. Uniformly high positive (and substantial) correlations would point towards a manifold and thus to a potential g-factor of consistency underlying the covariation of individual simple consistency scores. Low correlations would suggest otherwise. Fourth, we report findings for residual consistency using the same data-analytical steps as for simple consistency (i.e., quantification of residual consistency, intercorrelations of state residual scores, and intercorrelations of individual differences in residual consistencies). Doing so enables comparing simple and residual consistencies head to head, with the expectation that residual consistencies will be somewhat lower than simple ones yet still substantial. Further, we make use of two different ways of computing state residual scores (reliable change indices vs. regression-based residuals; see *Data Analyses*) to provide estimates of the robustness of our findings (i.e., our pattern of findings should hold regardless of how state residuals are computed). Notably, the use of reliable change indices was pre-registered, while the regression-based estimation of residuals (by predicting states from traits and using the resultant residual scores) was not.

Within a *person-oriented approach* that examines findings at the individual-level across all Big Five aspect items, we report estimates of *q*-correlations that are conceptually analogous to variable-oriented *r* correlations. Specifically, first, we report trait expressions as within-person correlations of a trait profile at T0 with a state profile at Situation A and A' ($q_{T0.A}$ and $q_{T0.A'}$). Second, we report simple consistency ($q_s$) as the within-person correlations between state profiles at Situation A and A'. The standard deviation of the resultant individual *q*-correlation scores indexes individual differences in person-oriented simple consistency. Third, as with simple consistency, we report residual consistency ($q_r$) as the within-person correlations between residual state profiles at

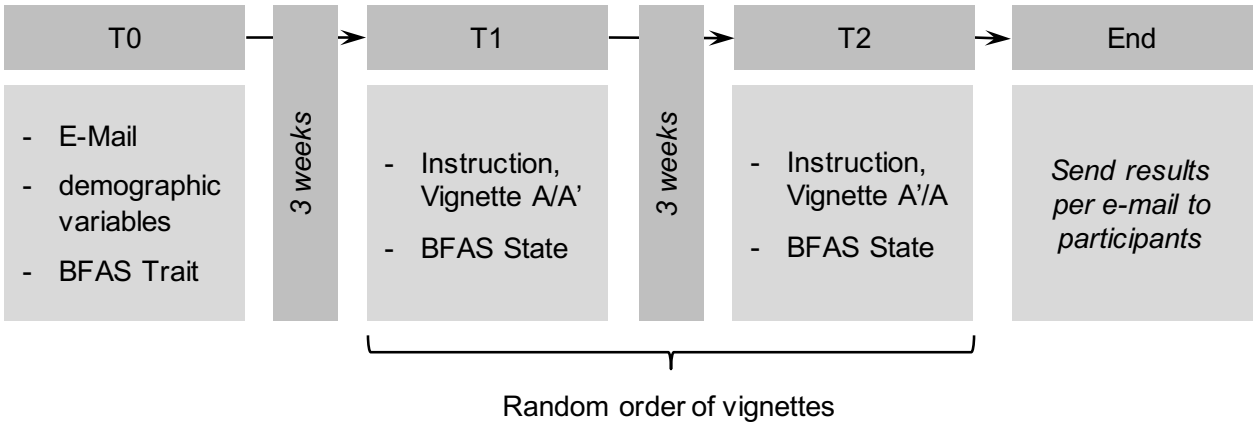
Situation A and A' (i.e., individual state profiles centered on individual trait profiles). Again, the standard deviation of resultant *q*-correlations indexes individual differences. Together, this program of analyses (Table 1.2.1) estimates all coefficients of Figure 1.2.1 and casts a differentiated perspective on the level and individual differences of variable-oriented and person-oriented forms of simple and residual state consistency.

## Methods

### Participants

A total of 101 participants completed all stages of the study (see Figure 1.2.2 for details). However, two participants had no variance across all items and were thus removed[3]. Thus, the final sample contains $N = 99$ participants (80.81% female; age: $M = 26.01$, $SD = 7.33$, range = 17-53 years). No other socio-demographic characteristics of the participants were assessed. Participants received detailed feedback on their personality traits. If participants were psychology undergraduate students, they could also receive course credit.

*Figure 1.2.2.* Process of Data Collection




*Note.* BFAS = Big Five Aspect Scale (DeYoung et al., 2007). "Trait" indicates that the scale was administered with the instruction "How are you in general"; "State" indicates that participants were asked how they behaved momentarily in the specific situation. Vignette A and A' refer to the hypothetical situations that were presented to the participants. Although we refer to T1 and T2 here, the two measurement occasions were balanced across participants (i.e., some obtained A first and then A', others first A' and then A).

[3] Removing these participants was not pre-registered, and we ran all analyses with the participants as well which did not change any results.

**Procedure**

The full study process is depicted in Figure 1.2.2. First, participants were informed about the study, gave their consent, and were registered with their e-mail address if they wanted to participate. 931 participants clicked on the study platform, of these, 178 participants continued with the study. Second at Stage T0, participants indicated their age and sex and responded to items measuring trait-version of the Big Five Aspect Scales (BFAS; DeYoung et al., 2007). After three weeks, participants were invited via e-mail to participate again and randomly assigned to either of the hypothetical Situation Vignettes A or A' (see *Materials*).[4] Participants had to read the situation and rate their hypothetical behavior in this situation on a state-version of the BFAS. After another three weeks, participants were invited again and rated their hypothetical behavior in the *other* situation (either A or A'). 126 participants completed state measures for Situation A, 123 participants for Situation A', and 99 participants provided usable data for all three measurement points.

**Materials**

All measures were administered online using the platform formr.org (Arslan & Tata, 2015). Participants were not allowed to skip items (hence, there were no missing data), but could abort at any point. All materials necessary for a direct replication of this study (situation vignettes, translated items, questionnaire, study set-up, data and code) can be found at the OSF page: osf.io/xfhdu.

**Situation Vignettes.** The hypothetical situation we chose was a party situation because most participants in our intended sample could be expected to have been in one (high familiarity, close to real life). Further, a party situation is relevant to the expression of at least two important interpersonal domains (high content specificity)[5]: agreeableness and extraversion (Gurtman, 2009). The participants were instructed to imagine a situation where they had been taken along to a private party. They only know one person there and have to start a conversation with other people. Vignette A is 261 words long, and Vignette A' 244 words. The vignettes can be found at the OSF page. They were constructed to be similar, though somewhat different in wording (to avoid strong memory effects after three weeks).

**Traits and States.** Traits at T0 and states in Situations A and A' were assessed with 100 items from the Big Five Aspect Scale (BFAS) by DeYoung and colleagues (2007), using a 6-point Likert-type response scale (1 = *disagree*, 6 = *agree*). Each Big Five domain is measured with two aspects and ten items each: Openness with Openness and Intellect, Conscientiousness with Industriousness and Orderliness, Extraversion with Assertiveness and Enthusiasm, Agreeableness

[4] As the order of the vignettes was randomly assigned, we do not refer to the measurement points as T1 and T2, respectively. We simply refer to A and A'.

[5] It is difficult creating a vignette with a real-world, familiar situation that activates all Big Five domains at once.

with Compassion and Politeness, and Neuroticism with Volatility and Withdrawal. As no German version of the BFAS was available, we translated the items with multiple translators (see OSF page). We chose the BFAS since it allows a more detailed look at personality than pure Big Five measures, while being reasonably short at the same time. We report variable-oriented findings at the level of aspects, not domains.

**Data Analyses**

All analyses as well as all anonymized data can be found at the OSF page. We used the software R with RStudio (R Core Team, 2016; RStudio Team, 2015) and the following packages: *dplyr* (Wickham & Francois, 2016), *stringr* (Wickham, 2016b), *purrr* (Wickham, 2016a), *psych* (Revelle, 2016), and *car* (Fox & Weisberg, 2011).

All coefficients in Figure 1.2.1 and questions in Table 1.2.1 were addressed (see *The Current Study*). Most data-analytical issues involved are straightforward and will be referenced in *Results*. However, two special issues for variable-oriented analyses are addressed here: the derivation of (a) residual state scores and (b) inter-individual differences in consistencies.

**Deriving Residual State Scores.** We used two types of residual state scores, and thus we could estimate residual consistencies twice (see $r_{RCI}$ and $r_{RES}$ in Figure 1.2.1 and Table 1.2.1). First, as pre-registered, we computed for each person their *reliable change index* (RCI; Jacobson & Truax, 1991) for each BFAS aspect from T0 to A and T0 to A'. To compute the RCIs (Equations 1 and 4), we used the retest reliabilities provided by DeYoung et al. (2007) for undergraduate students as we assumed that most of our participants would also be students. There are different ways to compute RCIs based on the standard error (*SE*) of the measurement (Maassen, 2004). We assumed a priori that the *SE* at T0 will be the best estimate (computed across all 176 participants who completed this stage, see Equation 3) and therefore relied on this *SE* for each aspect. The RCI for the difference between two measurement occasions is defined as:

$$RCI = \frac{x_{T0} - x_{A/A'}}{S_{diff}} \quad (1)$$

where $x_{T0}$ = trait score on T0, $x_{A/A'}$ = state score at Situation A or A', and $S_{diff}$ = standard error of the difference. $S_{diff}$ is defined as follows:

$$S_{diff} = \sqrt{2(S_E)^2} \quad (2)$$

where $S_E$ = standard error of the difference score.

Based on the assumption that the variances at both measurement occasions are the same and both aspect scales used are parallel, $S_E$ is defined as follows:

$$S_E = s_{T0}\sqrt{(1 - r_{tt})} \quad (3)$$

where $S_{T0}$ = standard deviation at the initial trait measurement T0, $r_{tt}$ = re-test reliability (as reported by DeYoung et al., 2007, p. 889, Table 5, $r^{\prime}$).

Plugging Equation (3) in (2) and the result into (1) will give the final RCI equation:

$$RCI = \frac{x_{T0} - x_{A/A\prime}}{\sqrt{2\left(s_{T0}\sqrt{1 - r_{tt}}\right)^2}} \tag{4}$$

Second, we used an additional common but not pre-registered procedure to derive residual state scores. Specifically, we regressed a BFAS state at Situation A or A', respectively, on the corresponding BFAS trait at T0. We then extracted the regression residuals (RES) which are state residuals.

**Simulations and validity of findings.** As pre-registered, we planned to use latent change score models (LCM) in addition to RCI to examine the correlation of residuals. Because our sample size was too small for sensible structural equation model analyses, we opted not to use LCMs. However, to show that LCMs would arrive at the same conclusions as manifest models using RCI or RES scores, we compiled simulations that can be accessed at OSF[6]. As can be seen there, LCMs and RCI-based models yield the exact same estimation for the correlation under the assumption that the latent variables are a good representation of the manifest variables.

In addition to this issue, we also address a second one that became apparent to us after pre-registration: Correlations based on difference scores may be over- or under-estimated, depending on the correlation of the variables used to compute the difference score (Vickers & Altman, 2001). However, as we show in the simulation, RES scores do not suffer from this limitation. Further, RCI and RES scores yield similar estimates for residual consistency *only under certain circumstances*. If they do not replicate, then findings may not be valid. However, in the case of replication, we have strong evidence for actual residual consistency.

**Deriving inter-individual differences in consistencies.** Asendorpf (1990) provided an index (*I*) for the computation of individual consistency scores across two measurements where their mean is virtually identical to the sample-level observed rank-order consistency:

$$I_{A.A\prime} = 1 - \frac{(z_A - z_{A\prime})^2}{2} \tag{5}$$

where $I_{A.A'}$ = individual consistency score for one person regarding one BFAS state, $z_A$ and $z_{A'}$ = $z$-scores on the state at Situation A and A, respectively.

To use $I_{A.A'}$ scores (e.g., to correlate them), they needed to be transformed due to their heavy skewness (for the exact procedure, see Asendorpf, 1990, p. 9), which is nearly identical to

[6] We include the simulation study as a supplement due to the limited space available. It can be accessed at osf.io/xfhdu.

the fisher *r*-to-*z* transformation, but is also defined for $r = 1$. The resultant transformed scores $TI_{A.A'}$ were approximately normally distributed and thus could be used for further analyses. $TI_{A.A'}$ scores were derived for simple ($r_{A.A'}$) and residual consistencies ($r_{RCI}$ and $r_{RES}$). Intercorrelations of $TI_{A.B}$ scores allow to examine the structure of individual differences in consistencies. For example, a highly correlated structure points towards a possible g-factor of consistency.

**Deviations from the Pre-registration**

Although we pre-registered to sample $N = 250$ participants, only 101 participants completed the final study. It is therefore not possible to run latent change score models (McArdle, 2009). However, as mentioned above, we found that change score models and change scores based on reliable change indices yield the exact same parameters. Table 1.2.1 alerts to elements that were pre-registered and those that we added after the pre-registration.

## Results

Descriptive statistics and reliabilities for all ten aspects assessed at the three measurement occasions (trait at T0, state at Situation A and A') are displayed in Table 1.2.2. The reliabilities (Cronbach's $\alpha$) for all scales were similar to the ones presented by DeYoung and colleagues (2007, Table 1.2.5).

Table 1.2.2
*Descriptive Statistics and Correlations*

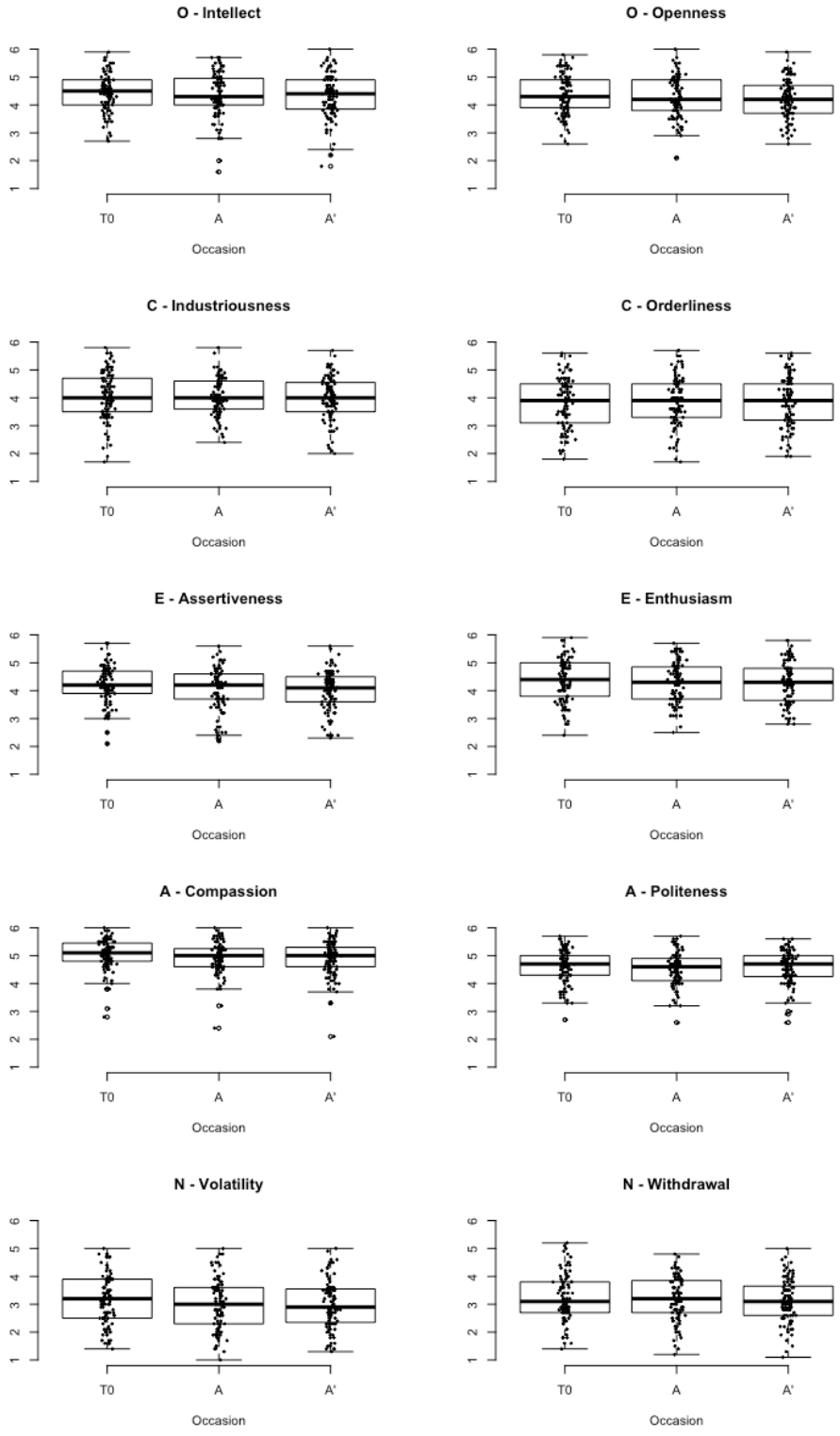
| Domains and Aspects | Trait at T0 | | | | State at Situation A | | | | State at Situation A' | | | | $\Delta M_{A\text{-}A'}$ | | | | Trait Expressions | | Consistencies | | | Convergence |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | α | ω | *M* | *SD* | α | ω | *M* | *SD* | α | ω | *t* | *d* | *p* | 95% CI | $r_{T0.A}$ | $r_{T0.A'}$ | $r_{A.A'}$ | $r_{RCI}$ | $r_{RES}$ | $R_{RCI/RES}$ |
| *Openness* | | | | | | | | | | | | | | | | | | | | | | |
| Intellect | 4.43 | 0.67 | .79 | .84 | 4.37 | 0.80 | .87 | .91 | 4.32 | 0.80 | .87 | .90 | 1.33 | 0.06 | .186 | [-0.02; 0.12] | .87 | .86 | .89 | .58 | .58 | .997 |
| Openness | 4.35 | 0.72 | .75 | .82 | 4.26 | 0.73 | .77 | .85 | 4.21 | 0.73 | .76 | .82 | 1.32 | 0.08 | .190 | [-0.03; 0.14] | .78 | .81 | .83 | .59 | .54 | .997 |
| *Conscientiousness* | | | | | | | | | | | | | | | | | | | | | | |
| Industriousness | 4.06 | 0.84 | .85 | .88 | 4.03 | 0.68 | .74 | .76 | 3.99 | 0.77 | .82 | .86 | 1.04 | 0.06 | .299 | [-0.04; 0.13] | .79 | .84 | .84 | .65 | .55 | .839 |
| Orderliness | 3.82 | 0.90 | .86 | .90 | 3.88 | 0.89 | .85 | .90 | 3.83 | 0.89 | .85 | .90 | 1.54 | 0.07 | .128 | [-0.02; 0.13] | .85 | .83 | .91 | .72 | .69 | .998 |
| *Extraversion* | | | | | | | | | | | | | | | | | | | | | | |
| Assertiveness | 4.22 | 0.67 | .80 | .86 | 4.05 | 0.73 | .85 | .88 | 4.01 | 0.70 | .85 | .88 | 1.16 | 0.07 | .248 | [-0.03; 0.13] | .77 | .80 | .83 | .60 | .57 | .991 |
| Enthusiasm | 4.39 | 0.78 | .86 | .90 | 4.29 | 0.73 | .85 | .89 | 4.28 | 0.74 | .86 | .90 | 0.33 | 0.02 | .741 | [-0.06; 0.08] | .80 | .79 | .88 | .72 | .67 | .997 |
| *Agreeableness* | | | | | | | | | | | | | | | | | | | | | | |
| Compassion | 5.03 | 0.55 | .82 | .87 | 4.92 | 0.56 | .84 | .87 | 4.89 | 0.62 | .86 | .89 | 0.77 | 0.05 | .440 | [-0.04; 0.10] | .68 | .71 | .83 | .70 | .67 | .892 |
| Politeness | 4.60 | 0.60 | .72 | .81 | 4.52 | 0.59 | .71 | .79 | 4.58 | 0.61 | .76 | .82 | -1.8 | -0.11 | .075 | [-0.14; 0.01] | .76 | .71 | .82 | .66 | .60 | 1.000 |
| *Neuroticism* | | | | | | | | | | | | | | | | | | | | | | |
| Volatility | 3.15 | 0.90 | .88 | .92 | 2.98 | 0.90 | .89 | .92 | 2.99 | 0.88 | .90 | .93 | -0.31 | -0.02 | .761 | [-0.12; 0.09] | .85 | .78 | .83 | .56 | .51 | .959 |
| Withdrawal | 3.20 | 0.83 | .84 | .88 | 3.19 | 0.78 | .82 | .85 | 3.12 | 0.77 | .82 | .87 | 1.57 | 0.09 | .120 | [-0.02; 0.16] | .84 | .80 | .83 | .58 | .50 | .986 |

*Note.* $N = 99$. α = internal consistency (Cronbach's alpha), ω = factor reliability (McDonald's omega). $\Delta M_{A\text{-}A'}$ = Mean difference of an aspect scale between Situations A and B. $r_{T0.A}$ = bivariate correlation of an aspect at T0 with state at Situation A, $r_{T0.A'}$ = bivariate correlation of an aspect at T0 with state at Situation B. $r_{TA.A'}$ = simple consistency, as the bivariate correlation of a state at Situation A with itself at Situation A'. $r_{RCI}$ = residual consistency, as the correlation of a state residual (here a reliable change index score) at Situation A with itself at Situation A'. $r_{RES}$ = residual consistency, as the correlation of a state residual (here a regression residual score where a state was predicting from the corresponding trait) at Situation A with itself at Situation A'. $r_{RCI/RES}$ = bivariate correlations of the individual consistency scores based on reliable change index and regression residual scores.

**Variable-Oriented Analyses**

**Estimation of Mean Level Change.** We examined (as pre-registered) to what extent state means differed between Situations A and A'. The results are displayed in Table 1.2.3 under "$\Delta M_{A\text{-}A'}$" and in Figure 1.2.3. As can be seen, there were no significant mean-level differences (dependent-sampled *t*s ranged from -1.88 to 1.57, *d*s = -0.11-0.09, all *p*s > .075). We interpret this pattern of findings as support for the functional equivalence of the two situation vignettes.

*Figure 1.2.3.* Box-plots of Individual Scores of Trait and State Aspects




*Note. N* = 99. O = Openness, C = Conscientiousness, E= Extraversion, A = Agreeableness, N = Neuroticism.T0 = Trait measurement; A = state measurement in Situation Vignette A; A' = state measurement in Situation Vignette A'.

**Trait Expressions.** Correlations between the trait scores at T0 and the corresponding state scores at Situations A and A' ($r_{T0.A}$, $r_{T0.A'}$) are presented in Table 1.2.2 under "Trait Expressions." As can be seen, trait expressions were uniformly high, ranging from .71 (Compassion and Politeness) to .87 (Intellect). Additionally, they were similar across Situations A and A', again speaking for functional equivalence of the vignettes.

**Simple Consistencies.** Correlations between the state scores at Situation A and A' ($r_{A.A'}$) are presented in Table 1.2.2 under the first column of "Consistencies" (see first column there). As can be seen, all simple consistencies were substantial, ranging from .82 (Politeness) to .91 (Orderliness). Thus, participants showed strong levels of simple rank-order consistency, which is in line with the functional equivalence of situation vignettes (i.e., rank-orders among participants are preserved in same situations).

Next, we examined the full intercorrelations of state scores for Situations A and A', as presented in the upper half of Table 1.2.3 under "For State Scores." The correlation matrix allows drawing conclusions based on convergent and discriminant associations. First, the gray-shaded cells capture convergent associations, which have already been detailed as $r_{A.A'}$ in Table 1.2.2. These correlations were on average .85 (*SD* = .13), while the off-diagonal discriminant correlations (i.e., different aspects correlating at the same or different situations) amounted to an average *absolute* correlation of only .23 (*SD* = .20). Thus, we can conclude that each state aspect was uniquely associated with itself when measured twice, suggesting convergent and discriminant validity of state aspects.

Lastly, we analyzed how inter-individual differences in simple consistencies of each aspect were related to each other. These findings are presented in the upper half of Table 1.2.4 under "For State Scores." As can be seen, correlations were in general rather small (and failed to reach conventional levels for statistical significance in all but one case), and the average absolute intercorrelation amounted to .09 (*SD* = .09), with the highest correlation being .34 (between Volatility and Assertiveness). This lack of covariation does not speak for g-factor of simple consistency.

Table 1.2.3

*Full Intercorrelation Matrix of State and Residual State Scores (Variable-Oriented)*

| Scales | O: $I_A$ | O: $I_{A'}$ | O: $O_A$ | O: $O_{A'}$ | C: $I_A$ | C: $I_{A'}$ | C: $O_A$ | C: $O_{A'}$ | E: $A_A$ | E: $A_{A'}$ | E: $E_A$ | E: $E_{A'}$ | A: $C_A$ | A: $C_{A'}$ | A: $P_A$ | A: $P_{A'}$ | N: $V_A$ | N: $V_{A'}$ | N: $W_A$ | N: $W_{A'}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *For State Scores* | | | | | | | | | | | | | | | | | | | | |
| O: $I_A$ | - | | | | | | | | | | | | | | | | | | | |
| O: $I_{A'}$ | **.89*** ** | - | | | | | | | | | | | | | | | | | | |
| O: $O_A$ | .21 | .20 | - | | | | | | | | | | | | | | | | | |
| O: $O_{A'}$ | .26 | .23 | **.83*** ** | - | | | | | | | | | | | | | | | | |
| C: $I_A$ | .28 | .34 | .04 | .01 | - | | | | | | | | | | | | | | | |
| C: $I_{A'}$ | .28 | .42** | .05 | -.04 | **.84*** ** | - | | | | | | | | | | | | | | |
| C: $O_A$ | -.17 | -.15 | .01 | .00 | .46*** | .35 | - | | | | | | | | | | | | | |
| C: $O_{A'}$ | -.11 | -.08 | -.04 | -.06 | .45*** | .37* | **.91*** ** | - | | | | | | | | | | | | |
| E: $A_A$ | .40** | .46*** | .22 | .18 | .40** | .42** | .03 | .07 | - | | | | | | | | | | | |
| E: $A_{A'}$ | .36* | .47*** | .13 | .21 | .44*** | .45*** | .03 | .05 | **.83*** ** | - | | | | | | | | | | |
| E: $E_A$ | .15 | .15 | .16 | .18 | .06 | .18 | -.02 | -.04 | .21 | .16 | - | | | | | | | | | |
| E: $E_{A'}$ | .05 | .11 | .09 | .15 | .12 | .21 | .05 | .00 | .14 | .23 | **.88*** ** | - | | | | | | | | |
| A: $C_A$ | .18 | .20 | .39** | .37* | .15 | .07 | .04 | .03 | .25 | .22 | .33 | .29 | - | | | | | | | |
| A: $C_{A'}$ | .05 | .10 | .34 | .35 | .18 | .11 | .10 | .05 | .20 | .30 | .26 | .34 | **.83*** ** | - | | | | | | |
| A: $P_A$ | -.22 | -.13 | .03 | -.02 | .38* | .38* | .33 | .33 | -.20 | -.11 | .02 | .08 | .13 | .22 | - | | | | | |
| A: $P_{A'}$ | -.23 | -.20 | .10 | .05 | .32 | .28 | .38* | .34 | -.22 | -.22 | -.01 | .02 | .22 | .29 | **.82** ** | - | | | | |
| N: $V_A$ | -.34 | -.40** | -.03 | .02 | -.55*** | -.55*** | -.02 | -.05 | -.29 | -.31 | -.19 | -.16 | .01 | -.02 | -.40** | -.20 | - | | | |
| N: $V_{A'}$ | -.36* | -.39** | -.07 | .07 | -.50*** | -.55*** | .01 | .02 | -.31 | -.23 | -.16 | -.11 | -.03 | -.07 | -.34 | -.25 | **.83*** ** | - | | |
| N: $W_A$ | -.34 | -.41** | .03 | .04 | -.52*** | -.53*** | .10 | .05 | -.46*** | -.46*** | -.25 | -.24 | -.03 | -.03 | -.21 | -.07 | .81*** | .71*** | - | |
| N: $W_{A'}$ | -.29 | -.40** | .03 | .11 | -.58*** | -.63*** | .00 | -.02 | -.40** | -.38* | -.22 | -.23 | -.03 | -.08 | -.31 | -.22 | .76*** | .80*** | **.83*** ** | - |
| *For Residual State Scores* | | | | | | | | | | | | | | | | | | | | |
| O: $I_A$ | - | **.58*** ** | -.10 | .05 | .26 | .16 | .16 | .18 | .33 | .31 | .34 | .16 | .16 | .04 | -.03 | .04 | -.26 | -.19 | -.23 | -.11 |
| O: $I_{A'}$ | **.58*** ** | - | .04 | .14 | .14 | .27 | -.06 | .03 | .18 | .31 | .24 | .23 | .11 | .09 | .14 | .05 | -.23 | -.08 | -.18 | -.22 |
| O: $O_A$ | -.18 | .00 | - | **.54*** ** | .06 | .18 | .22 | .07 | .10 | -.02 | .17 | .16 | .21 | .35 | .34 | .22 | .00 | -.02 | .01 | -.02 |
| O: $O_{A'}$ | -.03 | .11 | **.58*** ** | - | .15 | .13 | .20 | .03 | .06 | .29 | .18 | .31 | .06 | .27 | .17 | .06 | .06 | .28 | .02 | .15 |
| C: $I_A$ | .25 | -.01 | -.10 | .05 | - | **.55*** ** | .18 | .01 | .24 | .30 | .12 | .24 | .09 | .10 | .11 | .07 | -.36 | -.22 | -.40** | -.35 |
| C: $I_{A'}$ | .18 | .16 | .06 | .07 | **.65*** ** | - | -.03 | -.15 | .21 | .23 | .22 | .30 | .05 | .08 | .20 | .05 | -.26 | -.28 | -.27 | -.36* |
| C: $O_A$ | .20 | -.06 | .20 | .19 | .17 | .01 | - | **.69*** ** | -.01 | .00 | .14 | .03 | .22 | .16 | .16 | .20 | .07 | .07 | .20 | .14 |
| C: $O_{A'}$ | .22 | .03 | .07 | .03 | -.03 | -.14 | **.72*** ** | - | .09 | .07 | .16 | -.01 | .19 | .02 | .16 | .09 | .02 | .13 | .14 | .17 |
| E: $A_A$ | .37* | .16 | .00 | -.02 | .20 | .18 | .00 | .09 | - | **.57*** ** | .59*** | .29 | .34 | .10 | -.14 | -.14 | -.16 | -.20 | -.41** | -.21 |
| E: $A_{A'}$ | .34 | .28 | -.09 | .22 | .23 | .18 | .02 | .08 | **.59*** ** | - | .49*** | .56*** | .25 | .28 | -.03 | -.25 | -.13 | .09 | -.34 | -.08 |
| E: $E_A$ | .35 | .22 | .02 | .01 | .16 | .17 | .16 | .20 | .59*** | .52*** | - | **.67*** ** | .40** | .24 | -.03 | -.07 | -.28 | -.20 | -.28 | -.15 |
| E: $E_{A'}$ | .18 | .22 | .06 | .17 | .24 | .23 | .01 | -.01 | .31 | .56*** | **.72*** ** | - | .29 | .40** | .04 | -.07 | -.17 | -.05 | -.25 | -.16 |
| A: $C_A$ | .13 | .05 | .17 | .00 | .06 | .08 | .29 | .26 | .32 | .22 | .34 | .23 | - | **.67*** ** | .08 | .13 | -.07 | -.09 | -.10 | -.03 |
| A: $C_{A'}$ | .02 | .04 | .34 | .25 | .04 | .07 | .18 | .05 | .08 | .24 | .20 | .34 | **.69*** ** | - | .19 | .16 | .01 | -.06 | -.04 | -.10 |
| A: $P_A$ | -.02 | .11 | .34 | .17 | -.06 | .07 | .08 | .07 | -.12 | -.06 | -.02 | .02 | .15 | .19 | - | **.60*** ** | -.32 | -.24 | -.18 | -.27 |
| A: $P_{A'}$ | .04 | .03 | .17 | .02 | -.06 | -.02 | .10 | .00 | -.11 | -.24 | -.05 | -.06 | .12 | .13 | **.65*** ** | - | -.08 | -.26 | -.08 | -.26 |
| N: $V_A$ | -.25 | -.17 | .08 | .12 | -.30 | -.20 | .08 | .04 | -.15 | -.11 | -.30 | -.21 | -.02 | .06 | -.26 | -.09 | - | **.50*** ** | .46*** | .35 |
| N: $V_{A'}$ | -.18 | -.01 | .11 | .35 | -.18 | -.20 | .06 | .12 | -.17 | .11 | -.23 | -.09 | -.03 | .01 | -.19 | -.27 | **.55*** ** | - | .33 | .61*** |
| N: $W_A$ | -.23 | -.12 | .12 | .11 | -.34 | -.20 | .15 | .12 | -.38* | -.31 | -.32 | -.29 | -.04 | .01 | -.15 | -.12 | .43** | .33 | - | **.50*** ** |
| N: $W_{A'}$ | -.12 | -.14 | .07 | .20 | -.23 | -.21 | .13 | .19 | -.19 | -.06 | -.20 | -.20 | .04 | -.04 | -.17 | -.22 | .34 | .58*** | **.58*** ** | - |

*Note.* $N$ = 99. Gray-shaded and bold-faced: convergent correlations between the same aspect. Subscripted A and A' mean "at Situation A" and "at Situation A' ", respectively. O: I = Openness Intellect, O: O = Openness Openness, C: I = Conscientiousness Industriousness, C: O = Conscientiousness Orderliness, E: A = Extraversion Assertiveness, E: E = Extraversion Enthusiasm, A: C = Agreeableness Compassion, A: P = Agreeableness Politeness, N: V = Neuroticism Volatility, N: W = Neuroticism Withdrawal. Upper half: Correlations of state scores. Lower half: Correlations of state residual scores. The lower triangle is based on reliable change index scores (RCIs), and the upper triangle on regression residual scores (RES).

Table 1.2.4

*Intercorrelations of Individual Differences in Simple and Residual Consistency Scores (Variable-Oriented)*

| Domains and Aspects | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| *For State Scores* | | | | | | | | | | |
| *Openness* | | | | | | | | | | |
| 1. Intellect | – | | | | | | | | | |
| 2. Openness | -.01 | – | | | | | | | | |
| *Conscientiousness* | | | | | | | | | | |
| 3. Industriousness | .12 | .07 | – | | | | | | | |
| 4. Orderliness | .07 | -.05 | .06 | – | | | | | | |
| *Extraversion* | | | | | | | | | | |
| 5. Assertiveness | .15 | .24 | .04 | -.03 | – | | | | | |
| 6. Enthusiasm | .23 | -.01 | .20 | .00 | .14 | | | | | |
| *Agreeableness* | | | | | | – | | | | |
| 7. Compassion | -.02 | .16 | .18 | .11 | .04 | .06 | – | | | |
| 8. Politeness | .21 | .16 | .09 | .23 | .02 | .13 | .08 | – | | |
| *Neuroticism* | | | | | | | | | | |
| 9. Volatility | .11 | .16 | .03 | .25 | .34* | .20 | .06 | .06 | – | |
| 10. Withdrawal | -.05 | .23 | -.11 | .05 | .18 | -.04 | .19 | .01 | .13 | – |
| *For State Residuals* | | | | | | | | | | |
| *Openness* | | | | | | | | | | |
| 1. Intellect | – | -.01 | .11 | .07 | .13 | .25 | .02 | .21 | .05 | -.04 |
| 2. Openness | -.02 | – | .00 | -.10 | .26 | .00 | .17 | .16 | .22 | .21 |
| *Conscientiousness* | | | | | | | | | | |
| 3. Industriousness | .19 | .01 | – | .05 | -.01 | .16 | .18 | .15 | .05 | -.09 |
| 4. Orderliness | .07 | -.08 | .08 | – | -.04 | .01 | .07 | .22 | .15 | .05 |
| *Extraversion* | | | | | | | | | | |
| 5. Assertiveness | .13 | .23 | -.03 | -.03 | – | .11 | .06 | .02 | .30 | .23 |
| 6. Enthusiasm | .24 | .00 | .18 | .00 | .10 | – | .09 | .15 | .17 | -.05 |
| *Agreeableness* | | | | | | | | | | |
| 7. Compassion | .01 | .14 | .17 | .05 | .08 | .09 | – | .12 | .07 | .23 |
| 8. Politeness | .21 | .15 | .19 | .23 | .02 | .15 | .13 | – | .05 | -.02 |
| *Neuroticism* | | | | | | | | | | |
| 9. Volatility | .06 | .24 | .07 | .17 | .31 | .18 | .04 | .08 | – | .09 |
| 10. Withdrawal | -.04 | .24 | -.02 | .05 | .23 | -.08 | .22 | -.05 | .14 | – |

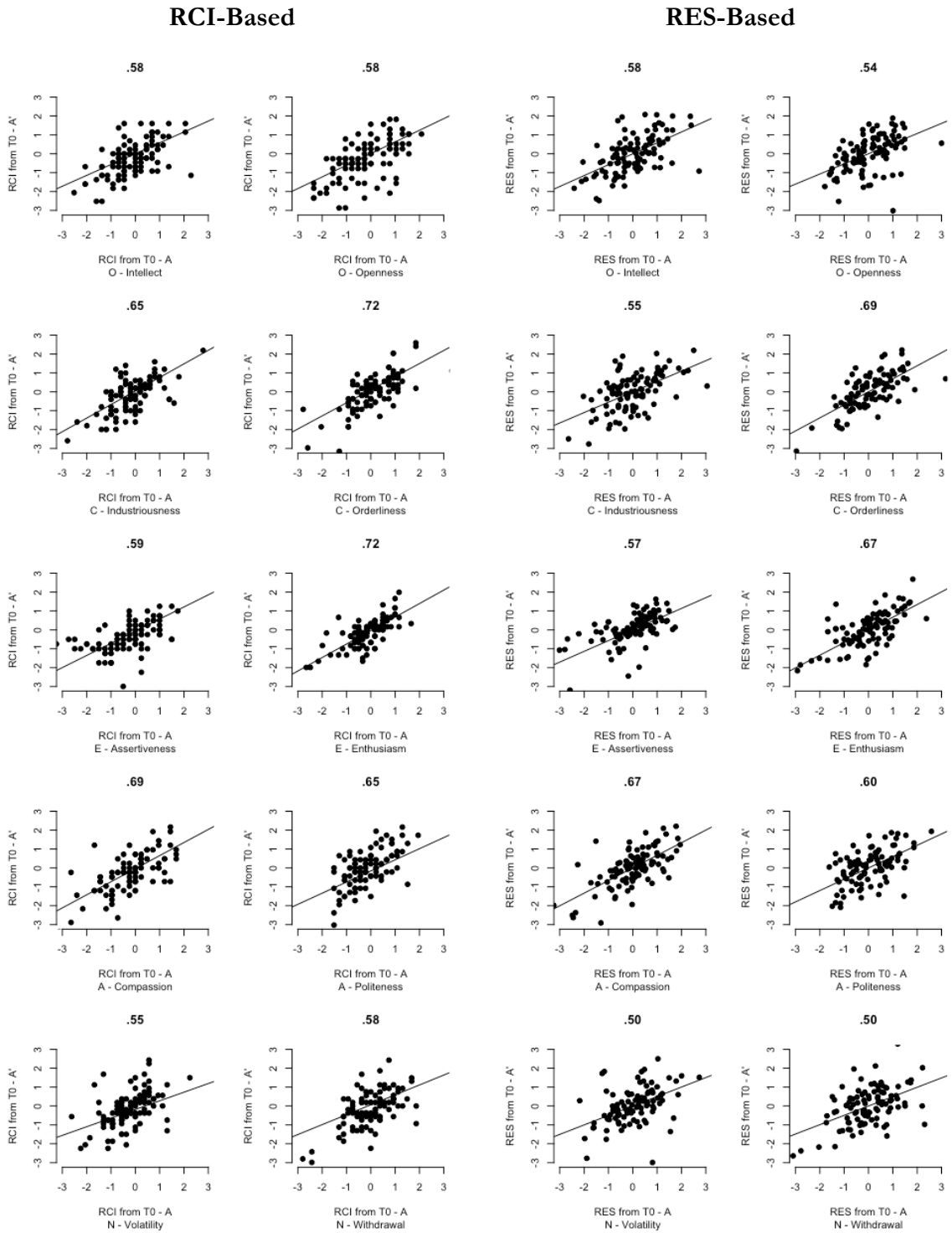
*Note.* $N = 99$. Upper half: Correlations of individual simple consistency scores (based on states). Lower half: Correlations of individual residual consistency scores (based on state residuals). The lower triangle is based on individual consistency scores derived from reliable change index scores (RCIs), and the upper triangle on those derived from regression residual scores (RES).

*$p < .05$.

**Residual Consistencies.** As we have explained under *Data Analysis*, we used two types of residual state scores: reliable change indices (RCI) and regression residuals (RES). Participants' RCIs and RES scores were generally highly correlated within the same aspect, ranging from .81 (Industriousness, at Situation A) to 1.00 (Intellect, at Situations A and A'). This suggests that all subsequent findings should replicate regardless whether RCI or RES scores are used.

First, residual consistencies ($r_{RCI}$ and $r_{RES}$), as the correlations between state residual score at Situation A and A', can be found in Table 1.2.2 under the second and third column of "Consistencies." Additionally, the scatterplots of these correlations are presented in Figure 1.2.4. As can be seen, residual consistencies were generally high (RCI and RES scores yielded comparable results), ranging from .50 (Withdrawal) to .72 (Enthusiasm). Thus, participants showed strong levels of residual rank-order consistency which means that residual state variance was reliable across functionally equivalent situations.

*Figure 1.2.4.* Scatterplot of Residual State Scores per Aspect




*Note.* N = 99. RCI = reliable change index score (see two left plots); RES = regression residual score (see two right plots). O = Openness, C = Conscientiousness, E = Extraversion, A = Agreeableness, N = Neuroticism. T0 = Trait measurement; A = state measurement in Situation A; A' = state measurement in Situation A'.

Second, we again inspected the convergent and discriminant correlations of residual state scores, which are presented in Table 1.2.3 under "For Residual State Scores" (lower triangle under the diagonal: RCI scores used; upper triangle above the diagonal: RES scores used). As can be seen, almost none of the residual state scores from different aspects were substantially correlated with each other, indicating good discriminant validity of the scores. Specifically, for RCI scores, convergent correlations (as already presented in Table 1.2.2 under $r_{RCI}$) amounted on average to .64 (*SD* = .11), while the off-diagonal discriminant ones amounted to an average *absolute* correlation of only .16 (*SD* = .13). This pattern was replicated for RES scores, with convergent correlations (as already presented in Table 1.2.2 under $r_{RES}$) amounting on average to .59 (*SD* = .11), while the off-diagonal discriminant ones amounted to an average *absolute* correlation of only .18 (*SD* = .13). These patterns suggest that state residuals are aspect-specific and intercorrelations not likely reducible to method or artifact effects.

Lastly, we also analyzed how inter-individual differences in residual consistencies were related to each other. As one would expect, the individual consistency scores derived from RCI and RES correlations showed high convergent correlations when the same facets were correlated, with *r*s ranging from .84 (Industriousness) to 1.00 (Politeness), as can be seen in Table 1.2.2 under "Convergences ($r_{RCI./RES}$)." However, as our simulations show, this is not a good estimate for true convergence – these correlations will always be rather high due to the common variance in the residuals introduced when subtracting the trait score from the state score. A better indicator for the question if RCIs and RES scores yield the same results, though, is the absolute correlation of the difference score RES and RCI for different aspects, as presented in the lower half of Table 1.2.4 under "For State Residuals" (lower triangle under the diagonal: RCI scores used; upper triangle above the diagonal: RES scores used). As can be seen, correlations were rather small (as has already been the case for inter-individual differences in simple consistencies) but similar for RES and RCI. For RCI-based scores, the average *absolute* intercorrelation amounted to .12 (*SD* = .08), with the highest correlation being .31 (between Volatility and Assertiveness). For RES-based scores, the average *absolute* intercorrelation amounted to .11 (*SD* = .08), with the highest correlation being .30 (again, between Volatility and Assertiveness). In both cases, the lack of covariation does not speak for g-factor of residual consistency.

**Person-Oriented Analyses**

**Trait Expressions.** Correlations between the profile of trait scores at T0 and the corresponding profile of state scores at Situations A and A' ($q_{T0.A}$, $q_{T0.A'}$) are presented in Table 1.2.5 under "Trait Expressions." As can be seen, profile-based trait expressions were high and

similar between Situations A and A' (on average, .70 and .71). Additionally, there were some individual differences in these trait expressions (see *SD*s of .25 and .26 in Table 1.2.5).

Table 1.2.5
*Descriptive Statistics of Simple and Residual Consistency (Person-Oriented)*

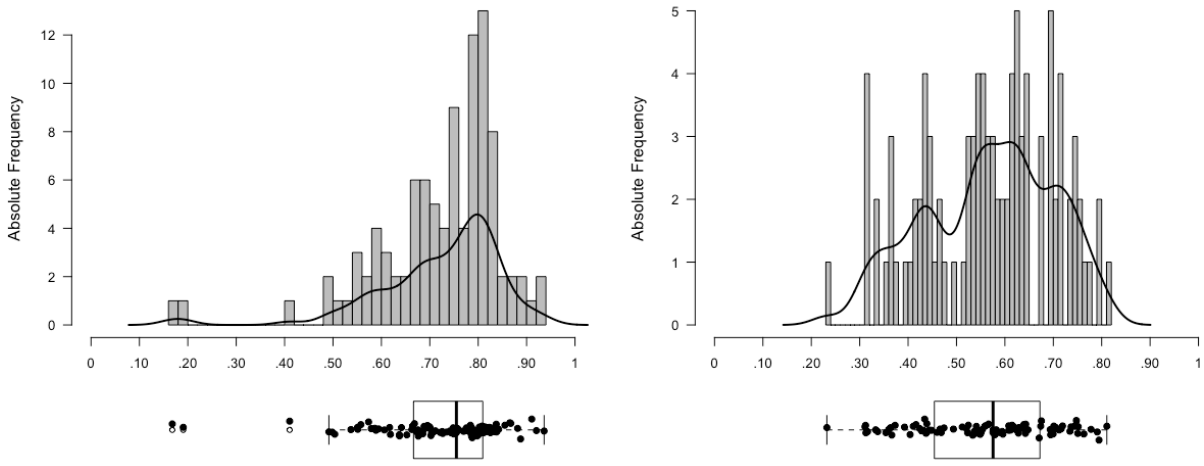
| Descriptive Statistics | Trait Expressions | | State Consistencies | |
|---|---|---|---|---|
| | $q_{T0.A}$ | $q_{T0.A'}$ | $q_s$ | $q_r$ |
| Median | .70 | .70 | .76 | .58 |
| *M* | .71 | .70 | .74 | .58 |
| *SD* | .25 | .26 | .25 | .20 |
| Minimum | .29 | .19 | .17 | .23 |
| Maximum | .94 | .94 | .94 | .81 |

*Note.* $N = 99$. Simple consistency is the within-person correlation on item level, residual consistency is the within-person correlation of residuals of the prediction of item at Situation A and A' with the item at T0.

**Simple Consistencies.** The correlation between the state profile scores at Situations A and A' ($q_o$) is presented in Table 1.2.5 under "State Consistencies" (first column). It amounted on average to .74, with noticeable individual differences (*SD* = .25). The distribution of *q*-correlations can be found on the left side in Figure 1.2.5, better showing the magnitude of individual differences in simple profile consistencies.

**Residual Consistencies.** First, we computed state residual profiles by subtracting each individual's trait profile from both their state profiles at Situation A and A'. Next, we examined the correlation between the residual state profile scores at Situations A and A' ($q_r$), as presented in Table 1.2.5 under "State Consistencies" (second column). It amounted on average to .58, again with noticeable individual differences (*SD* = .20). The distribution of *q*-correlations can be found on the right side in Figure 1.2.5, demonstrating the magnitude of individual differences in residual profile consistencies.

*Figure 1.2.5.* Distributions of Person-Oriented Consistency Scores




*Note.* Individual *q*-correlations (across 100 items) are given. Left: Simple *q*-correlations between state profiles at Situations A and A'. Right: Residual *q*-correlations between residual state profiles between at Situations A and A'. The upper half of the figures depicts histograms and smoothed density distributions, the lower half depicts box-plots where each individual *q*-correlation score is depicted as a dot.

## Discussion

Using a longitudinal design and two functionally equivalent situation vignettes, we showed that trait expressions, or states, consist of reliable trait variance as well as systematic situation variance. The concept of residual consistency represents a valuable extension to the concept of simple consistency and suggests the simultaneous systematic influence of the situation and personality traits on behavioral states.

### Simple and Residual Consistencies

As we have shown in Table 1.2.2, trait expressions were uniformly high, meaning that state variance was tied to trait variance, which also means that any simple consistency estimate derived from raw state scores contains trait variance. Simple consistency could thus be high due to either stable tendencies within the person (i.e., traits) or stable tendencies within the situation (i.e., functionally equivalent situations harbor situation characteristics that evoke similar mental and behavioral reactions). Indeed, several studies have sought to rule out the second explanation of consistency by having participants engage in *functionally different* situations (Back, Schmukle, & Egloff, 2009; Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2004; Fleeson & Law, 2015; Leikas, Lönnqvist, & Verkasalo, 2012; Morse, Sauerberger, Todd, & Funder, 2015; Riemann, Angleitner, Borkenau, & Eid, 1998; Weisbuch, Slepian, Clarke, Ambady, & Veenstra-VanderWeele, 2010) so that any observed consistency would be due to the actor or traits within individuals. If there was no trait expression, simple and residual consistency would yield identical results (taking the method-artifacts into account). Interestingly, one would surmise that, with high levels of trait expression in both situations and thus high levels of simple consistency, any residual consistency should be low. As we have shown (see Table 1.2.2 as well as Figures 4 and 5), this was neither the case nor is this due to auto-correlations. Though residual consistencies, whether computed variable- or person-oriented, were smaller than simple ones (which they necessarily need to be there is any trait-expression), they were not much smaller. Together, this pattern of findings suggests that state residuals – or "states freed from traits" – are indeed reliable, and distinctive consistencies can be traced back to the functional equivalence of two situations.

Although we deem our estimates of residual consistency relative high, they did not reach 1.0. This means that there was unsystematic variance which could be traced back to measurement error and/or systematic but unreliable person-specific effects, perhaps stemming from short-term affective or physiological fluctuations that were present while responding to the situational vignettes (Horstmann & Ziegler, 2018).

**Residual Consistency: What Does It Show Us?**

Substantial estimates of residual consistency, even if somewhat smaller than simple consistency estimates, can show at least two things. First, they provide evidence that controlling for trait levels of participants decreases but does not extinguish consistency. This should decidedly not be interpreted as the situation being more powerful than the person or traits in shaping behavior. Rather, this finding needs to be considered with the simultaneously high trait expressions across functionally equivalent situations in mind. Thus, it is safe to conclude that both traits and situations drive consistency, and thus that we may miss out on important information if we prioritize one over the other. This is consistent with recent theorizing and empirical results that have shown how persons and situations transact (e.g., Funder, 2016; Rauthmann & Sherman, 2016b; Rauthmann, Sherman, & Funder, 2015b; Wrzus, Wagner, & Riediger, 2016).

Second, our work may also be interpreted as a powerful and stringent test of the functional equivalence of situations. Computing simple consistency estimates between two supposedly functionally equivalent situations will not be enough because they contain substantive amounts of trait variance (see, e.g., our trait expression estimates). However, controlling for trait levels and using resulting state residuals provides a stronger test of situations' functional equivalence. In other words, the higher the functional equivalence of situations is[7], the higher residual consistencies need to be. We used situation vignettes that were on face validity alone pretty equivalent, and thus our working assumption was that the situations were indeed functionally equivalent. Because we were interested in disentangling simple from residual consistency *given that* our situations were supposed to be functional equivalent, we did not interpret our findings as evidence for the functional equivalence of our situations (which would be circular and then only a psychometric exercise of testing situation vignettes). However, it has not escaped our notice that the data-analytical procedures presented here can be used in other research as stringent tests for situational functional equivalences, especially when taking into account the results from the accompanying simulation.

**Intercorrelations of Individual Differences in Consistencies: No underlying g-factor?**

Our findings allow new insights into the structure of variable-oriented consistency. For simple and residual forms, inter-individual differences in consistency scores from one aspect were barely correlated with another aspect. Indeed, the covariation among individual consistency scores is too low to warrant fitting a g-factor underlying it[8]. Thus, we failed to find support for broad cross-aspect consistency in variable-oriented analyses; rather, our findings point towards *aspect-*

[7] Note that we assume that situational functional equivalence resides on a continuum. Situations are not either equivalent or not; rather, they can be more or less equivalent.

[8] We compute the Kaiser-Mayer-Olkin measure of sample adequacy, which ranges from .61 to .66. However, additional exploratory factor analyses revealed that no clear one-factor solution could be established.

*specific* forms of individual differences in consistency. Note, however, that such a lack of cross-aspect intercorrelations need not coincide with substantial profile consistencies at the individual level (see Table 1.2.5) as variable- and person-oriented consistencies have different units of analysis and thus address each different consistency questions (Fleeson & Noftle, 2008).

Bem and Allen (1974) noted that asking the participants about their variability over time in one domain "permit[s] the individual to employ his [*sic*] own concept of the trait dimension, to average across situations he [*sic*] sees as pertinent and to ignore the situations he [*sic*] sees as irrelevant" (p. 512). Given the aspect-specificity of individual differences in consistency, it is possible that such a *broad* self-reported consistency measure contains different content and meaning-systems for each individual and thus appears to be unreliable across participants. Fleeson and Noftle (2008) have alerted us to the nuances of defining and measuring consistency in different ways. If consistency is not clearly defined by the similarity computation, competing determinants, and behavioral enactments, it is unclear what such a consistency is supposed to constitute. We have defined different types of consistency clearly here, yet individual differences in them do not seem to cluster strongly together.

**Content Specificity of Consistencies**

Contrary to pre-registered predictions about the situation vignettes primarily tapping or activating extraversion and agreeableness (e.g., Tett & Guterman, 2000), we found substantial simple and residual consistencies for *all* Big Five aspects. This means that the situation we selected was relevant to the consistencies we computed. Nonetheless, for variable-oriented analyses, there were slight differences in consistencies among aspects. The most consistent aspect was Orderliness and the least consistent ones Politeness, Volatility, and Withdrawal (though they still showed substantive levels of consistency). As the boxplots (Figure 1.2.3) and descriptive statistics (Table 1.2.2) show, this cannot be explained away by bottom or ceiling effects in the aspects activated.

As we have further shown, state and residual state scores obtained in Situations A and A' of different aspects were not just all similarly positively correlated, which would speak for common method biases and artifacts. Rather, we found clearly demarcated convergent correlations between the same state aspect measured at Situation A and A' which were much higher than the off-diagonal discriminant correlations (i.e., one aspect with other aspects in the same or the other situation). Thus, states and state residuals were specific to each Big Five aspect. While we were interested primarily in drawing substantive conclusions here, these findings may also be taken as evidence for desirable psychometric properties of state-versions of the BFAS (DeYoung et al., 2007).

**Limitations and Future Directions**

The limitations to this study point towards future research that would ideally replicate, corroborate, and also extend our findings. First, we aimed at sampling 250 participants, but only 99 participants completed the whole study. Due to the within-person design and the clear patterns established here and in the simulations (see supplemental materials), findings are unlikely to change using a larger sample size with the data-analytical plan presented here. However, we could not make use of latent change score models, as initially pre-registered.

Second, we only used one situation that was encapsulated in two vignettes. Future research may seek to employ different kinds of situations (i.e., variation in situation content) and examine whether our findings can then be replicated. Notably, it will be interesting to estimate to what extent different combinations of different person-populations (e.g., pure community sample) and situation-vignette contents would lead to similar or different patterns of findings.

Third, this study resorted to hypothetical situation vignettes and people's self-reports of their imagined Big Five aspect states in those situations. Thus, our findings hold for people's self-views. Future studies should employ actual situations, likely standardized within a laboratory setting (because at least two functionally equivalent situations need to be administered to each person), and measure behavior not just via self- or even experience-sampling reports, but also via behavioral observation or peer-ratings of actual behavior shown in the situations.

Fourth, in the context of the previous point it is important to note that laboratory studies will yield opportunities for easily standardizing situations and thus ensuring that people can enter a certain situation twice. Assigning participants randomly to such experimental situations, their traits and the characteristics of the situations will not be substantially correlated. However, in real life people may encounter functionally similar situations that they may have chosen or changed (Emmons & Diener, 1986; Rauthmann & Sherman, 2016b) – and thus such situations may be aligned with their personality traits (Ickes et al., 1997; Rauthmann, Sherman, Nave, & Funder, 2015). Once traits and situations are correlated, associations between traits with states (trait expressions) and states with states (simple consistency) or state residuals with state residuals (residual consistency) will also change. Thus, any findings on the forms of consistencies we presented need to be replicated also in naturalistic settings within people's everyday lives.

Lastly, we have shown that variable-oriented inter-individual differences in both simple and residual consistencies were barely correlated with each other and that person-oriented inter-individual differences in both simple and residual profile consistencies were substantial. For both variable- and person-oriented consistencies, we could not examine to what extent the inter-individual differences were actually stable, meaningful, and consequential. To do so, future studies

should extract individual difference scores (as we did) and (a) insert them into a nomological network of sensible correlates (e.g., inter- and intrapersonal adjustment variables) and (b) examine their rank-order stability across extended periods of time. Conceptually plausible nomological networks and cross-temporal stability of inter-individual differences in simple and residual consistencies would suggest meaningful individual differences.

## Conclusion

We have examined several research questions for variable- and person-oriented forms of simple and residual consistency (Table 1.2.1). We found that all forms were relatively high, and we were especially interested in residual forms of consistency: enacting certain states consistently in the same functionally equivalent situations independent of one's trait levels. In other words: *every time* someone steps into cold water, then she or he shrieks – regardless of her or his general tendency to shriek.

## 1.3 Personality Theories after the Person-Situation Debate

The recognition that persons vary systematically across time, behavior content, and/or situations has influenced and advanced many personality theories or frameworks in the recent years (Baumert et al., 2017; Fleeson & Noftle, 2009; Sherman et al., 2015). These theories recognize that both traits and states are an essential part of personality.

**Trait Theories.** Trait theories assume that stable personality traits are responsible for reliable inter-individual differences in behavior (Sherman et al., 2015). The most commonly known and probably most widely applied personality theory is the Five Factor Theory (FFT, Costa & McCrae, 1995; John, Naumann, & Soto, 2008; McCrae & Costa, 2008; Pervin & John, 1999). Measures of personality traits were shown to be cross-temporally stable, heritable, and related to a variety of life outcomes (Funder, 2001; Ozer & Benet-Martínez, 2006; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). By using reliable and valid measures, FFT or other similar trait theories (such as the HEXACO model, Ashton & Lee, 2007; Lee & Ashton, 2008) therefore accomplish to a certain extent what was initially required of personality measures, namely the prediction of aggregated behavior.

**Latent State-Trait (LST).** LST (Deinzer et al., 1995; Steyer, Ferring, & Schmitt, 1992; Steyer et al., 2015, 1999) was developed to "provide a solution to the person-situation debate" (Steyer et al., 1992, p. 79). This theory suggests that momentary instances of personality expressions (personality states) can be decomposed into a latent trait (which is stable) and an "occasion specific residual" (Steyer et al., 1999, p. 392) that represents effects of the situation as well as the person x situation interaction. The strength of this theory lies in the fact that it is explicitly testable and can model personality traits and occasion-specific states at once. At the same time, this theory does not provide an explanation for the differences in traits and states, other than broad occasion-specific effects. Further, the situation per se is usually not measured but just assumed to be existent and effectual.

**Cognitive Affective Processing System (CAPS).** Similar to LST, CAPS (Mischel & Shoda, 1995) was also suggested to "reconceptualize[…] situations, dispositions, dynamics, and invariance in personality structure" (Mischel & Shoda, 1995, p. 246, slightly adapted for readability). Although Mischel's book in 1968 has been most widely recognized as key element of the person-situation debate because of its focus on behavioral variability (among other things), Mischel and Shoda stated that "variability reflects some of the essence of personality coherence" (Mischel & Shoda, 1995, p. 246). The authors argued that certain situations lead to certain behaviors, and that such situation-behavior contingencies reflect stable interindividual differences,

mediated by stable affective-processing units. These patterns were called *if…then contingencies: If* Situation A occurs, *then* Person Y will enact Behavior B. Interindividual differences in these processing units thus explain why people act differently in similar situations. Recognizing these stable patterns as personality itself thus unifies both stable and variable aspects of personality.

**Trait Activation Theory (TAT).** TAT (Tett & Burnett, 2003; Tett & Guterman, 2000) posits that certain personality traits are only expressed in relevant situations. The activation of traits in certain specific situations resembles *if…then* contingencies of behavior (Mischel & Shoda, 1995) and also leads to consistency in behavior across situations. Contrary to LST and CAPS, TAT is explicitly concerned about the interaction of traits with certain situations (Sherman et al., 2015). An activation can be interpreted as a moderation: The trait (or, "propensities" to act, Tett & Guterman, 2000, p. 398) is present in all situations but only active (via behavior) in some situations. This theory therefore also combines stable and variable aspects of the person, and clearly states that situations are relevant to the expression of personality.

**Whole Trait Theory (WTT).** WTT is the most recent theory that unifies stable personality traits with intra-individual variability (Fleeson, 2001; Fleeson & Gallagher, 2009; Fleeson & Jayawickreme, 2015). WTT suggests that personality traits have two parts: explanatory and descriptive. The explanatory part connects input, which can be external events (e.g., situation cues) and internal events (e.g., cognition), to behavioral outputs, which are conceptualized as momentary manifestations of traits (which are states). Mediating units (such as the persons' goals, but also physiological states) connect the input with the output. This explanatory part therefore allows explaining how persons behave in a situation. Fleeson and Jayawickreme (2015) clearly state that this explanatory part resembles cognitive-affective processing units suggested by Mischel and Shoda (1995) and is thus concerned with the mechanisms of trait-expression.

The descriptive part of traits can be represented in density distributions of states. If persons act across several occasions in their everyday life, a distribution of their states can be formed. The mean of this distribution then corresponds best with other trait measures, which has received strong empirical support (Augustine & Larsen, 2012; Fleeson & Gallagher, 2009; Horstmann & Rauthmann, in preparation; Rauthmann, Horstmann, & Sherman, 2018). One could even go so far and say that the descriptive part of traits incorporates ideas formulated much earlier about the relation of traits and aggregated states by Epstein (Epstein, 1980, 1983a, 1983b), but also by LST or trait theories. In LST, the common variance of several states is conceptualized as the personality trait (Steyer et al., 1992, 2015, 1999). Note that the difference between Epstein's ideas and WTT/LST is that Epstein thought of variability as measurement error (similar to classical test-theory), whereas both WTT and LST assume that variability is systematic.

**Further theories.** Many other theories or models were proposed in recent years that describe the relation of personality traits and states. DeYoung (2010, 2015), for example, focused more on biological or evolutionary aspects of personality traits. McAbee and Connelly (2016) emphasized the role of contexts, others, and identity to understand personality traits. Read, Smith, Droutman, and Miller (2017) proposed approach and avoidance as mediating units between situational affordances and behavior. Revelle and Condon (2015) brought attention to the fact that different behaviors exclude each other, thus leading to variability in behavior over time, and Wrzus and Roberts (2017) recognized stable and variable aspects of personality in a developmental framework. Without going into too much detail, it is apparent that modern personality theories and frameworks seek to combine intra-individual variability with inter-individual stability in behavior (Baumert et al., 2017).

## 1.4 Conclusion Part 1

The person-situation debate revealed that situations and personality traits must be considered simultaneously to understand variability and consistency in behavior. Although this idea was not new (e.g., Cattell, 1963; Lewin, 1936; Magnusson, 1971; Murray, 1938), the debate had brought forth several personality models and theories that advanced situations and their effects as integral parts. Although this recognition carried the chance to benefit the field in many ways, a new challenge presented itself, namely systems of description and measurement of situations. The lack of such systems does not only hinder the examination of personality theories, it also limits the generalizability of research findings. One limitation of our previous study (Horstmann et al., in preparation) was, for example, that we did not measure the level of functional equivalence of situations or even the reason for their equivalency. Although the situations were written in a way such that they were very similar, it could also be the case that slightly different situations have a comparable functionality.

In order to obtain a new "locus of consistency" (Mischel et al., 2002), it was thus necessary to describe and measure situational influences in a broad and generalizable manner by recognizing the influences of situations on behavior. This research is summarized in Part 2.

# 2 Part 2: The Description and Measurement of Situations

What is a situation, and how can it be measured? In 1971, Magnusson wrote "concerning the study of situational variation, we find ourselves at the same stage as that concerning the study of individual differences at the initial development period of differential psychology" (p. 852). He states that reaching this stage took about 50 to 60 years of research – and the same time may be required with regard to the description and measurement of situations. Forty-seven years after this statement, considerable progress has indeed been made, and the description and measurement of situations has received a lot of additional and renewed attention in recent years (Horstmann, Rauthmann, et al., 2018). Substantial progress was possible because general frameworks and principles for studying situations were developed (Rauthmann, 2012, 2015; Rauthmann et al., 2014; Rauthmann & Sherman, 2018a, 2018b; Wagerman & Funder, 2009), and several situational taxonomies were published and applied in different areas of research.

## 2.1 Background: Situation Research

Rauthmann, Sherman, and Funder (2015b) developed *Principles of Situation Research* that present a framework for the examination of situational information. These principles and the terms defined by Rauthmann and colleagues are essential pre-requisites for the remainder of Part 2 and for the interpretation of the results from Part 3.

**Cues, Characteristics, and Classes.** First, and important for understanding subsequent research on situations, is the definition of different types of situational information that can be conceptualized and assessed: Cues, Characteristics, and Classes. These reflect three broad ways of thinking about situations, and they have been focused on and utilized to different extents in different sub-fields of psychology.

***Cues.*** Situation cues describe physically present elements in a situation. These can, for example, be the number of people in a room, the objects in a room, the place or location, temperature, but also the time of the day. Cues are – in principle – measurable. If two perfect and omniscient raters were tasked with listing all cues in a situation, they should come up with the same list of objects (Horstmann, Ziegler, & Ziegler, 2018). Situation cues have been extensively used in psychological research, and some of the most famous experiments in social psychology relied on the manipulation of cues (Horstmann & Ziegler, 2016; Krueger, 2009; F. D. Richard et al., 2003). Examples include the examination of the bystander effect (Darley & Latané, 1968), the effect of conformity (Asch, 1956), or the (alleged) effect of holding a warm coffee cup on ratings of other persons (Lynott et al., 2014; Williams & Bargh, 2008). Some of the strong effects of

situation cues have even been cited to support the situation side of the person-situation debate (Bargh, 2007).[9]

***Characteristics.*** Situation characteristics refer to the psychologically meaningful aspects of a situation stemming from perceived and interpreted cues. For example, a situation may be stressful, boring, playful, or full of potential. This interpretation is of course not independent from any perceiver (Funder, 2016), and situation characteristics may not only reflect situation cues, but also the perceivers' experience with them, personality of a perceiver, or even cultural differences.

Recently developed situational taxonomies usually focus on the description and assessment of situation characteristics (Horstmann, Rauthmann, et al., 2018). The DIAMONDS, for example, describe eight dimensions of situation perception: Duty, Intellect, Adversity, Mating, pOsitivity, Negativity, Deception, and Sociality (Rauthmann et al., 2014; Rauthmann & Sherman, 2016a). Additionally, measurement tools for the assessment of situation characteristics have been developed (N. A. Brown et al., 2015; Gerpott et al., 2017; Parrigon et al., 2017; Rauthmann & Sherman, 2016a, 2016c; Ziegler, Horstmann, & Ziegler, submitted), and situation characteristics have been linked to other outcomes, such as behavior (Sherman et al., 2015) and age (N. A. Brown & Rauthmann, 2016).

**Classes.** Situation classes represent groups of situations that are either similar in their cues and/or characteristics (Rauthmann, Sherman, & Funder, 2015b). Examples for situation classes could be "at home" or "at work". Several previous situational taxonomies defined situations in terms of their class, yet an exhaustive descriptive system of situation classes still does not exist (Horstmann, Rauthmann, et al., 2018). One possible reason could be that most situations may not simply fall in one class. For example, I am currently at home, writing my dissertation, and this situation might fall in both classes, "at work" and "at home". Nevertheless, classes – usually generated ad hoc – were successfully used in psychological research and related to behavioral outcomes (e.g., Geukes et al., 2017; Oud, Voelkle, & Driver, 2017).

**Principles.** How are situation cues related to situation characteristics? Or, in other words, how can person characteristics, such as the interpretation of situation cues, be separated from the "true nature" of a situation? To guide situation research, Rauthmann and colleagues (2015b) have defined three principles of situation research that allow understanding the relation of objective cues and subjective perceptions of situations: the Processing, Reality, and Circularity principles.

***Processing Principle.*** This principle states that psychological experiences of situations matter. Simply put, a situation that is not perceived does not play a role in a psychological sense.

[9] One might cautiously argue that the situation side of the debate would have lost many of its arguments if the replication crisis in psychology had struck earlier as many of the situational effects could not be replicated (e.g., Lynott et al., 2014; Open Science Collaboration, 2015; Wagenmakers et al., 2016).

Thus, for behavior to be instigated, there needs to be some explicit and/or implicit processing of situation cues.

***Reality Principle.*** This principle states that situations have three realities: Physical, consensual, and idiosyncratic. The physical reality refers to the actually present cues in a situation, the consensual reality refers to the shared perception and interpretation of a situation by multiple perceivers, and the idiosyncratic reality refers to what each person uniquely perceives in a situation. This principle also comes with three corollaries. The first, the agreement corollary, states that people will, to some extent, share a certain interpretation of a situation. The second, the variation corollary, states that persons will differ in their interpretation of situations. The third, the componentiality corollary, states that situation perceptions consist at least of perceiver and situation variance as well as their interaction (Rauthmann & Sherman, 2018b).

***Circularity Principle.*** This principle states that situations and persons are conflated if a situation is defined via a person variable. For example, describing situations in terms of what people are doing (e.g., playing lacrosse) violates this principle, and such situations should instead be defined in terms of their cues (e.g., a lacrosse field, persons, a ball, a goal) or by their characteristics (e.g., the situation requires physical exercise). Several corollaries again extend this principle. First, the state corollary states that situations should not be defined in terms of mental or behavioral states. Second, the consequences corollary states that situations should not be defined in terms of their consequences. Third, the approximation corollary states that the true nature of a situation can best be approximated if a situation is described from different perspectives and sources, as averaging multiple ratings will cancel out idiosyncratic variance of perceivers.

## 2.2 Assessing Situational Information

We have highlighted some of the challenges researchers face during the construction of assessment tools for situational information – specifically situation characteristics – in our book chapter *Assessment of Situational Perceptions: Measurement issues and a joint taxonomization of persons and situations* (Horstmann, Ziegler, et al., 2018). We exemplify there how different approaches to the taxonomization of situations can lead to different results and different assessment tools.

### *2.2.1 Book Chapter Summary: Assessment of Situational Perceptions*

**Background.** The person-situation debate revealed the necessity to define and measure situational information (e.g., Magnusson, 1971; Rauthmann, Sherman, & Funder, 2015b). Taxonomies allow describing and classifying information, which in turn enables the generalization and communication of research findings. Rauthmann (2015) outlined that to assess situational information, six different questions have to be answered: 1) Theory: Which theories and concepts underlie the research?, 2) Information: Which situational information is in the focus?, 3) Aims: How should the situational information be structured?, 4) Generation/Pool: How is the situational information generated or sampled?, 5) Assessment: How is the situational information assessed and judged?, and 6) Analysis: How is the situational information analyzed? (Rauthmann, 2015, p. 177). Given that there are many different answers to each question, a plethora of different situational taxonomies is possible. However, the taxonomization alone does not yet allow assessing the situational information, and assessment tools were therefore needed as well.

**Review.** To exemplify the different approaches that can be taken, we reviewed three different situational taxonomies and their corresponding measurement tools. To do so, we followed recommendations by Ziegler (2014c), who suggested that for each test-construction, the following three questions (referred to as the ABC of test-construction) should be answered:

1. What is the construct being measured?
2. What is the intended use of the measure?
3. What is the targeted population?

Based on the ABC of test-construction, we reviewed the Situational Eight DIAMONDS taxonomy (Rauthmann et al., 2014) and the corresponding measurement tool (Rauthmann & Sherman, 2016a), the Situational Affordances and Adaptive Problems scale (SAAP; N. A. Brown et al., 2015), as well as the Situation Five (Ziegler, 2014a; Ziegler et al., submitted).

Each taxonomy and measurement tool provide different answers to the three questions. For example, the DIAMONDS refer to dimensions of situation perception in every-day situations, the SAAP apply to dimensions of perception of evolutionary relevant situations, and the Situation Five describe situation perception at trait level in an occupational context. Similarly, the intended use of each measure was different: Whereas the DIAMONDS were intended to describe every-day situations, the SAAP can mainly be applied to evolutionary relevant situations, and the Situation Five can best be applied in an occupational setting. Finally, the targeted population also differed substantially: The DIAMONDS and Situation Five may be applied to all populations, but the SAAP apply to some specific populations only, for example parents.

Furthermore, the measurement of situations is different to the measurement of personality. Specifically, ratings of situations or measurements of situation perception are per definition not unidimensional, as they contain situation-specific variance. This unique variance may lead to reduced model fit (and thus lower factor reliability) for measurement models of situation perception – a fact observed in many models of situation perception.

**Conclusion.** The measurement of situational information is a complex endeavor, and principles that apply to the assessment of personality may not readily be applied to the assessment of situational information. Similar to the application of personality measures, each measurement tool for situations has its strengths and weaknesses that must be considered when assessing situational information.

### *2.2.2 Book Chapter: Assessment of Situational Perceptions*

**Assessment of Situational Perceptions**

**Measurement issues and a joint taxonomization of persons and situations**

Kai T. Horstmann, Johanna Ziegler, & Matthias Ziegler
Humboldt-Universität zu Berlin, Germany

This book chapter was published as:

Horstmann, K. T., Ziegler, J., & Ziegler, M. (2018). Assessment of Situational Perceptions: Measurement issues and a joint taxonomization of persons and situations. In D. C. Funder, J. F. Rauthmann, & R. A. Sherman (Eds.), *The Oxford Handbook of Psychological Situations*. Oxford University Press.

This is the final author copy before copy editing.

**Abstract**

The assessment of situations and especially situational perceptions is the focus of this chapter. Based on general principles of test construction, the ABC of test construction (Ziegler, 2014c), and the road map to the taxonomization of situations (Rauthmann, 2015), this article shows how situational taxonomies and their assessment tools can be developed. These principles are exemplified by presenting three recent situational taxonomies and the effect different approaches have on the resulting taxonomy. Similarities and differences to established taxonomizations of personality traits (such as the Big Five) are discussed. Furthermore, a new taxonomy and assessment tool is presented which captures personality traits and situational perception at the same time. Finally, challenges of future situational taxonomization, especially the need to establish a nomological net of situational perception and other, related constructs and psychological processes are discussed.

To get a good understanding of "situational assessment," we start with a simple, brief example. Imagine you and two good friends are all invited to a barbecue. You have been to barbecues before with each of your friends, so you have an idea of how they are likely to behave. However, what if someone asked you how your friends are going to behave at this very specific barbecue? Which information would you need to predict their behavior at this specific barbecue?

Of course, this will largely depend on your friends' personality (Paunonen & Ashton, 2001; Roberts et al., 2007; Ziegler, Bensch, et al., 2014), which may even not be very different from yours (Maaß, Lämmle, Bensch, & Ziegler, 2016; Youyou, Stillwell, Schwartz, & Kosinski, 2017). Your extraverted friend is very likely to talk to a lot of people, while your overly agreeable friend might help out in the kitchen. However, your friends' behavior will also depend on their current state, for example their affect: Maybe one of your friends recently broke up with his or her partner and is therefore likely to get either completely drunk or very flirty (Ching et al., 2014; R E Lucas, Diener, Grob, Suh, & Shao, 2000). Your other friend could just have received a good job offer, which might make him or her more confident or happy (Wrzus, Hänel, Wagner, & Neyer, 2013).

What is still lacking, though, to predict your friends' behavior, is information on the actual situation itself. First, imagine the barbecue as very formal. The prospective behavior of your friends is going to change, and maybe you consider not even taking them along. On the other hand, the barbecue could be a very informal party. There is no need to go into details: The effect of the situation should be obvious, and its effect on your friends' behaviors may be very extensive.

It has to be added, though, that a simple main effect of "a barbecue situation" on several persons' behaviors is not relevant when it comes to the prediction of behavior by personality scores. If the barbecue situation increased behavior X of every person by the same amount, there would be no need to include information about the situation at all in this case. In other words, you could be sure that, for example, your most extraverted friend would still exhibit the most extraverted behavior as the rank ordering among individuals remains preserved despite mean level differences between situations (Mischel & Shoda, 1995; Sherman et al., 2010). However, this does not allow predictions about the exact level of behavior (e.g., how much time will your extraverted friend spent on the dance floor, and which exact behavior will he or she exhibit?). Thus, focusing on situational effects only makes sense if the exact behavior should be predicted, or if people vary in their response to different situations. In other words, the effects of the situation and the person × situation interactions need to be considered to predict and explain variance in behavior.

As seen in this brief thought experiment, it can be argued that the explanation of behavior can best be achieved by considering the personality of a person, the current state of a person (e.g., affect), as well as the actual situation the person finds him- or herself in. We will come back to

personality and the current state later. The main part of this chapter is devoted to the assessment of the situation and peoples' perceptions thereof.

**Principles of Situation Research**

How could a situation, such as the barbecue, best be described? Why do we all share a common idea about what a barbecue is? First, if we want to describe something, a taxonomy (as a descriptive system) can be helpful.

Rauthmann, Sherman, and Funder (2015) defined three types of situation information that can be used to describe situations: *Cues*, *Characteristics*, and *Classes*. Cues are physical elements of a situation, characteristics refer to the psychologically meaningful features of a situation, and classes describe groups of situations that are similar in their cues or characteristics (Rauthmann & Sherman, 2016b). Each of these types of situation information can come with their own taxonomic systems.

**Situational cues.** Cues are the objectively describable things in any situation, for example, other people or a barbecue, but also the time and the place. Cues constitute the physical setting of a situation, and the manipulation of situational cues is a core paradigm of social psychology (Horstmann, Rauthmann, & Sherman, in press; Krueger, 2009). For example, the number of bystanders (a very easy to manipulate cue) has consequences for helping behavior, which is known as the bystander effect (Darley & Latané, 1968; Latané & Nida, 1981).

One could conclude that information on *actual* cues in a situation is sufficient for situation assessment, which is referred to as the objective tradition to situation taxonomization (Rauthmann, 2015). However, there are several shortcomings to this kind of assessment. First, the effort required to list all cues of a situation is very huge, and there are only very few examples of studies that have tried this (e.g., Barker, 1968). One challenge with this approach is that assigning a psychological meaning that applies to all perceivers of these cues is extremely difficult, if not impossible. For example, whether there are a few or many people in a situation or whether a room is crowded or cozy are differences within the eye of the beholder (Rauthmann, 2012). In other words, while it might seem easy to count the number of people or to measure the room temperature, using these mere facts for psychological research questions is at least difficult. However, it is known that *perceived* cues might have of course a general psychological meaning (e.g., many books in a room indicate that the inhabitant of the room might score high on the trait Openness, see Gosling, Ko, Mannarelli, & Morris, 2002). Note the distinction between actual cues (what is present and could be perceived), perceived cues (what is perceived persons inside or outside the situation), and interpreted cues, which are, in fact, situational characteristics. Future technological advances might make actual situational cues directly available for psychological

research (N. A. Brown, Blake, & Sherman, 2017), for example by showing that a specific cue almost always leads to certain behavior, *independent* of the observer of that particular cue. Until then, assessing situational characteristics as the results of individual information processing (i.e., perceived cues) appears to be the most promising way to assess situational information.

Furthermore, future situations and their constellation are usually unknown, and the application of future situations for the prediction of behavior is therefore very difficult (Hogan, 2009). However, as noted above, the most important argument against such an objective assessment of situations is that the psychological meaning of actual cues is unknown to an independent observer (Rauthmann, Sherman, & Funder, 2015b): Any cue could have nearly any meaning, and thus it is the subjective reality of a situation that matters from a psychological perspective (i.e., when we aim for the prediction and explanation of affect, behavior, cognition, or motivation) and is therefore relevant for a taxonomy of situations. The principle that connects objective cues and subjective situation *characteristics* is called the Processing Principle, which states that only "[p]sychological experiences of situations matter" (Rauthmann, Sherman, & Funder, 2015b, p. 367).

**Situational characteristics.** Accepting the Processing Principle opens the door to a completely different – albeit not new – approach to the assessment of situations: the subjective tradition. The general idea is to ask a person how s/he perceives a situation. Collecting a variety of situation ratings either across people within the same situation or across situations for the same people allows capturing specific characteristics appearing across persons and situations. Yet, if only the subjective experiences matter, how is it possible that we all have probably more or less the same image of a barbecue situation?

Rauthmann and colleagues (2015) give the answer to this in the Reality Principle. Each situation has three realities: A physical reality, which depends on the physical cues present; a consensual reality, which depends on the shared interpretation of situational cues; and lastly an idiosyncratic reality, which is each perceiver's individual and unique interpretation of the situation. It is the consensual reality that reflects the shared understanding and interpretation of the same situation across participants. For example, most of us think that a barbecue is a positive situation, whereas cleaning up the kitchen afterwards is not. Yet, some – for example, very introverted persons – may think the opposite and prefer (i.e., evaluate as more positive) most other situations over a barbecue.

The latter example alludes to another issue in psychological situation research: people's perceptions of their situations and their momentary affect in them will likely be closely linked, though different from each other (Horstmann & Ziegler, 2018). The confound between the

perceiver and the situation that arises when a situation is defined via the perceiver is called the Circularity Principle. It follows from this principle that only more than one perceiver can approximate the *true* nature of any given situation, at best from different perspectives. Ratings from persons *in situ* (within the situation), ratings from bystanders *juxta situm* (around the situation), and ratings from people *ex situ* (outside of the situation, for example based on videos or other records) can be used to triangulate the consensual characteristics of a common situation (Rauthmann, Sherman, & Funder, 2015b).To clarify, we will extend the above example and assume someone got seriously harmed or killed during the barbecue. To get the best picture of the barbecue, the police would ask all participants, thereby obtaining ratings from people *in situ* as well as from all possible observers *juxta situm* who were not directly involved, such as neighbors. Finally, they could collect expert ratings of barbecue situations (e.g., an expert assessment on the safety of gas containers), rate accounts of the participants, or assess data from other sources (e.g., videos from the scene), thereby obtaining different *ex situ* ratings and data. Via this variety of approaches to assessing the situation, it would be possible to distill the commonly shared notions of a barbecue and separate the specific cues and characteristics of the one specific barbecue situation in question. This approach is similar to the idea of other-ratings in personality assessment: Systematic yet unique biases from individual sources should be cancelled out across several raters, which thereby offers a better picture of the rated person (Hofstee, 1994).

**Situational classes.** Classes of situations refer to situations that are similar in their cues or perceived characteristics. For example, it is very likely that the class "at work" differs substantially from the class "at home" – at least for most people. Still, there can also be more specific and very narrow types of situations (e.g., specific crime scenes or surgeries).

Despite the intuitive appeal of classes and their ease of efficient communication, the assessment of classes makes their application difficult. If a participant is asked "in which situation(nal class) are you currently?," the selection of the class (or category) most likely depends on the perceived characteristics of the situation (Rauthmann & Sherman, 2016b). A barbecue situation may be fun for everyone who is invited and thus fall into one class, for example, "leisure." But for the employee who meets his boss, the situation might be of another class, for example, "work." These people would categorize the barbecue situation into different classes because they construe the situation differently. This information about the characteristics of the situation would be lost when using situational classes only and choosing the same situational class for all persons in this situation. This is, however, not to say that classes cannot be very valuable (for an application, see for example Geukes et al., 2017; Oud et al., 2017). A robust but also nuanced taxonomy of

situational classes as well as a means for their assessment would indeed be highly desirable, as it would allow comparing many studies that have worked or strive to work with situational classes.

**Cues, characteristics, or classes?** So far, most progress was made on the identification, description, and measurement of situational characteristics (for a review, see Horstmann, Rauthmann, et al., 2018). Recently, several taxonomies based on situational perception were developed. If perceptions of situational characteristics are to be used in situation research, situational characteristics have to be measured or assessed. The next sections describe how this can be achieved.

## Constructing Measures of Situational Perception

The construction of a tool for the assessment of situational perception is generally similar to the assessment of individual difference characteristics of a person. However, there are notable differences that can have important consequences for the resulting measurement tool or situational taxonomy. Throughout the remainder of this chapter, we will therefore first review general recommendations made for the construction of any questionnaire and then go into the specifics for the construction of assessment tools for situational perceptions. Finally, we will present a recently developed taxonomy for the assessment of situational perceptions. We then conclude the chapter with recommendations and outlines for future research.

### The ABC of Test Construction

The ABC of test construction is a set of three questions proposed by Ziegler (2014c). If a psychological construct (or latent, not directly observable variable) is observed, then an answer to the following three questions provides clarity about the measurement of said construct:

A. What is the construct being measured?

B. What is the intended use of the measure?

C. What is the targeted population?

Answers to these questions help test constructors and test users better understand the specific tool in question, its application, and its limitations. Thus, with reference to Question A, defining the construct(s) to be measured as well as their nomological network (Cronbach & Meehl, 1955) not only help to understand and interpret the actual test score, but also inform item construction and the validation strategy of the resulting measurement tool. Defining how the construct(s) relate(s) to other constructs aids selecting tools for the examination of convergent and mostly discriminant as well as evidence for test-related criterion validity.

The answer to Question B, the intended use of the measure, mostly decides which test-criterion relations are important. A test claiming to be useful for personnel selection should provide some empirical evidence showing its relation to job success. Tests that should be used in

scientific settings should show good discriminant validity (Horstmann & Ziegler, 2018), and tests that are used in experience sampling should have a reasonable number of items (Horstmann & Rauthmann, in preparation).

Finally, and related to Question C, defining the target population means to define in which aspects samples for construction and possibly the norm-population of the sample should vary. Furthermore, for the construction of questionnaires of situational perception, not only the population of participants, but also the population of the measured situations is relevant. Both aspects, persons and situations, are important for the construction of the items and the applicability and generalizability of the scores derived from the measure.

To sum up, the ABC of test construction is a guideline specifying important steps on the way to a psychometrically sound test. Likewise, test users can use the information contained in the answers to these three questions to form an opinion on the quality of a specific test.

**Road Map to Situational Taxonomization**

For the assessment of situational information, detailed answers to these three main questions can be found by following Rauthmann's (2015) road map to situational taxonomization. This road map consists of six steps, which can be directly related to the questions above, specifically to Question A, the construct that is being measured. The construct that is measured is defined during the construction process of the taxonomy. Answers to Question B and C strongly depend on the decisions made during the initial construction process. The six steps proposed by Rauthmann are 'multiplicative forks' – this means that each decision taken can be combined with almost any (but not necessarily all) of the decisions made at the other steps, thus leading to a plethora of possible situational taxonomies with different underlying assumptions, properties, and areas of application. Making explicit statements about which step was taken at the assessment of situational information is therefore important. The six steps presented by Rauthmann (2015, p. 177) are:

1. Theory: Which theory underlies the construction of the situational taxonomy?
2. Information: Which situational information (e.g., cues, characteristics, classes) is investigated?
3. Aims: How should the situational information be structured?
4. Generation/Pool: How is the situational information obtained or sampled?
5. Assessment: How is the situational information assessed and judged?
6. Analysis: How is the obtained information analyzed and structured?

All of these steps have an influence on the construct that is measured, and thus, on the measurement tool and the population that is targeted. To exemplify the effects different decisions can have at each step, we will briefly introduce and review three different and recently developed situational taxonomies: The Situational Affordances for Adaptive Problems (SAAP; N. A. Brown et al., 2015), the Situational Eight DIAMONDS (Rauthmann et al., 2014), and the Situation Five (Ziegler, 2014a). An overview is presented in Table 2.2.1.

Table 2.2.1
*Steps involved in the construction of situational taxonomies and their assessment tools*

| **A: What is the construct being measured?** | | | | |
|---|---|---|---|---|
| **Road Map** | **Relevant question** | **SAAP** | **DIAMONDS** | **Situation Five** |
| Theory | Which theories and concepts underlie the research? | Evolutionary Theory | No explicit theory | Lexical Approach |
| Information | Which situational information is the focus? | Characteristics, content dimensions | Characteristics, content dimensions | Characteristics, mainly content dimensions |
| Aims | How is the situational information structured? | Hybrid of inter- and intra-situational | Inter-situational | Hybrid of inter- and intra-situational |
| Generation | How is the situational information generated? | Post-Hoc generated by laypersons, later in the construction by researchers | Post-Hoc generated by laypersons | Post-Hoc generated by laypersons, later in the construction by researchers and experts |
| Assessment | How is the situational information presented and judged? | Raters that were in situ, recollected information, rated in on a-priori defined dimensions | Raters that were in situ, recollected information, rated situations on items derived from the RSQ | Raters that were in situ, recollected information, rated these on adjectives extracted from the lexical approach |
| Analysis | How is the situational information analyzed? | Factor analysis | Factor analysis | Factor analysis |
| Uses | How is the situational taxonomy applied? | Mainly research | Mainly research | Personnel selection |
| **B: What is the intended use of the measure?** | | **SAAP Questionnaire** | **S8-I, II, III** | **B5PS** |
| Intended Areas of Use | | Rate evolutionary relevant situations on SAAP dimensions | Rate nearly all daily situations on DIAMONDS dimensions | Obtain general tendency of persons to perceive different situations (situational perception trait) |
| **C: What is the targeted population?** | | **SAAP Questionnaire** | **S8-I, II, III** | **B5PS** |
| Persons[a] | | For some dimensions people at all ages, for other only people in a certain period of their lives (e.g., spouses, parents, sexually active persons, etc.) | General population | Persons qualified and experienced in the job-market, especially white-collar workers |
| Situations | | All regular and day-to-day situations | All regular and day-to-day situations | Taxonomy describes any situation, but measurement tool is focused on hypothetical situations from an occupational context |

*Note.* a = all measures have been validated for gown-ups only. So far, no versions for children exist at all; SAAP = Situational Affordances and Adaptive Problems (N. A. Brown et al., 2015); DIAMONDS = The Situational Eight DIAMONDS (Rauthmann et al., 2014); S-I, II, III = Brief- and ultra-brief measures for the DIAMONDS (Rauthmann & Sherman, 2016a, 2016c); B5PS = The Big Five Inventory of Personality in Occupational Situations (Ziegler et al., submitted).

**Underlying theory.** The construction of the SAAP relied on evolutionary theory, which suggests that during human evolution at least seven broad adaptive problems had to be overcome in everyday life of human beings. According to Brown and colleagues (N. A. Brown et al., 2015), these seven adaptive problems are Self-Protection (obtaining physical protection), Disease Avoidance (avoiding sick people and getting ill), Affiliation (getting along with others and cooperating), Status (being respected), Mate Seeking (finding a suitable mate), Mate Retention (keeping a mate, including making the mate happy and holding on to a relationship), and Kin Care (caring for relatives, such as children or siblings). Addressing these challenges has shaped how situations are perceived and thus should allow describing everyday situations along these dimensions.

For the Situational Eight DIAMONDS, no explicit theory was used. However, the authors reviewed existing literature on situational perception and predicted a number of corresponding dimensions that they would expect (Rauthmann et al., 2014).

The Situation Five are based on a lexical approach. The lexical hypothesis assumes that important and distinguishable characteristics of objects found their way into human language (Allport & Odbert, 1936; Klages, 1926). It has also been used in the construction of other situational taxonomies (Parrigon, 2017; Parrigon et al., 2017). Analyzing and structuring adjectives will thus allow a construction of situational dimensions that are important to most humans. This means such dimensions may not necessarily be exhaustive (e.g., they might not apply to extraordinary, rare situations) but are for the same reason more applicable across wider set of everyday situations.

**Situational information.** Situational information refers to the question which situational information (cues, characteristics, classes) is at the focus of the taxonomy. All three situational taxonomies focus on situational characteristics at the core of their description (Horstmann, Rauthmann, et al., 2018). Furthermore, Rauthmann (2015) also distinguished content and style characteristics. Style characteristics describe in an abstract manner "how" a situation is (e.g., active or passive, formal or informal). Content characteristics describe "what a situation is about", for example, whether something needs to be done (e.g., Duty from the DIAMONDS taxonomy). SAAP, DIAMONDS, and the Situation Five do not make explicit statements about these two types of characteristics, yet they focus mainly on content dimensions. Only the Situation 5 (and most recently, another situational taxonomy, the CAPTION by Parrigon et al., 2017) includes a style-dimension: Lack of Stimuli (i.e., something is normal, dull, monotone).

**Aims.** In deriving a taxonomy, one can aim for a structure of situational information that is either intra-situational or inter-situational. Intra-situational means that information from the

same or very similar situations is clustered, for example, when one wants to find relations between situational cues within the same (class of) situation. Inter-situational means that different situations are sampled (for the sampling of situations, see next step), which is directly relevant to the B question "What is the intended use of the measure?" If the taxonomy should be generalizable to many situations, then it is necessary to sample many situations and compare many different situations during the process of its development. On the other hand, if the taxonomy should be used to describe only very few, specific situations, then the intra-situational approach should be preferred. Notably, an intra-situational structuring corresponds to a more idiographic, whereas an inter-situational one to a more nomothetic approach.

The DIAMONDS taxonomy clearly identifies as an inter-situational taxonomy since it was constructed on inter-situational comparisons. Indeed, several applications show that the application of the DIAMONDS is feasible across many research contexts (Horstmann & Ziegler, 2018; Rauthmann & Sherman, 2016c; Sherman et al., 2015). The SAAP and the Situation Five are, on the other hand, a hybrid of inter- and intra-situational comparisons. Both allow inter-situational comparisons: By using self-reported, random everyday situations as a situation pool they automatically apply to a wider context of situations than a taxonomy that is based on a fixed set of situations. However, their theory (evolutionary theory: SAAP) or targeting towards a specific context (occupational settings: B5PS) bends them towards intra-situational comparisons, with a narrower focus on comparisons of the same or similar situations.

**Generation and sampling of situational information.** The sampling of the situations (situational information or situational material) is particularly relevant to the intended use of the measure. At this stage, two different steps have to be completed.

First, items on which situations are rated have to be generated. There are at least five different methods to generate items, and these are of course not independent: Information may be generated based on researchers' expertise, laypersons' nominations, theory, existing lists of stimuli, or adjectives from lexica.

The SAAP, for example, used a theory-based generation of items. The authors generated a set of questions that could represent the theoretically assumed dimensions (N. A. Brown et al., 2015). For the construction of the Situation Five, a German dictionary was scanned and all adjectives (around 15,000) were extracted. Because 15,000 adjectives are too many to apply within any sensible study, these adjectives were reduced to about 300 adjectives useful to describe situations (Ziegler, 2014a). The DIAMONDS used existing items from the Riverside Situational Q-Sort (RSQ; Wagerman & Funder, 2009). The RSQ is based on the CAQ, a tool initially developed by clinical practitioners to describe their patients (Block, 1978; Funder, 1997).

Second, situations or situational material that should be rated has to be sampled (or generated). During the construction of a personality questionnaire, this second step would be comparable to the selection of the participants *that are rated*. In the context of situations, it therefore means which situations are used that should later be presented to and rated by perceivers. Similar to the effects the construction- or norm-sample has on the psychometric properties of a personality questionnaire, the selection/generation of situational information can have strong implications for the resulting situation taxonomy. During the construction process of the B5PS (the assessment tool for the Situation Five), human resource managers generated information about critical situations of their workplace (critical incident technique: Flanagan, 1954). For the construction of the DIAMONDS and SAAP (but also for the Situation Five at an earlier stage), participants were asked to "rate a situation from their everyday life," which means participants were "generating" the situational material ad hoc by remembering it. Note that sampling situations by asking participants about a situation they remember may introduce a bias, as certain extreme situations may not be reported. Even though this step may not receive much attention during the construction process, it surely has consequences for the applicability and generalizability of the resulting taxonomy and its assessment tool. To clarify, when constructing a personality taxonomy, selecting the persons or behaviors targeted will define the degree of generalizability and most likely affect the number of resulting domains as well as their breadth. In the same way, selecting only certain situations during the process of constructing a taxonomy of situational perception will potentially limit the number of resulting dimensions as well as their heterogeneity. Thus, in the same way a homogeneous rater sample restricts variance, a homogeneous set of situations will also affect the validity and generalizability of the resulting taxonomy.

**Assessment.** The next step during the definition of a situational taxonomy has probably the biggest impact on the final result, namely the sampling of raters, as well as the presentation and rating of situational material. To put things into perspective, the usual procedure at this stage in personality research would be the recruitment of participants (i.e., sampling of raters), the generation of items on which the situations are rated, the presentation of the item to the raters, the introspection of the rater ("Who am I?"), and the answer on a personality item (i.e., judgment).

At this stage, several decisions have to be made. The first decision involves the selection of the raters: Who rates the situational information? For example, who answers the items that were generated in the previous step? The selection of raters or judges needs to be considered when applying the taxonomy or resulting tool later and is of course in some cases not independent from the sampling of situational material.

As mentioned above, three different types of raters exist. Relative to the situation that is judged, raters can be *in situ* (within the situation), *juxta situm* (around the situation), or *ex situ* (outside the situation). Raters *in situ* take actively part in the situation. Raters *juxta situm*, such as bystanders, do not have an active part in it. Finally, raters *ex situ* access accounts of a situation that they themselves were not involved with. Each of these raters has access to different information about the situation, which could be compared to the self-other-knowledge asymmetry in personality psychology (Vazire, 2010): Neither rater has better or worse access to the situation, but each has a different and in parts unique perspective.

Note that the rater and the situational material to be rated can be conflated: A rater who is *in situ* and reports their self-experienced situation (i.e., generates the situational material to be rated; see previous step) is usually also required to rate this material. However, this does not necessarily have to be the case, and one could imagine that a first person generates (recalls) a situation and a second one rates this situational material, thereby separating rater from perceiver variance.

Second, the medium of presentation has to be decided on. This can be a recollection of previous everyday situations (the "medium" here would be thoughts, imagination). This was used in the construction of all three taxonomies SAAP, DIAMONDS, and Situation Five) or the presenting of vignettes or hypothetical situations, which was both used in a later construction phase of the SAAP and Situation Five, respectively. Again, the medium of presentation from the construction phase should resemble that in the application. Note that this step is different from the sampling of situational information, even though these two steps are usually very similar: Sampling refers to the selection of situations, presentation means "how are they presented to the participant?" In some cases, these two steps will be nearly indistinguishable, but they can be different in other cases.

The third decision concerns the exposure to situations – how is the situational material (e.g., the account of the situation or the actual cues of the situation) presented *to the judge* or perceiver? This could be *in vivo* (while being at the situation, which is only applicable for certain types of raters: *in situ* and *juxta situm*), but could also include written accounts of the situation. The B5PS and the SAAP, for example, used situational vignettes (i.e., written accounts of hypothetical situations) in the latter stage of the validation. Other methods, such as recordings of situations (e.g.,picutres taken in daily life: Brown et al., 2017; acoustic snapshots: Mehl, Vazire, Ramirez-Esparza, Slatcher, & Pennebaker, 2007) could also be presented to all three types of judges.

The fourth decision concerns the judgment dimensions on which the situations that are selected are rated. Some recent taxonomies have focused on an adjective-like assessment of the situation (Horstmann et al., 2017). However, there are other taxonomies that have, for example,

focused on the assessment of behavior or trait-expressions that are evoked by a specific situation (Ten Berge & De Raad, 2001, 2002). Other dimensions on which situations could be rated are the appearance of the physical environments (e.g., rough, wet, cold), mental processes of the persons in the situation (e.g., this situation makes me think about food), or other's behavior (e.g., this situation makes people in general aggressive).

The criteria on which situations are judged is the next decision that has to be made. Criteria could be, for example, the similarity of one situation to another situation or to a typical situation; the frequency of occurrence of a situation; or expectancies someone has about a situation. Most recent taxonomies focus on the psychological characteristics of situations (e.g., this situation requires that work is being done) and their valence (how positive or negative is a situation?).

The last decision that has to be made concerns the way the judgment is delivered, that is, the judgment method. The judgment could be quantitative, and thus, for example, a Q-Sort ranking method (e.g., characteristic A describes the situation better than characteristic B) or a Likert-type rating (e.g., characteristic A describes the situation very well with a "4" on a scale from 1 to 5). On the other hand, it could be a free response format, in which a person verbally describes his or her judgment. Whatever method is selected to obtain the judgments, it should reflect the purpose of the taxonomy and also be close to the later assessment procedure. Obtaining qualitative judgments of situations can be very informative, but a systematic and objective scoring method for future assessments needs to be available as well. For the construction of the SAAP, Situation Five, and DIAMONDS taxonomy, all these decisions were made, whether explicitly or implicitly, and all tools subsequently built require rating the situation on a Likert-type response scale.

**Analysis of situational information.** The analysis of the gathered information has a strong impact on the resulting scale and thus on every aspect of the ABC of test construction. Rauthmann (2015) lists three commonly used main approaches: Factor analyses, cluster analyses, and multidimensional scaling. Each of them results in a different interpretation of the results. Factor analyses and multidimensional scaling will yield a reduced number of underlying dimensions that capture most or all of the systematic variance in judgments to the items. Cluster analyses will result in hierarchical dimensions, or, depending on the kind of analysis, in situational classes.

**Uses.** A seventh and additional final step mentioned by Rauthmann (2015) is the future use of the scale. This is not only relevant for the application of the scale itself, but also for its validation (Rauthmann, 2015). If a situational taxonomy is mainly applied in a predictive context, validation via prediction should be the key focus. However, if the taxonomy is used to describe situations, a validation should focus on the correlation with other descriptive taxonomies. Even though this step is the last of the decisions that has to be made, it should be considered before

starting with the construction process at all because the use of the measure will substantially inform all other steps required (Ziegler, 2014c).

**Jingle-Jangle Fallacies and the Nomological Net**

Considering all these different paths through the road map to a situational taxonomy, it may well happen that researchers end up with one of two following cases: On the one hand, a construct may be "discovered" or defined that already exists but then given a different name. On the other hand, a construct may be named similarly or equally to one that already exists even though it is quite different. This confusion of constructs is called the jingle-jangle fallacy (T. L. Kelley, 1927). One example of this fallacy are the dimensions Negativity (*Situation entails or could entail stress or trauma*) and Adversity (*Another person is under threat*) from the DIAMONDS (Rauthmann et al., 2014) versus Negative Valence (*malicious, repulsive*) and Adversity (*stressful, frustrating*) from the CAPTION model (Parrigon et al., 2017). In the DIAMONDS, Negativity is what Adversity is in the CAPTION model.

To avoid jingle-jangle fallacies and guide construct validation – especially convergent and discriminant validity –, a clear definition of each construct's nomological net is required. A nomological network describes how psychological constructs are related to each other (Cronbach & Meehl, 1955). According to Cronbach and Meehl, these relations (or 'laws', p. 290) can be threefold. They can either connect observable properties to each other (such as a correlation between a construct as measured and relevant behavior), can relate a theoretical construct to an observation, or two theoretical constructs to each other. To make the nomological net visible, at least some nodes of the network have to be observable, that is, must allow linking the theoretical construct to the 'real' world. Cronbach and Meehl (1955) acknowledged that when a construct is fairly new (as is the case of recent dimensions of situation perception), "there may be few specifiable associations by which to pin down the concept. As research proceeds, the construct sends out roots in many directions, which attach it to more and more facts or other constructs" (p. 291). For the definition of the construct, this 'pinning down' is essential, especially in the early stages of construct-, taxonomy-, and test-development.

The first extensive nomological net of situation perception dimensions was recently presented by works from Horstmann, Rauthmann, and Sherman (Horstmann, Rauthmann, et al., 2018; Rauthmann & Horstmann, 2017; Rauthmann & Sherman, 2018a). Based on empirical findings, especially convergent and discriminant correlations, but also based on theoretical considerations (e.g., similar item-content, similar description of the constructs), six dimensions of situational perception were proposed. However, especially empirically established links between

situational perception dimensions are still scarce and thus, we need more studies looking at several taxonomies at once.

The definition and establishment of a nomological net has further advantages, not only for the understanding of the construct. The development of items to assess information on situational perception and the final item-selection also profits from a nomological net. It should, for example, be possible to locate each item within the nomological net, such that it captures the core of the construct (Ziegler, 2014c). Items that are accurately located in the nomological net and do not overlap with other constructs are more likely to lead to a unidimensional score (Ziegler & Bäckström, 2016; Ziegler & Hagemann, 2015), thus, increased model fit and potentially better reliability and validity.

**Concluding Thoughts – Using the Full Road Map**

Considering all the different possibilities presented by Rauthmann (2015), it is not unlikely that each combination of decisions made will lead to a different situational taxonomy, and none is necessarily better or worse than the other. Nevertheless, even different approaches have led to strikingly similar situational taxonomies (Horstmann, Rauthmann, et al., 2018). Despite this, it cannot be concluded that all approaches will ultimately lead to similar situational taxonomies. This might be compared to the difference between self- and other-reports of personality. Even though self-reports were considered the gold-standard of personality assessment, the comparison of self- and other-reports of personality provided new and valuable insights into the nature of personality constructs (e.g., Vazire & Mehl, 2008). We think it is therefore necessary to continue with the development of situational taxonomies and evaluate their respective convergences and divergences. Each new taxonomy as well as each new corresponding assessment tool will increase the knowledge about situational perceptions *if* these new taxonomies and their measures are compared and related to existing ones. This does of course not mean that we suggest a never-ending debate on which taxonomy is better or describes the world more accurately. Such a fruitless debate does not really advance the field. However, when proposing new taxonomies, the relation of each new taxonomy to existing ones should always be considered as already stated above. The aim is to keep the number of constructs and taxonomies as large as necessary and as small as possible.

## From Taxonomy to Measurement

A taxonomy and an assessment tool are in principle independent from each other. Each can exist without the other, although this is not very often the case. For example, the RSQ was a tool for the assessment of situations and existed without a clear taxonomy of situations. However, the DIAMONDS were based on data from the RSQ. The DIAMONDS taxonomy then went on to spawn own, specific assessment tools (Rauthmann & Sherman, 2016a, 2016c). On the other hand, several situational taxonomies exist that have not presented an assessment tool (Horstmann, Rauthmann, et al., 2018). Of course, the taxonomy has an influence on the assessment tool, and in most cases, the assessment tool is a "by-product" of the development of the taxonomy. If, for example, a taxonomy is developed based on the lexical approach and adjectives were used to rate situations, these adjectives could be used in a subsequent step for the development of a situational assessment tool (Parrigon et al., 2017; Ziegler et al., submitted). In the following, we will provide one example of a test developed to measure personality and situational perception at the same time. The example also includes the preceding development of a situational taxonomy.

### A Joint Taxonomization of Persons and Situations

The Situation Five are a taxonomy of situational perception dimensions that were tailored to be used in occupational situations (Ziegler, 2014a; Ziegler et al., submitted). This taxonomy is based on the lexical approach (see above), and the final measurement tool (Big Five Inventory for Personality in Occupational Situations: B5PS) aims at assessing the general tendency of a person to perceive and interpret situations. Thus, it aims at assessing a stable characteristic of the person and thus a personality trait (Funder, 2001). As argued earlier, constructing situational perception as stable characteristics of a person is reasonable when the future situations are unknown (Hogan, 2009).

**Development of the Situation Five.** The Situation Five were developed based on adjectives from a German dictionary, the Duden (Dudenredaktion, 2006). The authors extracted all adjectives (about 15,600). Raters then reduced this list to about 300 by excluding adjectives unsuited to describe situations (e.g., *schief*, meaning skewed), redundant (e.g., *schmerzhaft* and *schmerzvoll*, both meaning painful), rarely used (e.g., *abecelich,* a strange word for alphabetical), or outdated adjectives (e.g., *schnäkig*, being picky in the selection of food). Participants were then invited to rate self-reported situations from their everyday lives using these adjectives. Exploratory and confirmatory factor analyses were applied to identify and confirm five dimensions of situational perception across two independent samples: Outcome-Expectancy, Briskness, Psychological and Physical Load, Lack of Stimuli, and Cognitive Load (Ziegler et al., submitted). An additional dimension, Weather, was also found. However, due to the intended purpose of the

final taxonomy (personnel selection), this dimension was removed. The resulting taxonomy was then used to develop a test capturing personality and situational perception at the same time.

**Construction of the B5PS.** The B5PS is a questionnaire that assesses self-reported personality traits (the Big Five) as well as the Situation Five just described. The test was developed for use in personnel selection and development where traits are more suited than states. However, the general idea could also be applied to a clinical setting in which participants describe their general perception of situations (Ziegler et al., submitted). Accordingly, the definition of the Situation Five as stable traits was a good starting point. The key to connect personality and situational perception assessment was a special kind of item capturing both aspects in specific situations. First, based on interviews with human resource managers, hypothetical situations (vignettes) were developed that describe everyday situations in occupational settings (e.g., "just before finishing time, your supervisor assigns you a task that has to be finished the same day"[10]). After reading such a statement, participants are first required to assess how they perceive the situation (e.g., "I assess the situation as challenging" capturing Cognitive Load). Next, participants are asked how they would behave in that situation (e.g., "I keep calm" capturing Emotional Stability). Participants respond to 211 of these items, whereby scores of the Big Five, 42 facets, and the Situation Five are obtained.

Since each of the Situation Five scores for each person is based on the same set of hypothetical situations (vignettes), participants' scores can be compared inter-individually. The Situation Five thus represent a different yet valuable approach to the assessment of situational perception, which allows the use of the B5PS scores for purposes such as personnel selection. Using random situations (varying between participants) and assessing situational perceptions *in situ* or using external raters would be unsuited for such purposes. Moreover, interindividual comparisons would be problematic due to the different situations used. This advantage comes along with a limitation, though. The population of situations in the B5PS is fixed, so applicability of the assessment tool outside of occupational settings might be limited (even though the taxonomy itself can be used outside of occupational settings as it is based on self-reported everyday situations). The sample used for the construction and validation of the final scales was a representative group of the German-speaking populations (from Switzerland, Austria, and Germany), and the measure can therefore be used across a wide range of non-clinical participants. There is evidence supporting the test scores' reliability, stability, and validity (Ziegler, 2014a; Ziegler et al., submitted).

[10] Please note that this describes the generation of situational material that is rated.

Another interesting feature of the test is that it is possible to select situations (vignettes) from the test that are perceived by most people in the construction sample in a similar way. One could select situations that were assessed as highly outcome-oriented or very boring by most participants in the norm sample. It is then possible to analyze how the Big Five domain scores of an individual person change when only ratings from these specific situations are included in the scoring. For example, it could be an interesting question how a person describes her/himself in a situation that was assessed as stressful by most participants in the norm group. This way it is possible to gauge the variability of a person with regard to changing situational demands, which is essentially an examination of a person-situation interaction. To sum up, the B5PS is a measure of personality, meaning the Big Five as well as the Situation Five, thereby offering unique possibilities in applied settings.

Table 2.2.2
*The Situation Five dimensions of situational perception*

| **Dimension** | **Description: Situations that …** | **Sample Item** |
|---|---|---|
| Outcome-Expectancy | have potential, and may result in a positive outcome | *I assess the situation as full of potential.* |
| Briskness | are stimulating and lively | *I assess the situation as lively.* |
| Psychological and Physical Load | are stressful, burdening and mentally challenging | *I assess the situation as burdensome.* |
| Lack of Stimuli | are boring, lack input | *I assess the situation as boring.* |
| Cognitive Load | are challenging, demand thinking and analysis | *I assess the situation as challenging.* |

**Specific Challenges for Measures of Situational Perception**

If a situation is perceived and subsequently rated on an item, two systematic sources of variance will occur. On the one hand, there is the commonly shared perception of a situation which can be captured to some extent by nomothetic taxonomies as described above. Assuming that people use the dimensions captured in those taxonomies to describe situations, but consistently vary in the degree with which they ascribe the level of one particular dimension to a specific situation, systematic variance in situational perceptions should be observable. This systematic variance would then be reflected in the estimates of reliabilities or tests of factorial validity.

On the other hand, there is the idiosyncratic view of a person on each situation also creating variance (Rauthmann, 2012), which reflects the effect the person × situation interaction. It captures the unique but systematic perception of a specific situation by a specific person. Of

course, there is also unsystematic measurement error. Importantly, there are two sources of systematic variance: the commonly shared view and the idiosyncratic view. This theoretically sound conclusion poses a big problem for the psychometric evaluation of measures capturing situational perception (Horstmann & Ziegler, 2016; Ziegler & Horstmann, 2015), which we elaborate below.

**Reliability.** The reliability of a test score is usually either estimated in terms of internal consistency or stability (if the construct is believed to be temporally stable). Cronbach's alpha (Cronbach, 1951) probably still is the most widely used estimator of internal consistency. However, it has been criticized and other estimators such as McDonald's Omega have been proposed (Sijtsma, 2009; Ziegler & Brunner, 2016). When it comes to situational perception, though, the fact that its variance consists of two systematic sources poses a big problem. Strictly speaking, Cronbach's alpha should not be used at all because it requires tau-equivalent items or at least congeneric items. In any case, tau-equivalent items or at least congeneric items would mean (and the use of Cronbach's alpha would imply) that the items are unidimensional, that is, responses to the items are influenced by one systematic source of variance only. As outlined above, this is not the case with situational perception due to the two systematic variance sources. Specifically, the idiosyncratic variance presents a big challenge for reliability estimates. This variance is unique to each person and therefore does not contribute to the correlations between items across persons. Yet, it contributes systematically to the items' variance. The most important aspect of Cronbach's alpha is the difference between one and the ratio between the sum of item variances and the variance of the item sum. The latter also includes the item covariations and therefore surpasses the simple sum of item variances if the items are unidimensional and reliable. Consequently, the ratio decreases and gets closer to zero. The difference between one and the ratio remains large, yielding a large reliability estimate. The idiosyncratic variance, however, is captured in both (i.e., sum of item variances and variance of the item sum) and therefore yields a ratio closer to one and thus also a difference from one closer to zero (e.g., a bad reliability estimate). A similar problem occurs for McDonald's Omega. This estimator uses factor loadings derived from factor analysis. Here, the same problem arises: The idiosyncratic variance cannot be explained by a factor representing the variance shared by all items. This variance can be found in the residuals. Consequently, the loadings do not represent all systematic variance of the items. Again, the estimate is most likely too small and underestimates the true reliability. In sum, neither approach seems suited to estimate the complete amount of systematic variance. Instead, such estimators should be regarded as lower bounds of reliability.

The issues with internal consistency estimations imply that an estimator of stability might be better suited if the situation does not change from one measurement occasion to another

(McCrae, Kurtz, Yamagata, & Terracciano, 2011). However, a different problem arises here. It is difficult to experience the same situation twice as if it were new each time (Horstmann et al., in preparation). Yet, this is the basic idea of stability estimates. Otherwise, memory effects or something similar could distort the estimate. Using vignettes with hypothetical situations might help as these can be carefully constructed to be functionally identical. Still, estimations of reliability focusing on stability will likely not be fruitful going forward in situation research.

Finally, it has to be noted that the problem described above relates to the assessment of one situation by multiple raters, where each rater perceives the same situation. In an applied context (e.g., experience sampling), this would not be the case, and unique variance of each situations contributes to the variance of each item. Thus, the problem described above is even more pronounced.

What can be done? If each perceiver sees a different, unique set of situations, it means that these situations are nested within persons. Such a nested data structure across time allows using more sophisticated approaches to reliability estimation, such as latent state trait theory (Steyer et al., 2015; Ziegler, Ehrlenspiel, & Brand, 2009), multilevel modeling (e.g., Nezlek, 2017; Sherman et al., 2015), or generalizability theory (Brennan, 1992; Cronbach, Gleser, Nanda, & Rajaratnam, 1972; Ziegler, Poropat, & Mell, 2014). Whereas latent state trait theory does not require many measurement occasions, it allows estimating variance systematically occurring across situations and within each situation. Both can be combined to an estimate of reliability. Finally, generalizability theory allows the specification of each source of systematic variance and estimation of its contribution to a test score interpretation. Thus, a better estimate of all systematic variance sources might be possible. It is beyond the scope of this chapter to delve deeper into these issues, but test developers are warned that classical approaches to estimating reliability might be ill-suited for measures of situational perception.

**Validity.** The same problems affecting reliability estimates also affect some validity indicators. Starting with factorial validity, it should be clear that systematic variance which cannot be accounted for by latent variables (i.e., idiosyncratic variance) lowers model fit in structural equation modeling. Thus, typical model fit indicators might indicate ill-fitting models (when evaluated with the cut-offs suggested by Hu & Bentler, 1999), but there is no way to further improve the model. This should be viewed as evidence for idiosyncratic variance (reflected in unaccounted but unique variance per item and ultimately in a lower value for the Comparative Fit Index (CFI) in confirmatory factor analyses). This should, in fact, be embraced and not avoided, as it points to idiosyncratic situation variance that is reflected in the items. This effect, which is

common in measures of personality (Heene, Hilbert, Draxler, Ziegler, & Bühner, 2011; Hopwood & Donnellan, 2010), is therefore even more pronounced in measures of situational perception.

Similarly, idiosyncratic variance is problematic when it comes to test-criterion correlations, such as, for example, the prediction of behavior. Here, the level of symmetry is very important (Wittmann & Süß, 1999; Ziegler & Brunner, 2016). Aggregating situational perceptions across different situations reduces the influence of idiosyncratic situation variance and maximizes systematic, commonly shared variance. However, such an aggregate represents situational perceptions shared by many people across a variety of situations. Thus, the criterion needs to be equally abstract. For the B5PS, for example, work engagement was used as one criterion hoping to establish a similar level of symmetry. Nevertheless, oftentimes specific behaviors in specific situations are used as criteria. As was the case in the barbecue example, we might be interested in using situational perception scores to explain behavior within a specific situation. Here, it is vital that idiosyncratic variance is utilized, but the criterion measure (i.e., the actual behavior) needs to be susceptible to these idiosyncratic perceptions. Such criterion relevance needs to be considered just as much as criterion contamination (i.e., variance not representative of behavioral differences, such as rater effects) and criterion deficiency (i.e., behavioral differences not captured, such as lacking content validity of criterion measure; Brogden & Taylor, 1950). Thus, judging test-criterion correlations simply by their size is bound to lead to erroneous conclusions. It is necessary to carefully consider the level of aggregation as well as the make-up of the criterion and the predictor.

## General Conclusion

The current chapter described one approach to constructing taxonomies of situational perceptions. Using the examples of the DIAMONDS, SAAP, and Situation Five taxonomies, important decisions to be made in the construction process of a situational taxonomy and associated measures were outlined. At the end of such a development process, there often is not only a taxonomy but also a measurement tool that can be used for further research or applied questions (e.g., personnel selection). The B5PS was described as one such tool.

Most importantly, no existing taxonomy alone can claim to be applicable to all possible situations. Further, despite some differences, there are striking similarities between the taxonomies. Future projects looking to expand the nomological net and establish new dimensions of situational perception should try to avoid jingle-jangle fallacies by paying close attention to what already exists and choosing labels wisely.

One aspect that has not been paid much attention to so far is the degree of detail with regard to a hierarchy. It might be possible that situational perception is just as facetted as personality itself. This brings up another challenge not mentioned thus far. Situational perception

is a result of an individual and yet in part commonly shared interpretation of situational cues. The intricate processes that actually lie behind these shared interpretations and their relation to personality taxonomies should be explored in more depths.

**Outlook**

Going back to our initial example of the barbecue situation, we see that there is more than just the situational perception and the personality of a person shaping human behavior. As outlined, previous research has already focused on the taxonomization of situational information, the taxonomization of personality variables, as well as the interplay of persons and situations. Questions such as "which situational perception dimensions entail which kinds of behavior?" or "which situational perception dimensions are linked to which personality traits?" have been posed and empirically investigated (Rauthmann, Sherman, Nave, et al., 2015; Sherman et al., 2015). For example, Sherman and colleagues (2015) could show that perceiving momentary deception in situations is tied to behaving less honest and humble. This effect was significant after controlling for the influence of trait honesty and humility. Yet, it is reasonable to assume that there are further person factors which also play a crucial role in the interplay of situational perception and behavior. As presented in the introductory example, current mood or affect are linked to the perception of the situation. Wilson and colleagues (Wilson et al., 2017) have shown that behavioral variability is more than variability in mood and affect. Thus, they suggest that the effect of the situation should be considered to explain behavioral variance. For example, one question so far unresolved is whether mood influences situational perceptions or whether it is the other way around. Both options are possible, even at the same time (Horstmann & Ziegler, accepted).

The next step in situation research will be an integration of these different findings. How do affect, situational perception, personality, social roles, and other constructs explain behavior in a situation? Several models or conceptual equations have been proposed that include a different number of these constructs (e.g., Bond, 2013; Cattell, 1963; Lewin, 1936; Rauthmann, Sherman, & Funder, 2015b; Westhoff & Kluck, 2008). Even though these models are not new, the possibility to measure situations is. We therefore think, that – with the advent of situational perception measures – the time is right for a systematic examination of these models. If these models, and thereby the constructs are successfully examined in more detail, then this will increase knowledge about their nomological net and ultimately inform new and better measures.

Based on the accumulated evidence so far, we would like to make a proposition summarizing some of these ideas and extending Lewin's (1936) equation. Behavior was defined by Lewin as a function of momentary person states and one's actual environment: $B = f(P, E)$. We

suggest extending by plugging in situational perception, both as a state (s) as a proxy for E, and a trait (S), personality (P), state affect (a), and trait affect (A) as a proxy for P:

$$B = f(S, s, P, A, a)$$

Note that we do not consider personality states (p) as a predictor of behavior. Instead, we would like to suggest breaking down personality states into more nuanced pieces. Here, as a first step, we suggest affect (a), and behavior (B). If behavior should be predicted, and behavior is conceptualized as a personality state, this personality state should not be included on the right-hand side of the equation. Of course, other so far unheeded states (e.g., emotion, motivation, social roles) could also be included on the right-hand side, and the current equation servers as a validated starting point for future research that includes situational perception as a predictor for behavior.

The hypothesis is not only that behavior can be predicted by S, s, P, A, and a, but that all of these elements are constructs in their own right as well and therefore independent from each other to serve as distinct predictors. This also means that it is worthwhile and important to further taxonomize each element, to develop measurement instruments capturing their unique variances, and to only then examine the relations of the constructs and their interactions.

## 2.3 The Situation Five

As the componentiality corollary states, situation perception consists of three sources: situation variance, person variance, and their interaction (Rauthmann, Sherman, & Funder, 2015b). The person variance can assumed to be stable across different situations and thus be considered a personality trait (Horstmann & Ziegler, 2016; Ziegler & Horstmann, 2015). The assessment of such a person variable could prove useful for the prediction of behavior in future, unknown situations. Even if the situation itself (i.e., constituting situation cues) was unknown, situation perception as a person-variable could explain variance in the perception of specific situations and thus ultimately in a person's reaction to them. The aim of the article *Personality in Situations: Going Beyond the OCEAN and Introducing the Situation Five* (Ziegler et al., submitted) was first to develop a taxonomy of situation characteristics for an occupational context, and second, to develop a measurement tool that allows assessing situation perceptions on a trait level. Based on the lexical approach (Allport, 1936; Parrigon, 2017), we first developed the Situation Five: Outcome-Expectancy, Briskness, Cognitive Load, Psychological and Physical Load, and Lack of Stimuli. Second, we developed a situational judgment test, the Big Five Inventory for Personality in Occupational Situations (Ziegler, 2014a) that allows assessing situation perception and personality simultaneously.

The development of the Situation Five filled a gap left open by other situational taxonomies as they were explicitly developed to describe occupational situations (Horstmann, Rauthmann, et al., 2018). Further, this research showed that situation perception has indeed a stable trait component that allows predicting variance in work-engagement. This trait-component, which is independent to and incrementally valid over the Big Five personality traits, should thus be considered in future applications of situation perception measures in order to disentangle effects of trait situation perception from effects of state situation perception.

### *2.3.1 Article Summary: Personality in Situations*

**Background.** Situation perception is the result of an individual process during which situation cues are perceived and interpreted (Rauthmann, 2012; Rauthmann, Sherman, & Funder, 2015b). As such, it is likely that situation perception has a stable, trait-like component (Horstmann & Ziegler, 2016; Ziegler & Horstmann, 2015). One of the challenges of situation research has been that future situations are usually unknown (Hogan, 2009). Assessing general tendencies to perceive situation perception, however, eliminates this shortcoming and makes situation perception available for the prediction of future behavior.

The aim of the current study was first to develop a taxonomy that can be applied to an occupational context (the Situation Five), and second to develop a measurement tool that allows assessing trait situation perception as well as the Big Five personality traits. To this end, several independent studies were conducted.

**Study Design.** The development of the situation perception taxonomy is based on the lexical approach (Allport & Odbert, 1936; Parrigon, 2017). Initially, all adjectives (about 15,000) from the German spelling dictionary (Dudenredaktion, 2006) were extracted. In two consecutive steps, three independent raters selected a subset of 300 adjectives suitable to describe situations. In a planned missingness design, $N = 521$ participants rated everyday situations on these adjectives. Using exploratory factor analyses, seven initial factors were extracted. In a second sample ($N = 387$), five of these seven factors were confirmed, which we referred to as the Situation Five: Outcome-Expectancy, Briskness, Cognitive Load, Psychological and Physical Load, and Lack of Stimuli.

**Measurement Tool.** For the assessment of situation perception as a personality trait, we first designed hypothetical situations that resembled typical situations from an occupational context. To this end, 16 Human resource managers were interviewed and interviewed about typical work-situations. These situations were adapted to form 211 different situation vignettes, that is, brief hypothetical scenarios at the workplace.

In the final assessment tool, the Big Five Inventory of Personality in Occupational Situations (Ziegler, 2014a), participants were first required to read these vignettes. After reading them, participants rated the situation on one of the Situation Five dimensions. Subsequently, they indicated how they would behave in each situation. This second item assessed one of the Big Five dimensions. The Big Five items were developed by Ziegler, Cengia, and Roberts (in prep.) and based on the International Personality Item Pool (IPIP, 2015). Using a third sample of $N = 389$, which was representative for the German speaking population, the structure of the Situation Five and Big Five could be confirmed.

**Reliability and Validity.** In a final step, we examined discriminant and convergent validity as well as test-criterion correlations of the Situation Five and Big Five. The Situation Five and Big Five show good reliability (internal consistency, for $N = 389$, as well as test-retest reliability, $N = 96$). Furthermore, the Big Five showed convergent and discriminant validity with another measure of the Big Five (Arendasy, 2011). The Situation Five displayed discriminant validity to both other Big Five measures. Finally, the Situation Five allowed explaining substantial variance in work-engagement (assessed with the Utrecht Work Engagement Scale, Seppälä et al., 2009).

**Conclusion.** Situation perception has a stable trait component. The perception of a situation and thus behavior may thus not only be explained with the current state perception of a situation, but also with stable, trait-like tendencies of individuals to perceive and interpret situation cues.

### *2.3.2 Article: Personality in Situations*

**Personality in Situations:**
**Going Beyond the OCEAN and Introducing the Situation Five**

Matthias Ziegler[1], Kai T. Horstmann[1], & Johanna Ziegler
Humboldt-Universität zu Berlin, Germany
[1]shared first authorship

This article was conditionally accepted in the journal *Psychological Assessment*:

Ziegler, M., Horstmann, K. T., & Ziegler, J. (conditionally accepted). Personality in Situations: Going Beyond the OCEAN and Introducing the Situation Five. *Psychological Assessment.*

This is the conditionally accepted version before copy editing.

**Funding:**
This research was supported by Schuhfried GmbH, Hyrtlstraße 45, AT-2340 Mödling.

**Author Note:**
The data presented in this manuscript were also used in the Manual of the proprietary test, The Big Five Inventory of Personality in Occupational Situations (Ziegler, 2014a).

**Abstract**

We present the psychometric evaluation of a personality measure that assesses the Big Five and situation perception based on a newly developed taxonomy of situation characteristics. Following the lexical approach, more than 15,000 adjectives were extracted from an authoritative German dictionary. In a first exploratory study, 521 participants rated every-day situations on 300 adjectives selected as potential situation descriptors. Seven dimensions of situation perception were initially extracted. In a second study with $N = 387$, five of these seven factors were confirmed: Outcome-Expectancy, Briskness, Cognitive Load, Psychological and Physical Load, and Lack of Stimuli, together referred to as the Situation Five. Finally, a measurement tool, the Big Five of Personality in Occupational Situations (B5PS), was constructed to assess the Big Five personality traits and the Situation Five simultaneously. We present evidence for the reliability, convergent and discriminant validity and predictive validity of the B5PS test scores. Our study highlights the relevance of situation perception as a trait and discusses their applicability in diverse contexts.

## Personality in Situations: Going Beyond the OCEAN and Introducing the Situation Five

It is widely accepted that the behavior of a person is influenced by internal aspects of the person, such as personality, but also by external factors such as aspects of the situation (Lewin, 1936; Mischel & Shoda, 1995; Rauthmann, Sherman, & Funder, 2015b). Personality theories have been exhaustively examined, and numerous well-accepted taxonomies for the description of personality have been developed (e. g., the Big Five, Goldberg, 1990). Recently, the convergence of such models with pathological traits has been established (Thomas et al., 2013). Furthermore, scores derived from measures of personality were successfully linked to external criteria for example career success or subjective well-being (e.g., Barrick, Mount, & Judge, 2001; Ozer & Benet-Martínez, 2006; E. Richard & Diener, 2009; Schmidt & Hunter, 1998). The situation, and especially the assessment of situational influences, has only recently received wider research attention (Horstmann, Rauthmann, et al., 2018). In recent years several taxonomies of situation (perception)s have been proposed, such as the DIAMONDS or CAPTION framework (Parrigon et al., 2017; Rauthmann et al., 2014), and it was shown that the measured perception of a situation is predictive for behavior in this situation (Sherman et al., 2015). Unfortunately, the development of measures incorporating both, situation perception and personality, is lacking behind. Considering the relevance of both constructs for assessment in clinical or organizational contexts, it seems important to show that a measure including both can be constructed. Most recently, Rockstuhl[11] and colleagues (Rockstuhl, Ang, Ng, Lievens, & Van Dyne, 2015) examined if the judgment of a situation and the response to the situation in situation judgment tests can be differentiated. Not only was this the case, but they demonstrated convincingly how each response predicts unique variance in meaningful outcomes.

So far, no standardized inventory exists that allows assessing situation perception tendencies as well as behavioral responses to situations as indicators of personality at the same time. The current project was undertaken to close this gap. To this end a two-stage research project was conducted. Within the first stage a taxonomy of situation perception was developed using a lexical approach. In the second stage, an inventory measuring situation perception and personality simultaneously was constructed and evaluated.

[11] We are thankful to a reviewer for directing us towards this study. Please note that we were unaware of this publication by the time we designed and analyzed the current study. Rockstuhl and colleagues' work therefore has not informed the data collection of the present work, yet their results support the necessity for this investigation, which is why we report their study here.

**What Is a Situation?**

Rauthmann and colleagues (Rauthmann, Sherman, & Funder, 2015b) proposed that situation information can be found on three ordered levels: Situation *Cues*, situation *Characteristics* and situation *Classes*. Cues describe the elements that constitute a situation objectively, such as a table or coffee cup. Characteristics describe the psychological meaningful information of a situation. A situation could for example be pleasant, stressful, deceptive, or demanding. It is important to note that the same situation can be intellectually demanding for one person, whereas it can be rather dull for another. On an even more aggregated level, classes refer to different situations that are similar in either their situation characteristics or situation cues. Thus, situations that are perceived as intellectually stimulating as well as cognitively demanding could be subsumed in a class that could be labeled "Learning". Yet, no comprehensive taxonomy for situation classes has been developed thus far (Horstmann, Rauthmann, et al., 2018). Although situation classes or contexts have been used in previous research (e.g., Geukes et al., 2017), their application does not allow a closer look at individual differences, as they assume that the psychological situation is similar for all persons within a class (Horstmann, Ziegler, et al., 2018). We thus argue that the most informative information on situations is that of situation characteristics, measured via situation perceptions.

**What influences situation perception?** Ziegler and Horstmann (2016; 2015) described a model of situation perception which underlies the current work. The idea is that situation cues are processed by each individual in any given situation. This information processing is in part idiosyncratic (Rauthmann, 2012; Rauthmann, Sherman, & Funder, 2015b). That means that any situation judgment by a person is influenced by the situation judged, by the person that judges, and by their interaction. However, utilizing this person by situation-specific variance would require single case assessments, which is not feasible in many settings (Hogan, 2009; Ziegler & Ziegler, 2015). Yet, as the proposed model suggests, underlying dimensions on the person-level exist that structure situation perceptions. They are the result of an individual information processing based on idiosyncratic patterns. In other words, we assume that the person-variance in situation perceptions can be measured reliably and is distinct from other psychological trait-like constructs such as the Big Five personality traits. It was empirically shown that situation perceptions on the person level allow explaining variance above and beyond established constructs (e.g., Rockstuhl et al., 2015; Sherman et al., 2015). Although the study by Rockstuhl and colleagues assessed situation perception on the person level to predict other outcomes, they did not employ a systematically derived and generalizable taxonomy of situations. On the other hand, Sherman and colleagues used an established framework to assess situation perception (the DIAMONDS by Rauthmann et

al., 2014), but they did not use a fixed set of situations. Thus, their situation perception scores were not comparable across participants, as each participant experienced different situations. The current study combines these two approaches: First, a systematic taxonomy is developed for the description and assessment of situations. Second, a questionnaire is developed that uses a fixed set of situations, thereby allowing to compare scores of situation perceptions between persons.

**The Importance of Situation Perception**

Situation perception has only recently been established as a construct on its own (Rauthmann, 2012). However, the principles of situation perception – namely the evaluation of external stimuli based on previous knowledge, personality, and current states – can be found in most areas of psychology. First, any study that contains ratings by participants about their "current situation" contains a situation judgment. For example, in Oud et al.'s study (Oud et al., 2017), participants had to state whether they were at home or at work. This judgment would most likely be based on their location, which is a situation cue. In some cases, though, people might work at home, which makes these judgments less valid. Ratings of participants' perceptions of work or home situations in general would thus be more beneficial. Second, appraisal theories of emotions are concerned with the perception of cues, their interpretation, and their resulting effect on affect (Horstmann & Ziegler, 2018; Kuppens, 2009; Sander, Grandjean, & Scherer, 2005). Measures of situation perception could shed light on this process, which ultimately makes situation perceptions also relevant for clinical psychology. Schwartz and Weinberger (1980) for example examined how different anxiety evolving situations influenced emotions. Similarly, Keller and Nesse (2006) examined how depression is affected differentially by varying situations. Other examples include the examination of anxiety in different naturally occurring situations (Chen et al., 2010) or even the use of mobile sensing that informs interventions for depressed patients (Burns et al., 2011). The interest in the "current situation" of a person is further reflected in the widely used SORC-model known from cognitive behavioral therapy (Kanfer & Saslow, 1965). SORC stands for Stimulus, Organism, Reaction, Consequence, and Stimulus can be understood as the perception of situation cues. Although all of the above studies shared a combined interest in the role of the situation in a clinical intervention, they also shared a lack of a systematic and generalizable way to assess the situation or even situation perception. As such, clinical research could benefit from a common framework that allows describing situations and assessing perceptions thereof. Moreover, the measure introduced here showcases the feasibility and advantages of measuring personality and situation perception in one instrument and being able to compare the obtained scores across participants.

Furthermore, situation judgment plays and important role in occupational psychology (Lievens, 2017; McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001), and not only in the role of situational judgment tests. Predictive validity of ratings can be enhanced if these ratings are obtained in situations that are similar to later work environment (e.g., Shaffer & Postlethwaite, 2012), and can thus also be applied to behavioral tasks in assessment centers or interviews. However, these tests usually rely on implicit assumptions about the similarity of situations designed for the hiring process and those encountered later. A systematic descriptive system of occupational situations could thus be used during the analysis of job requirements, which in turn could inform the situations used in the hiring process. Thus, the current project in its first stage also aimed at uncovering and structuring dimensions of situation characteristics.

### Situation Perception – Taxonomy and Assessment Tool

To assess trait situation perception two conditions must be fulfilled: First, a taxonomy must exist that describes which dimensions situations can be described by, and second, an assessment tool must exist to measure intra-individual differences in situation perception.

**Existing taxonomies of situation perception.** Several taxonomies to describe situations have already been proposed (Horstmann, Rauthmann, et al., 2018; Rauthmann et al., 2014; Yang, Read, & Miller, 2006), and most of them focus on describing differences in situation characteristics. On the broadest level, these existing taxonomies can be classified into two groups. First, taxonomies that were constructed based on data, that is, bottom up (e.g., Rauthmann et al., 2014; Yang et al., 2006), second, taxonomies construed to explicitly reflect theoretical assumptions, that is, top-down (e.g., N. A. Brown et al., 2015; Gerpott et al., 2017). It is important to note that these different approaches to the development of situation perception taxonomies converge on a broader level (Horstmann, Rauthmann, et al., 2018; Rauthmann & Horstmann, 2017), and that overarching dimensions of situation perception can be found (Rauthmann & Sherman, 2018a).

## Stage 1: Development of the Situation Five

Thus, the first aim was to develop a taxonomy for situations as a vital component of a questionnaire assessing personality and situation perception at the same time.

### Study 1 - Development of the Situation Perception Taxonomy

For the construction of the final assessment tool (see Study 3), the Big Five in Psychological Situations (B5PS), two different taxonomies and related scales had to be constructed and then combined. The first taxonomy was meant to portray dimensions of situation perception. The second taxonomy was meant to establish a broad facet structure of the Big Five. The construction of this latter taxonomy and the related scale is reported elsewhere (Cengia, Ziegler, & Roberts, in prep.; MacCann, Danay, Ziegler, & Roberts, 2011).

**Lexical approach**. The development of the situation perception taxonomy was grounded in the lexical approach. The lexical approach is most widely known with respect to the construction of the Big Five (John & Srivastava, 1999). Its general assumption is that relevant and distinguishable features of an object are reflected in human language (Allport & Odbert, 1936; Deary, 2009), but there is also a general agreement that this also applies to characteristics of situations (e.g., Edwards & Templeton, 2005; Parrigon, 2017; Van Heck, 1984, 1989; Yang et al., 2006). Thus, if a feature of a situation is perceived as relevant, a word should exist to describe it.

**Sample.** For the initial construction of the situation perception dimensions, we used a sample of $N = 521$ participants, who were mainly students. Their mean age was $M = 25.58$ (median = 24, $SD = 6.59$), $N = 135$ were male (25.91%), 379 were female (72.74%).

**Materials and procedure.** We first selected adjectives that could be applied to the description of situations. These adjectives were then used in a survey to rate perceptions of different situations.

***Adjectives for situation ratings.*** To obtain adjectives that were suitable for the ratings of situations, we first extracted all 15,679 adjectives from the German spelling dictionary, the Duden (Dudenredaktion, 2006). These adjectives were then reduced to 300 by three independent raters. First, two raters examined the initial list of adjectives. Adjectives were excluded if they were generally not suitable to describe situations (e.g., scavenging, massless, careless, or uvular), were outdated or rarely used (e.g., haggard or earthy). Examples of items that were unknown to one of the two coders include the German *vif* (meaning lively), or *adiabatisch*. Across both raters, 228 adjectives were rated "unknown" by at least one of the raters, 4826 adjectives were rated as suitable by at least one of the raters, and 1934 adjectives were rated as suitable by both raters simultaneously. Removing the items that were marked as unknown, the raters achieved an agreement of kappa = .46.[12] All 4,826 adjectives that were rated as suitable by at least one rater were selected for the next step. Using an online dictionary, synonyms or hypernyms were identified and removed. For example, *abenteuerlich* (adventurous) has about 15 hypernyms. Two raters then removed all words that had a very similar meaning from this initial list, thereby reducing it to 2,720 words. These adjectives were then rated by a new rater using the same system as above, identifying 200 words as suitable. The third rater's goal was to reduce the number of items to a manageable size. To this end, he examined the list several times, iteratively. Finally, based on the ratings of each individual rater and the elimination of synonyms, all three raters discussed each of the 2,720 adjectives again and came up with a final list of 300 adjectives. Based on the ratings from and

[12] Note that this interrater agreement must not necessarily high: For example, if one rater was simply not familiar with an adjective, but the other rater definitely knew the adjective would not apply to situations the inter-rater agreement would decrease.

discussion among all three raters, these adjectives were not redundant, not out of use or barely known, and could be used to describe a situation. The final list of adjectives is available in the OSM.

***Data collection.*** Participants completed an online questionnaire. First, they indicated their age, gender, and other demographic variables such as educational background. They subsequently had to answer the following question: "What did you do yesterday at 11a.m./4p.m./9p.m.?" (the respective time was randomly selected for each participant). After giving a brief verbal description of the situation, participants evaluated how well each adjective described their selected situation on 75 out of the 300 adjectives, on a 4-point Likert-type scale ranging from strong disagreement (1) to strong agreement (4). All in all, nine item blocks were created and randomly assigned to keep participant burden minimal. We used nine item blocks to ensure that each participant had some overlap in the adjectives used with participants that responded to another item block (e.g., block 1 includes items 1-75, and block 2 items 51-125, and so on), thus "linking" the item blocks to each other. Since we specified the missing adjectives in advance, the missing data (i.e., 225 adjectives that were not rated in a given block by one person) are missing completely at random (Rubin, 1976) and can therefore be imputed. We used an expectation maximization algorithm in SPSS to impute the missing data resulting in, thereby obtaining a full data set with $N$ = 521 participants and ratings across 300 adjectives. These ratings were analyzed using exploratory factor analysis.

**Results of exploratory factor analyses (EFA).** We used the psych-package (Revelle, 2014) in R (R Core Team, 2016) and Mplus 7.1 (Muthén & Muthen, 2015) to analyze the data. To estimate the number of factors to extract, we examined the distribution of Eigenvalues, the minimum average partial test (MAP) and parallel analysis (1000 iterations). The first nine Eigenvalues were as follows: 8.99, 7.67, 5.79, 5.63, 5.09, 4.94, 4.24, 3.89, 3.64 (see Figure 1, Online Supplemental Material [OSM], for the first 40 Eigenvalues). The MAP-test suggested 34 factors and parallel analysis 23. Based on content and, most importantly, interpretability of the factors and with the goal to extract broader, more heterogeneous constructs, we finally extracted seven oblique factors: Valence, Temporal environmental conditions, Cognitive Load, Psychological and Physical Load, Briskness, Lack of Stimuli, and Outcome-Expectancy. The items with the highest loadings are presented in Table A (OSM). Temporal environmental conditions, that is, the weather, did not have a psychological relevance for the final assessment tool and its items were therefore excluded from all further analyses. Another factor called Valence correlated highly with the remaining factors and was thus examined in later analyses as two higher order factors, named Burden and Vigor (see Figure 2, OSM).

**Study 2 - Replication in an Independent Sample**

The aim of Study 2 was to replicate and test the individual factors established in Study 1 in an independent sample and obtain scales with a balanced number of items for each factor.

**Sample.** The sample consisted of $N = 387$ participants chosen to representatively match the German speaking population. Their mean age was 45.59 (median = 44, $SD = 17.49$), $N = 197$ participants were female (50.90%).

**Materials and procedure.** Participants were invited to a laboratory and first provided information on demographic variables (age, gender, and level of education). Participants were required to name a situation from the previous day ("What did you do yesterday at 11am./4p.m./9p.m.?") and the time was randomly chosen for each participant. Afterwards, each participant evaluated this situation on the same 4-point Likert-type scale used in Study 1. All in all, 59 adjectives from Study 1 were selected based on their factor loadings in the exploratory factor analyses and interpretability. Selection criteria were a balanced representation of each dimension as well as a broad coverage of the adjectives used. Moreover, item difficulties and loading sizes were considered. For example, items that had lower loadings, but extreme item difficulties were not removed to obtain a measure that is able to differentiate between persons with different trait levels. In some cases, adjectives were replaced by more colloquial alternatives to obtain higher discriminant validity of the factors. Examples of this include inauspicious (replaced with burdensome), barren (replaced with depressing), delightful (replaced with full of expectation), professional (replaced with productive and prolific). Furthermore, additional items were added in Study 2 to obtain an almost balanced number of items for each factor.

A confirmatory factor analysis was conducted for each measurement model of each of the initial six dimensions (without temporal environmental conditions). The factor Valence correlated highly with all other factors and was thus modeled as a higher order factor.

Five factors remained after this initial analysis. Some items were additionally deleted during the CFA, so that finally 48 adjectives remained. Selection was again based on loadings, item difficulties, and prototypicality for the specific factor (Ziegler, 2014b). The selected adjectives are displayed in Table A (OSM).

**Results of confirmatory factor analysis (CFA).** We first computed the CFAs as planned and subsequently added exploratory analyses of higher order factors.

***Measurement models.*** Table 2.3.1 shows the results for each of the five measurement models. Based on the criteria for model fit by Hu and Bentler (1999), each model has an excellent fit, meeting the requirements for the CFI, RMSEA, and SRMR. The final dimensions are

Outcome-Expectancy, Briskness, Psychological and Physical Load, Lack of Stimuli, and Cognitive Load. Combined, we refer to them as the Situation Five.

***Higher order factors.*** Additionally, and in an exploratory manner, we estimated a higher order factor model of the Situation Five. To this end, we computed item parcels containing 2-3 items. The items were assigned to their parcels based on the loading in an exploratory factor analysis (Little, Cunningham, Shahar, & Widaman, 2002). This approach allowed estimating the full model while keeping the number of participants reasonably low. This model described the covariances of the Situation Five reasonably well (Figure 2, OSM). It had an acceptable model fit ($\chi^2$ = 530.84, *df* = 86, $p < .001$, *CFI* = .89, *RMSEA* = .12, *SRMR* = .13, estimator: MLR).

Table 2.3.1
*Model-Fits for the Situation Five and Big Five Models*

| | Study 2 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Situation Five* | $X^2$ | *p* | *df* | CFI | TLI | RMSEA | SRMR | Estimator |
| Outcome-Expectancy | 124.44 | < .001 | 27 | .98 | .97 | .097 [.080; .114] | .04 | WLSMV |
| Briskness | 107.95 | < .001 | 20 | .96 | .94 | .107 [.087; .127] | .06 | WLSMV |
| Psych. and Phys. Load | 116.99 | < .001 | 27 | .98 | .98 | .093 [.076; .111] | .04 | WLSMV |
| Lack of Stimuli | 86.50 | < .001 | 27 | .99 | .99 | .076 [.058; .094] | .03 | WLSMV |
| Cognitive Load* | 70.94 | < .001 | 16 | .98 | .97 | .094 [.073; .117] | .05 | WLSMV |
| | Study 3 | | | | | | | |
| *Situation Five* | $X^2$ | *p* | *df* | CFI | TLI | RMSEA | SRMR | Estimator |
| Outcome-Expectancy | 3424.81 | < .001 | 819 | .74 | .72 | .090 [.086; .093] | .09 | WLSMV |
| Briskness | 3920.56 | < .001 | 860 | .67 | .66 | .095 [.092; .098] | .10 | WLSMV |
| Psych. and Phys. Load | 2270.24 | < .001 | 819 | .88 | .88 | .067 [.064; .070] | .07 | WLSMV |
| Lack of Stimuli | 2707.28 | < .001 | 819 | .71 | .71 | .076 [.073; .079] | .09 | WLSMV |
| Cognitive Load | 3107.94 | < .001 | 629 | .58 | .55 | .100 [.096; .103] | .11 | WLSMV |
| *Big Five* | $X^2$ | *p* | *df* | CFI | TLI | RMSEA | SRMR | Estimator |
| Emotional Stability | 42.74 | < .001 | 14 | .96 | .94 | .072 [.050; .095] | .04 | MLR |
| Extraversion | 104.60 | < .001 | 25 | .91 | .88 | .090 [.074; .106] | .05 | MLR |
| Openness/Flexibility | 59.90 | < .001 | 27 | .95 | .93 | .055 [.039; .072] | .04 | MLR |
| Agreeableness/Team-Orientation | 48.56 | < .001 | 18 | .89 | .83 | .065 [.045; .086] | .05 | MLR |
| Conscientiousness | 113.69 | < .001 | 27 | .88 | .84 | .090 [.075; .105] | .06 | MLR |

*Note.* $N$ = 387 (Study 2), $N$ = 398 (Study 3).
The Situation Five were modeled on the item level in both studies, whereas the Big Five were modeled on the facet level.

**Discussion Studies 1 and 2 – Development of a Situation Taxonomy**

In the first two studies, we reported the development of a taxonomy for situation characteristics of everyday situations, the Situation Five. The five dimensions are Outcome-Expectancy (adjectives include "professional", "confident", "promising results"), Briskness (e.g., "vivid", "enchanting", "lively"), Psychological and Physical Load (e.g., "irksome", "woebegone", "tense"), Lack of Stimuli (e.g., "boring", "barren", "dull"), and Cognitive Load (e.g., "mind-wracking", "demanding", "excruciating").

Despite the fact that some adjectives were added after the initial exploratory study (Study 1), the results of the confirmatory analyses showed that these adjectives were useful to capture the psychologically salient aspects of situations. Four out of five confirmatory models computed showed acceptable model fit. One exception, though, was the model of Cognitive Load: Four adjectives ("*geistig stimulierend"* [mentally stimulating], "*herausfordernd" [*challenging, demanding*]*, "*geistig anregend"* [mentally inspiring], and "*anspruchsvoll"* [demanding, challenging]) had correlated residuals, possibly due to their shared content of 'cognitive stimulation'. This shared variance was therefore modeled in a bi-factor. Differences between the models of Study 1 and Study 2 were small, although some adjectives were changed and the student sample (Study 1) was replaced with a representative sample (Study 2).

The five dimensions of situation perception fit nicely into other existing taxonomies of situation characteristics (Horstmann, Rauthmann, et al., 2018) and thereby broaden and extend the nomological net of situation perception. Horstmann and Ziegler (2018) presented evidence that these dimensions are meaningfully related to the DIAMONDS (Rauthmann et al., 2014), and explain variance in positive and negative affect within a given situation.

The most prominent other recent situation taxonomy that examined the structure of situations in a lexical approach is the CAPTION taxonomy by Parrigon and colleagues (Parrigon et al., 2017). The authors identified seven situation dimensions: Complexity, Adversity, Positive Valence, Typicality, Importance, Humor, Negative Valence. As Horstmann and colleagues argued, these dimensions fit reasonably well to the five dimensions extracted in the current study (Horstmann, Rauthmann, et al., 2018; Rauthmann & Horstmann, 2017; Rauthmann & Sherman, 2018a). Negative Valence corresponds to Psychological and Physical Load, Adversity to Cognitive Load, Importance to Outcome-Expectancy, Positive Valence and Humor to Briskness, and Typicality to Lack of Stimuli. Especially these two last dimensions – Typicality and Lack of Stimuli – were so far only identified when using a lexical approach, thus lending further credibility to the current findings.

Furthermore, the Situation Five were shown to have two higher order factors of situation perception, which we named Burden (loading on Psychological and Physical Load, Cognitive Load, and Lack of Stimuli), and Vigor (loading on Outcome-Expectancy and Briskness). Although the model fit of this model did not meet the requirements set by Hu and Bentler (1999), it is very similar to the fit presented by Rauthmann and Sherman (2016) for a CFA model of the DIAMONDS. These factors may represent positive and negative valence. Even though it is likely that higher order dimensions of situation perception exist (Horstmann, Rauthmann, et al., 2018; Rauthmann & Sherman, 2018a), the two dimensions presented here have to be replicated and examined in more detail before they can be clearly interpreted. We will therefore focus on the Situation Five domain level in the current article.

## Stage 2 – Development and Validation of the B5PS

The objective in this stage was to develop a questionnaire that measures the Situation Five as traits as well as the Big Five personality traits. The decision to measure situation perceptions as traits was due to a specific criticism often brought forward against situation research, namely that the prediction of future behavior with measures of state situation perception (i.e., perception of individual situations) is difficult. Hogan (2009) correctly stated that predicting behavior from one situation experience would lead to fruitless point predictions: One could only predict behavior of a person when both the person as well as the future situation are known. However, it is of course nearly impossible to predict how a person will perceive a future situation. *In situ* situation perceptions are vitally important for understanding how and when a situation will influence behavior, that is, for the *explanation* of behavior; yet they are less useful for the *prediction* of behavior in future unknown situations. We therefore argue that the general tendency of a person to perceive certain situations is relevant for the prediction of future behavior. As pointed out above, showing that situation perceptions can be measured like personality traits would have implications for all fields of psychology, especially organizational and clinical where the prediction or malleability of behavior is often focused.

### Comparison of Trait Situation Experiences

The existing tools for the assessment of situation perception often measured situation perception of a participant by asking how a randomly selected situation was perceived, either by using an experience sampling design or by asking something along the lines of "how did you perceive the situation you were in yesterday at 12 a.m.?". This is a valid approach for the construction of a situation taxonomy, yet it is problematic when scores of situation experiences should be compared between participants (Lievens, 2017). For example, it is likely that participants will select situations according to their preferences (Rauthmann, Sherman, Nave, et al., 2015) and

thereby report a biased set of situations: It is, on average, more likely that an extraverted person will experience more situations that are perceived as, for example, sociable. This can be either due to situational selection or situational construal. Thus, as long as the situations are not fixed, scores on their perceptions are hard to compare. Being able to measure situation perception as a trait would eliminate this limitation.

**The B5PS and the ABC of Test Construction**

Whenever a new psychometric assessment tool (e.g., questionnaire, test, observation) is developed, some background information on this measure should be provided. Ziegler (2014c) presented a framework for this purpose, which can briefly be summarized with three questions:

A) What is the construct being measured?

B) What is the intended use of the measure?

C) What is the targeted population?

The answer to the first question requires that the construct is extensively described and integrated into a nomological net. This integration can happen on two levels: Either by linking it to theoretically relevant other constructs, or by providing empirical evidence for the relation to other constructs. Second, the intended use of the measure is important, since not every measure can or should be applied in all contexts. Some measures may be more relevant for empirical research (e.g., short scales), whereas other measures may be more relevant for an applied context. Whatever the intended purpose, sufficient empirical support should be provided that shows the applicability of a measure in its intended context. Third, the targeted population has to be clearly stated. A test that has been developed for one population is not necessarily applicable to another population.

With regard to the most recent situation taxonomies and their assessment tools, Horstmann and colleagues (Horstmann, Ziegler, et al., 2018) provided answers to these three questions. Measures of such taxonomies usually focused on A) the measurement of perceived state situation characteristics of random, everyday situations, B) the use of situation measures to predict in situ behavior, and C) the assessment of every-day situations, judged by a population of normal (i.e., non-clinical) participants. With the B5PS, we present a measurement tool that assesses situation perception and personality simultaneously as traits, is therefore applicable in an applied context (with a focus on personnel selection and occupational health) and can be used for all people representative of the German working population. On a more abstract level, we demonstrate a principle which can also be adopted in clinical, educational, or other assessment settings.

We first describe the development of a new questionnaire to assess Big Five personality traits. We then describe how fixed situations were generated so that the Situation Five can be assessed on the same set of situations for each participant. Finally, the integration of situation and personality measures will be described.

**Development of personality scales.** The personality dimensions were taken from another study (Ziegler et al., in prep.). Based on items from the International Personality Item Pool (IPIP, 2015), Ziegler and colleagues constructed scales that measure the Big Five Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Each of the domains has 7 to 9 facets, 42 all together, and each facet is measured with 5 items each. In a first step, the facet structure was derived from a sample of 726 US college students (MacCann, Danay, Ziegler, & Roberts, 2011). In a second step, the items were translated into German and the facet structure as well as measurement invariance with the US sample was confirmed in an online sample of $N$ = 387 (Ziegler et al., in prep.).

**Development of Situation Vignettes**. The comparison of inter-individual differences in situation perception is only possible if the participants perceive the same or very similar situations. As a remedy, brief descriptions of hypothetical, work-related situations were therefore created, so called situation vignettes. We focused on work-related situations for several reasons: First, similar to the argument made later by Rockstuhl and colleagues (2015), we sought to develop a tool that can be used in the applied context of situation selection and prediction of later behavior. Second, and despite the relevance for clinical psychology, the necessary samples that are relevant for a clinical scale are harder to obtain in sufficient quantity. Nevertheless, the principles presented here could also be applied to a clinical setting and should therefore be understood as a proof of concept.

For the construction of the situation vignettes, sixteen human resource managers were interviewed, asking them about typical situations within occupational settings. The resulting situations were re-framed so that they were brief (one or two sentences), unambiguous, and would relate well to one of the five dimensions of situation perception. The last point stresses that only certain behaviors are possible in certain situations, as postulated by the trait activation theory (Tett & Burnett, 2003; Tett & Guterman, 2000). It was also ensured that the vignettes allow a certain range of behavior: A fictional situation should not prompt the participant directly towards the "most appropriate" behavior in that situation, or it should not restrict the behavioral response options too much. Overall, 211 different situation vignettes were created. Each of the situation vignettes was expected to tap only one of the five dimensions of situation perception.

**Integration of Big Five and Situation Five**. To measure the Big Five and the Situation Five simultaneously, both were combined using an approach similar to a situation judgment task.

The general idea is to present a brief description of a situation. Participants are required to first make a statement about their perception of this hypothetical situation and subsequently indicate how they would behave in it. For the B5PS, 211 of such items were constructed with one item for situation perception and one item for personality per vignette, resulting in 211 items for situation perception and 211 items for personality.

**Item construction for the B5PS.** The 211 items capturing the Situation Five were created based on the adjectives identified as markers in the previous two studies. Each adjective was used to complete the sentence: "I perceive the situation as …". We adapted personality items marking the 42 facets previously identified.

**Method**

**Sample and procedure.** The final version of the questionnaire was applied to a representative sample of the German, Austrian, and Swiss population. The sample consisted of 173 women (43%) and 225 men (57%), ranging from 17 to 69 years. The mean age was 42.03 years (*SD* = 13.91). The data were collected in a controlled environment in the laboratories of Schuhfried GmbH in Mödling, Austria.

**Materials.** Participants responded to the final version of the B5PS. This included the 211 situation vignettes, for example *"You just had your annual appraisal interview with your manager, in which you received a lot of detailed feedback."* In a first step, participants are required to rate how they would perceive such a situation, such as "*I perceive this situation as challenging*" (the last word/phrase is always the adjective identified in Studies 1 and 2). Subsequently, participants indicated how they would behave in this situation: *"I reflect on the feedback".* Responses were given on a 6-point Likert-type scale (1 = "strong disagreement" to 6 = "strong agreement"). Each Big Five was combined with each Situation Five in a total of eight to ten vignettes, resulting in an overall total of 211 situation vignettes.

**Data analysis.** All data were analyzed using the *lavaan* package (Rosseel, 2015) in R. First, confirmatory factor analyses were computed for each measurement model of the Situation Five as well as each measurement model of the Big Five facets separately. For the Big Five, we tested measurement models for each domain using factor scores of the facets as indicators. Several models were then tested for the complete Big Five model, including all domains: The first reflects the assumption underlying all Varimax rotated Big Five analyses. Consequently, a model with uncorrelated domains was tested (Costa & McCrae, 1995; Goldberg, 1990). The second model allows correlated Big Five factors and is more in line with findings attesting correlations between Big Five scores (Mount, Barrick, Scullen, & Rounds, 2005). Finally, we tested a model, which is based on research trying to explore the nature of the Big Five score intercorrelations. By now there

is a body of research, both correlation- and experiment-based, that supports the notion that covariance between the Big Five domain scores is due to social desirability (Bäckström, 2007; Bäckström, Björklund, & Larsson, 2009; Biderman, Nguyen, Cunningham, & Ghorbani, 2011; Klehe et al., 2012; Oshio, Abe, Cutrone, & Gosling, 2014; Schmit & Ryan, 1993; Ziegler & Buehner, 2009). Following this research, an additional bi-factor, reflecting social desirable responding, was added to the first model with uncorrelated personality domains. In this model, all other factors are controlled for this common variance.

With regard to model fit, personality structure models typically fail to reach the standards formulated by Hu and Bentler (1999) because of substantially lower loadings (Heene et al., 2011). Hopwood and Donellan (2010) analyzed several questionnaires and reported CFIs between .65 and .79 and RMSEAs between .09 and .13. These values have been used by other researchers as references (Thalmayer & Saucier, 2014). We will also do this, at the same time following advice to explore sources of misfit (Greiff & Heene, 2017). Using the same procedure as in Study 2, we computed item-parcels for each of the Situation Five measurement models. Using these item-parcels, we examined the higher-order structure of the Situation Five.

**Results.**

***Confirmatory factor analyses – Situation Five.*** The results of the confirmatory factor analyses are displayed in Table 2.3.1, lower half. The models for the Situation Five showed mediocre model fit. The RMSEA values ranged from .067 to .100. The CFI values were low, especially compared to the values of the Big Five. However, CFI values reflect how much unique (error) variance remains unexplained in each item. Since the Situation Five items – and any measure of situation characteristics – capture the perception of a specific situation (i.e., the vignette), some of this variance must be specific (idiosyncratic) and therefore unexplainable by the trait situation perception. This is also why correlated residuals in the measurement models of the Situation Five which would improve model fit are scarce and mainly due to using the same adjective[13]. In the end, we refrained from altering the models.

For the complete model of the Situation Five, we used item-parcels as indicators as described above (Figure 3, OSM). It had a reasonable model fit, $\chi^2 = 460.82$, $df = 86$, $p < .001$, CFI = .91, RMSEA = .10 [90% CI: .10 – .11], SRMR = .16. Two higher order factors were again found: The first called Vigor, loading on Briskness and Outcome-Expectancy, and the second called Burden, loading on Cognitive Load, Lack of Stimuli, and Psychological and Physical Load. However, due to the exploratory nature of these additional analyses, these higher order factors

[13] In fact, when running these analyses in *M*plus, no modification indices could be obtained for these models. However, *lavaan* still returned modification indices that could theoretically have been used to improve these models.

must be interpreted cautiously. This model fit is similar to the fit of the model in Study 1 and comparable to model fit of other assessment tools of situation characteristics (e.g., N. A. Brown et al., 2015; Rauthmann & Sherman, 2016a).

***Confirmatory factor analyses – Big Five.*** The factor structure of each of the Big Five domains that were based on the factors scores of the underlying facets could be confirmed (Table 2.3.1), and all models showed a good model fit (see OSM, Table C, and reliabilities, OSM, Table D). The CFI ranged from .88 to .96, and the SRMR from .05 to .06).

We tested three different theoretically plausible models for the structure of the Big Five personality traits. The first model, which assumed uncorrelated factors, had the worst model fit ($\chi^2$ = 3,616.40, *df* = 815, RMSEA = .093 [.090 - .096], CFI = .54, SRMR = .23). The second model, which allowed correlated factors, had a slightly better model fit ($\chi^2$ = 2,669.75, *df* = 805, RMSEA = .076 [.073 - .079], CFI = .69, SRMR = .09). The model that assumed a bi-factor structure fitted best, even though the CFI was still comparatively low: $\chi^2$ = 2,273.03, *df* = 773, RMSEA = .07 [.067 - .073], CFI = .76, SRMR = .08. However, as explained above, this is typical for structural models of personality. In fact, the values found here can be regarded as good compared to the values reported by Hopwood and Donellan (2010). We therefore conclude that the bi-factor model describes the Big Five factor structure best.

**Reliability**

Reliability estimates reported for the scores of a measurement tool should depend on its intended use (Ziegler, 2014c). If a test score is supposed to be used for status assessments (i.e., the current state of the person), an estimate of internal consistency or other point-estimates of reliability (e.g., split-half or parallel-test reliability) are warranted, whereas a prediction of future outcomes requires the presentation of test-retest-reliability estimate. For the B5PS scores, we present Cronbach's alpha, McDonald's Omega, as well as the test-retest reliability estimates.

**Sample and procedure.** Cronbach's alpha and McDonald's Omega were estimated based on the norm sample of the B5PS (*N* = 398) described above. For the test-retest-reliability, *N* = 96 participants completed the B5PS a second time (~65% return rate of questionnaires in second wave; time between measurement occasions: *M* = 5.7 months, $SD_{interval}$ = 1 month, mean age of this subsample = 48.11 years, $SD_{age}$ = 15.9 years, 40.6% male). The data-acquisition procedure of the norm-sample is described above.

**Results.** Results of the reliability analyses are displayed in Table F, OSM. Estimates for the internal consistencies of the Situation Five scores and the Big Five personality scale scores are very good, and so are factor reliabilities McDonald's Omega (alpha: .85 to .91, and omega: .65 to .96). The estimates for the 6-months test-retest reliabilities are also sufficient and range from *r* =

.58 (Lack of Stimuli and Cognitive Load) to $r$ = .75 (Outcome-Expectancy). The estimates for internal consistencies of the Big Five scores are slightly lower compared to the Situation Five scores (ranging from .72 to .92 for alpha, and for omega from .66 to .86). However, the test-retest reliabilities were higher compared to the Situation Five, ranging from $r$ =.63 to $r$ =.85.

**Validity**

Next to factorial validity, we will present evidence supporting the convergent and discriminant validity, as well as test-criterion correlations for the Big Five and the Situation Five test score interpretations.

**Sample and procedure.** The sample used for the validation studies is the norm-sample of the questionnaire ($N$ = 398). Participants did not only respond to the B5PS, but additional questionnaires were also administered for the purpose of validation.

***Big Five.*** The Big Five Structure Inventory (Big Five Struktur Inventar, BFSI, Arendasy, 2011) was used to obtain measures for convergent (Big Five) and discriminant (Situation Five) validity. The BFSI uses 300 adjectives and short sentences modeling the same facets and domains suggested by Costa and McCrae (1995). Test scores are estimated using item response theory. Participants were required to rate the items on a four-point Likert-type scale (1 = "strong disagreement" to 4 = "strong agreement").

***Work-Engagement.*** Work-engagement was assessed using the Utrecht-Work-Engagement scale (Seppälä et al., 2009), which assesses satisfaction with work based on one overall score and three facets (vigor, dedication, and absorption). Each facet is measured with 3 items using a seven-point frequency scale ranging from 0 = "never" to 7 = "always". Work-engagement shows a mediocre stability over an extended period of time (Mauno, Kinnunen, & Ruokolainen, 2007; Seppälä et al., 2009), and it is reasonable to assume that its scores are influenced by the current context or the perception thereof. It is further related to perceived job-resources (Schaufeli & Bakker, 2004). High work-engagement is "assumed to be the positive antipode of burnout" (Schaufeli & Bakker, 2004, p. 294). Thus, the ability to predict work-engagement is not only relevant from an occupational perspective, but also from a clinical perspective, for example in burnout-prevention.

***B5PS factor scores.*** For all subsequent regression and correlation analyses, we used the factor scores of the Big Five and Situation Five from the B5PS. The factor scores are computed based on the final structural models of Study 3. We used multiple linear regressions to predict all UWES scores (including the total composite score) with the Situation Five and the Big Five scores to investigate test-criterion correlations (concurrent validity).

**Results.**

***Convergent and discriminant validity.*** Table E (OSM) presents the bivariate correlations between the Big Five and Situation Five, both assessed with the B5PS. The correlations among the Big Five were moderate (the highest between Conscientiousness and Emotional Stability, $r$ = -.44). The correlations between the Situation Five were somewhat higher (up to $r$ = .70 between Briskness and of Outcome-Expectancy), which was reflected in their shared higher order factor.

The correlations between the Big Five and Situation Five scores, as measured with the B5PS, and the Big Five scores from the BFSI are displayed in Table 2.3.2. The Situation Five scores correlated weakly to moderately with the Big Five scores from the BFSI. The highest correlation was between Emotional Stability and Psychological and Physical Load ($r$ = -.46). Overall, Outcome-Expectancy correlated highest with the Big Five scores ($r$ = .15 to .43, whereas Lack of Stimuli and Cognitive Load correlated lowest with the Big Five scores ($r$ = -.08 to -.25, and $r$ = -.04 to 24, respectively). Briskness and Psychological and Physical Load correlated weak to moderately with the Big Five scores ($r$ = .13 to .31 and $r$ = -.46 to -.06, respectively).

The Big Five scores (B5PS) correlated highest with the corresponding personality trait score from the BFSI (see diagonal in the lower half of Table 2.3.2). Two convergent correlations were moderate (Conscientiousness, $r$ = .28, and Agreeableness, $r$ = .28), yet none of the other correlations of Conscientiousness and Agreeableness was higher. Correlations with other personality traits were small to moderate, indicating discriminant validity.

***Test-criterion correlations - UWES.*** We predicted the overall UWES total score and its three facet scores (dedication, absorption, and vigor) using the Big Five scores assessed with the B5PS in a first block and the Situation Five scores in a second block within four separate hierarchical regression analyses. Since the UWES subscales were highly correlated ($r_{\text{Vigor.Dedication}}$ =.78, $r_{\text{Vigor.Absorpion}}$ = .79, and $r_{\text{Dedication.Absorption}}$ = .85), we also computed the UWES total score. The results of these analyses are displayed in Table G, OSM (for bivariate correlations with UWES and other outcome variables, see Table B, OSM). The Big Five scores explained only little variance in all of the UWES scores. However, adding the Situation Five significantly increased the explained variance in all cases. Between 17.11% (Dedication) and 25.23% (Vigor) could be explained using the Situation Five and the Big Five scores combined.

Table 2.3.2
*Bivariate Correlations between the Big Five (BFSI) and the Big Five and Situation Five Scores*

| | BFSI Dimensions | | | | |
|---|---|---|---|---|---|
| **Situation Five** | ES | E | O | A | C |
| Outcome-Expectancy | .35*** | .43*** | .37*** | .15** | .35*** |
| Briskness | .13* | .25*** | .31*** | .17*** | .23*** |
| Psych. and Phys. Load | -.46*** | -.23*** | -.12* | -.06 | -.13* |
| Lack of Stimuli | -.25*** | -.11* | -.08 | -.17*** | -.20*** |
| Cognitive Load | -.04 | .09 | .13** | .11* | .24*** |
| | | | | | |
| **Big Five** | ES | E | O | A | C |
| Emotional Stability | **.42*** ** | .11* | -.08 | -.07 | -.03 |
| Extraversion | .08 | **.40*** ** | .16** | .07 | -.22*** |
| Openness/Flexibility | -.24*** | .01 | **.35*** ** | .05 | -.02 |
| Agreeableness/ Team-Orientation | -.03 | -.17** | -.04 | **.28*** ** | .05 |
| Conscientiousness | -.18*** | -.13* | -.03 | -.01 | **.28*** ** |

*Note.* ES = Emotional Stability; E = Extraversion; O = Openness; A = Agreeableness, C = Conscientiousness; *N* = 387 of the 398 participants completed the BFSI. Correlations in bold are convergent correlations between Big Five scores from different questionnaires.

**Discussion - Study 3**

**Factor structure.** The structure of the Big Five items, facets, and domains showed a very good model fit. It has to be noted, though, that starting at domain level, models were fit to the factor scores from the facet models, not the individual items due to the restricted sample size.

For the measurement models of the Situation Five scores, the model fit was not perfect, and if judged by conventional levels that are for example applied to intelligence tests (Hu & Bentler, 1999), in some cases unsatisfactory. However, as we have already mentioned above, an even lower model fit than in personality measures had to be expected in measures of situation perception (Horstmann, Ziegler, et al., 2018). As explained, this is due to the unique variance that can be attributed to the unique situation and thus a situation by person interaction resulting in idiosyncratic variance that is present in each item. It is common that the CFI is lower in models of situation perception due to this situation (vignette)-specific variance, which cannot be explained by the general situation perception of a person. We furthermore consider the high construct reliabilities as evidence for the cross-situation consistency of these factors.

**Reliability.** Reliability estimates of all scores were satisfying. The internal consistency and construct reliability estimates were very high across all ten scale scores. The test-retest reliability estimates for the Big Five scores (with the exception of Agreeableness: $r = .63$) were higher than the test-retest reliability estimates for the Situation Five scores. There are two possible explanations for this: First, the trait situation perception may change more across six months than trait

personality, indicating that it is in itself a more unstable construct. Another explanation would be that situation perception is stronger affected by momentary states, such as affect, and that even the responses to a vignette-based questionnaire such as the B5PS are subjected to this influence. This effect has already been shown for *in situ* situation perception (Horstmann & Ziegler, 2018; Parrigon et al., 2017; Sherman et al., 2015; Wilson et al., 2017), but it has yet to be examined if this is also true for situation perception as a trait. Importantly though, empirical support for the temporal and cross-situation stability of the Situation Five scores was found which is a necessary condition when considering situation perception as a personality trait. This in itself is an important extension of the current status of research on situation perception. Clearly, this has potential implications for clinical psychology in particular, as we will elaborate below.

**Validity.**

***Convergent and discriminant validity - Big Five.*** The convergent validity of the Big Five scores overall was good when compared with typical convergent correlations for Big Five measures (Pace & Brannick, 2010). It has to be noted that the facets of the B5PS did not reflect the facet structure suggested by Costa and McCrae as is the case for the BFSI. Such differences in test family have been shown to be important factors lowering convergent validity (Miller, Gaughan, Maples, & Price, 2011). Further, the discriminant validities of the Big Five scores were good. Although some domains (e.g., Openness and Extraversion, Conscientiousness and Extraversion, or Extraversion and Agreeableness) were substantially correlated, were not higher than the convergent correlations.

***Convergent and discriminant validity - Situation Five.*** The discriminant validity of the Situation Five can be considered satisfactory with regard to Big Five scores. However, the correlations among the Situation Five scores were, in some cases, rather high (see Table E, OSM). These correlations could be modeled using higher order factors. As we have mentioned earlier, other measures of situation perception show similar correlations among their scores (N. A. Brown et al., 2015; Parrigon et al., 2017; Rauthmann & Sherman, 2016a). First, measures of situation perception and measures of in situ behavior usually display a positive manifold (Baird, Le, & Lucas, 2006), and situation perception measures are not orthogonal. It is an unresolved question, however, why this is the case: If a situation is assessed, it may simply be the case that participants are not able to distinguish sufficiently between situation aspects, and rather evaluate situations on a broader, more abstract level. Given that the Situation Five were assessed over many different situations, it may also be the case that momentary aspects of the person influenced the judgement. Indeed, the adjectives from the Situation Five were previously shown to correlate with measures of affect (Horstmann & Ziegler, 2018).

At the time of data collections, no other measure for situation perception was readily available (Horstmann, Rauthmann, et al., 2018), and so far, no other measure of trait situation perception exists. However, previous research showed that the adjectives extracted for the Situation Five align with some of the DIAMONDS (Rauthmann et al., 2014) dimensions, further supporting construct validity of the test scores (Horstmann, Rauthmann, et al., 2018; Horstmann & Ziegler, 2018; Rauthmann & Sherman, 2018a).

The Situation Five additionally showed good discriminant validities with the Big Five scores from the BFSI. Although some correlations were rather high (~.40), the Situation Five could still be considered independent constructs. This is comparable to other findings that demonstrated a correlation between situation perception and personality trait scores (Rauthmann, Sherman, Nave, et al., 2015; Sherman, Nave, & Funder, 2013; Sherman et al., 2015). Personality may therefore be involved in shaping situation perceptions (Rauthmann, Sherman, & Funder, 2015b).

***Test-criterion correlations.*** The main purpose for the construction of an assessment tool of situation perception in an occupational context was the prediction of work-related outcomes. Work-engagement and its three subscale scores were used as outcomes for an initial validation, and the results show that the Situation Five scores explain significant variance in all UWES scores, whereas the Big Five did not. This is in line with models of job engagement, such as the Job-Demand-Control-(Support) model (J. V Johnson & Hall, 1988; Karasek, 1979). A general tendency to interpret situations and perceive their demands, which can directly reflect one's available resources in this situation, could be more specific for the prediction of job-related behavior than personality scores. Although this would need further specific validation (i.e., are situation perceptions correlated with available resources or the lack thereof?), it could be the stepping-stone for future research.

Additionally, earlier studies also reported very small correlations between the Big Five and work engagement (Kim, Shin, & Swanger, 2009). Moreover, it needs to be kept in mind that the factor scores used here were controlled for common method variance thereby also correcting test criterion correlations inflated by such variance sources.

With regard to the Situation Five, these results underscore that situation perception scores, assessed across a set of hypothetical situations in an occupational context, are a valid and so far unrecognized predictor for work-engagement. These initial results are encouraging and stress the importance of that construct. Moreover, the criterion chosen is of relevance for the clinical context just like the evidence for the feasibility of a combined measure for situation perception and personality traits is.

## Overall Discussion

### Theoretical Implications – What is the Construct being measured?

The theoretical implications of these findings are manifold. First and most important, we demonstrate that situation perception has a stable trait component: The perception of a fixed set of situations is not only stable, but also predicts relevant outcomes in a work-related context. Similarly, Sherman and colleagues (2015) previously reported effects of trait-like situation perception on behavior in daily life. Dimensions of situation perception should thus be considered as a construct that enrich the nomological net of personality traits and allow predicting additional variance in behavior and other outcomes.

Second, the assessment of situation perception as a trait is possible and feasible. One argument against the use of situations in personality research was that considering only the situation would allow to only make point predictions, or in other words explanations of behavior *in* that very same situation (Hogan, 2009). Assessing situation perception as a personality trait allows *predicting* behavior across situations and thereby overcomes this strong criticism.

### Applied Contexts – What is the intended use of the measure?

The first implication for an applied context is that situation perception could be considered for selection and training procedures in an occupational setting (Lievens, 2017). Situation perception might serve as an additional predictor and thus selection criterion, supporting the evidence presented by Rockstuhl and colleagues (Rockstuhl et al., 2015).

Second, situation perception and the measurement thereof are suitable tools for clinical psychology. As stated earlier, work-engagement is closely tied to clinically relevant outcomes such as burn-out or depression (Schaufeli & Bakker, 2004; Schaufeli, Salanova, Gon Alez-ro, & Bakker, 2002). Assessing a person's general perception of situations can therefore be used as an outset for intervention or even therapy. If the Situation Five indeed reflected the *perceived* availability of job-resources or demands, and this perception then caused burnout, situation perception could be a link between the two. Situation perception of *in situ* ratings in specific situations might further be used to shape and change environments such that they are evaluated more positively by their perceivers. This has so far not been examined or considered but may be a valuable avenue for further research.

The B5PS can furthermore be understood as a proof of concept and blueprint for the creation of assessment tools for situation-contingent perception and behavior. For example, the situation vignettes used could easily be replaced with critical incidents provided by clinicians or therapists of their patients' lives. Such a situation rating tool for clinically relevant situations may suit a larger audience of participants, not only in an occupational setting. The availability of

situations that have received a normative rating can shed new lights on the processes involved in mental malfunctions and the occurrence of pathological behavior.

**Limitations**

**Model Fit.** The model fit of the presented measurement models of the Situation Five in the B5PS is not optimal. Especially the low CFI might be cause for concern. As mentioned above, this is rather typical for personality measures (Hopwood & Donnellan, 2010). The CFI is indicative of unexplained unique variance in manifest variables, and as such, it has to be concluded that the ratings in situation perception cannot completely be explained by situation perception as a trait (Horstmann, Ziegler, et al., 2018). The higher order factor models for the Situation Five were furthermore modeled using item parcels. This approach allows testing the structure of the Situation Five (Study 2 and Study 3), without having to sample too many participants. At the same time, model misfit can be masked using item-parcels (Little et al., 2002). Although the higher order factor structure of the Situation Five could be confirmed across two independent samples and using a different set of stimuli, these results must be interpreted cautiously. Most importantly, it should not be concluded that situations are primarily perceived on two dimensions only.

**Population.** One limitation of the assessment tool is that it only refers to white-collar workers their occupational environments. Other contexts, such as manual work, were not considered and the generalizability to these contexts may therefore be limited. Furthermore, the construction of the taxonomy initially relied on students who chose to report a self-selected situation from their previous day. Although this may lead to higher ecological validity (i.e., only situations are sampled that are commonly experienced), other, more extreme situations might be overlooked: The structure of situations that are rarely experienced during everyday life (e.g., a burning house, a surgery, a funeral) could differ from everyday self-selected situations, and the Situation Five taxonomy might therefore not be applicable to such contexts.

## General Conclusion

The current research article contributes several aspects to the research on person-situation transactions. First, the development of a new set of dimensions of situation characteristics provides further insight into the development of the structure and content of situations. Situation research may bring forward other taxonomies, and linking different taxonomies together examining their convergent and discriminant validity, which might be different on a trait and state level (Rauthmann et al., 2018), is the next task at hand.

Second, the B5PS is a measurement tool that allows assessing trait situation perception and Big Five personality scores of a person simultaneously in a reliable and valid form. This demonstrates that valid results based on ratings of vignettes can be obtained (Ziegler & Ziegler, 2015), thereby adding another method for future research and application of situation taxonomies. Besides the here shown application in organizational settings, this seems especially promising in clinical settings. A measure capturing situation perceptions and manifestations of pathological traits might contribute important insights for diagnoses and therapies.

Third, our research shows that situation perception has a stable trait component. Applications of other situation perception taxonomies could benefit from the recognition of these trait components of situation perceptions. Situation research has taken a new turn in the last years; however, the current results and theoretical developments clearly indicate that situation research still holds many secrets.

## 2.4 Replicable Dimensions of Situation Perception

In the last few years, several situational taxonomies have been published (Horstmann, Rauthmann, et al., 2018; Rauthmann et al., 2014; Yang et al., 2006). As elaborated above (see *Assessing Situational Information*, section 2.2), there are many different approaches to developing situational taxonomies, and each taxonomy may serve a different purpose and may be applicable to different situations.

In few cases, however, authors of situational taxonomies have suggested dimensions that (a) strongly resemble one another, yet were named differently, or (b) were given the same name, but referred to different content. This phenomenon is called jingle-jangle-fallacy (T. L. Kelley, 1927). Jingle-fallacies refer to cases in which two different constructs have the same name, and jangle-fallacies refer to cases in which the same construct bears two different names. These cases may occur during test construction and will remain undetected unless a nomological network of the constructs is established. A nomological network defines the nature of a construct by relating it to other constructs, for example, by means of theoretical argumentation and empirical validation (Cronbach & Meehl, 1955). For dimensions of situation perception, such a network did not yet exist. Given the novelty of this these dimensions, this is not surprising. As Cronbach and Meehl wrote in 1955, p. 291:

> "When a construct is fairly new, there may be few specifiable associations by which to pin down the concept. As research proceeds, the construct sends out roots in many directions, which attach it to more and more facts or other constructs."

The goal of our book chapter *Measurement of Situational Influences* (Horstmann, Rauthmann, et al., 2018) was to start such a nomological net and connect dimensions of situation perception to one another, possibly identifying overarching dimensions of situation perception. Comparable to the Big Five in personality psychology (Funder, 2001; Goldberg, 1990; John et al., 2008), the development of a nomological net and the discovery of replicable dimensions is likely to advance the field of situation research in many ways.

### *2.4.1 Book Chapter Summary: Measurement of Situational Influences*

**Background.** The person-situation debate and subsequent theoretical developments in the realm of personality psychology called for the description and measurement of situations. Beginning in 1962, the first systematic taxonomies for the description of situations were published (e.g., Endler, Hunt, & Rosenstein, 1962; Krause, 1970; Magnusson, 1971; Moos, 1973; Price, 1974). Still, Hogan wrote in 2009 that the "person situation debate is an empty exercise, because the perpetrators cannot define or measure situations" (Hogan, 2009, p. 249). In our book chapter, we reviewed the causes for this depressing assessment of the state of the field and gave an overview over existing situational taxonomies. Between 1962 and 2010 (Endler et al., 1962; Sherman et al., 2010), 21 different situational taxonomies had been suggested. However, none of these taxonomies had provided a comprehensive measurement tool that allowed assessing situations and could thus be hardly applied in a broader scheme. Although several definitions of situations existed, a way to measure situations had not yet been developed.

**Five New Taxonomies.** Rauthmann and colleagues published the first extensive taxonomy that describes situations with regard to their psychologically relevant features, or situation characteristics (Rauthmann et al., 2014), together with a tool to assess these characteristics (Rauthmann & Sherman, 2016a, 2016c). Subsequently, four other situational taxonomies were published: The Situation Five (Ziegler, 2014a; Ziegler et al., submitted), the Situational Affordances and Adaptive Problems scale (N. A. Brown et al., 2015), the CAPTION taxonomy (Parrigon et al., 2017), and the Social Interdependence Scale (Gerpott et al., 2017). The taxonomies were developed based on different theoretical approaches, different goals in mind, and in different cultures and populations. At the same time, though, the dimensions of situation perception described in these taxonomies showed remarkable theoretical and empirical overlap.

**An Overarching Structure of Situations.** All four situational taxonomies that were developed after the DIAMONDS used the DIAMONDS as convergent and discriminant measures. These correlations, but also theoretical consideration, allowed us to develop an initial nomological net of situation perception dimensions (Rauthmann & Horstmann, 2017). This network highlights links between dimensions from different taxonomies and points towards potential jingle-jangle-fallacies. These are cases in which two dimensions have been named differently, even though they measure the same construct, or have been named similarly, even though they describe two very different constructs.

Finally, we identified six overarching dimensions of situation perception: Threat, Stress, Tasks, Processing, Social Positive, and Mundane. Notably only the CAPTION framework covers all six dimensions of situation perception thus far. Yet, future studies must seek to validate and test this proposed structure, and it is likely that other emerging situational taxonomies will refine or extend this network (e.g., Oreg, Edwards, & Rauthmann, in preparation). Furthermore, it is possible that some dimensions may show a facetted structure, comparable to personality domains and facets.

**Conclusion.** This book chapter can be seen as the first, careful approach to an integration of existing measures of situation perception. Such an overarching framework will not only allow constructing better measures for situation perception dimensions and avoiding jingle-jangle-fallacies, but it will also contribute to a cumulative science by making research findings on situations comparable across studies.

### *2.4.2 Book Chapter: Measurement of Situational Influences.*

**Measurement of Situational Influences**

Kai T. Horstmann[1], John F. Rauthmann[1,2], Ryne A. Sherman[3]
[1]Humboldt Universität zu Berlin, Germany,
[2]Wake Forest University, United States,
[3]Texas Tech University, United States.

This book chapter was first published as:

Horstmann, K. T., Rauthmann, J. F., & Sherman, R. A. (2018). Measurement of situational influences. In V. Zeigler-Hill & T. K. Shackelford (Eds.), The SAGE Handbook of Personality and Individual Differences. SAGE Publications.

This chapter provides an overview on the current state-of-the art of assessing situational influences on behavior or personality expression. We first briefly review the history of situational assessment and recent developments of situational taxonomies. Extant taxonomies are then compared and integrated on a theoretical basis. Lastly, we provide recommendations for future research and discuss challenges of assessing situations.

## Reasons for Assessing the Situation

The first question that should be briefly answered is: Why should situational influences be assessed at all? The answer is that prediction of behavior is a core interest of psychology. Various aspects of the person predict different kinds of behavior, including aggressive behavior (Bettencourt, Talley, Benjamin, & Valentine, 2006; Lämmle, Oedl, & Ziegler, 2014), behavior in interpersonal relations (Cuperman & Ickes, 2009; Leckelt, Küfner, Nestler, & Back, 2015; Maaß & Ziegler, 2017), health-related behavior (Hall, Fong, & Epp, 2013), career success (Barrick et al., 2001; T. A. Judge, Heller, & Mount, 2002; Ones, Mount, Barrick, & Hunter, 1994; Ozer & Benet-Martínez, 2006; Schmidt & Hunter, 1998; Ziegler, Danay, Schölmerich, & Bühner, 2010), and behavior in everyday life (e.g., Sherman et al., 2015; S Vazire & Mehl, 2008).

However, personality is a not a perfect predictor of behavior. There are many reasons why a predictor (e.g., personality) may not be perfectly correlated with a criterion (e.g., behavior). Besides methodological (e.g., lack of reliability) and conceptual issues (e.g., differing levels of abstractness), a pervasive issue is that behavior is multiply determined. Although people display impressive amounts of consistency (both within and between themselves; Fleeson & Noftle, 2008a), people also vary in their behavior across situations and time (Bem & Allen, 1974; Fleeson, 2004; Fleeson & Jayawickreme, 2015). To understand, explain, and predict such variability, we not only need personality (Sherman, Nave, & Funder, 2010), but also knowledge about situations. Thus, assuming that variability in behavior over time and occasion is not just measurement error (Epstein, 1979, 1983a), considering the situation the person is currently in can enhance behavioral prediction. Indeed, recent years have seen increasing interest in the situation as a predictor of behavior, and there is overwhelming consensus and evidence that the situation is in fact a good predictor of behavior (Deinzer et al., 1995; Fleeson, 2004; Fleeson & Noftle, 2009; Funder, 2006; Furr & Funder, 2004; J. A. Johnson, 1999; Mischel, 1977; Mischel & Peake, 1982; Mischel & Shoda, 1995; Rauthmann, 2012; Sherman et al., 2010; Shoda, Mischel, & Wright, 1994; van Mechelen & De Raad, 1999; Wagerman & Funder, 2009). Further, most contemporary personality theories readily acknowledge the importance of situational influences, such as Whole Trait Theory (Fleeson, 2001, 2004; Fleeson & Jayawickreme, 2015), Latent-State-Trait theory (Steyer et al., 2015, 1999), Cognitive Affective Personality System (Mischel & Shoda, 1995), Trait Activation Theory (Tett &

Burnett, 2003; Tett & Guterman, 2000), and even the Five Factor Theory (McCrae & Costa, 2008). All of these models and theories include the role of external influences on behavior, beyond the influence of personality.

However, basic questions regarding the definition of a situation (i.e., what is a situation?) and the underlying process of situational influence on behavior (how does it influence behavior?) are still unresolved. This chapter will therefore focus on three elements of situational assessment:

1. The historical background and development of situational assessment and its theory.
2. Novel developments in situational assessments and their possible integration.
3. Issues of situational assessment and its application as well as future developments.

## A Brief History of Situational Assessment

The idea that the physical environment of a person can be used to predict his/her behavior in that situation was prominently featured in Kurt Lewin's work (e.g., Lewin, 1936). He stated that *B*ehavior is a function of the *P*erson and the *E*nvironment, B = f(P, E). This formula has later been re-interpreted as the personality triad (Funder, 2006, 2009), consisting of personality, situations, and behavior – with the conceptual idea that to understand any member of that triad one would need the other two. The person-part of the equation has traditionally received much more attention in psychology than the situation- or environment-part although social psychology is ostensibly concerned with situational influences (Krueger, 2009). However, the situation as a potent predictor of behavior has been (re-)discovered in response to a now famous book by Walter Mischel (Mischel, 1968), in which he claimed that personality can only predict a limited share of variance in behavior (Mischel, 2009). The term *personality coefficient* referred to an upper ceiling of $r \approx .30$ when it comes to the prediction of behavior by simple (one-time self-report) personality measures. Mischel's main critique was that the field of personality psychology focused too much on an "unconditionalized" conceptualization of traits (Mischel & Shoda, 1994, p. 156) which would be largely devoid of predictive power.[14]

Mischel's (1968) book and the subsequent person-situation debate (Fleeson & Noftle, 2008b; Funder, 2009) sparked attention for the situation and its measurement. There were – and still are – two basic ways to react to this critique. The first reaction is to discard it and consider deviations from predicted trait-relevant behavior as pure measurement error. Such error would cancel out over time and across measurement occasions (e.g., Epstein, 1983a). However, variations

[14] Note that .30 is actually meaningful (e.g., Ozer, 1985) and quite large as compared to other effect size sin the literature (e.g., Gignac & Szodorai, 2016; Hemphill, 2003). Further, the reasons for such a seemingly low correlation coefficients have been discussed elsewhere (e.g., Schmitt, 2009).

in behavior do not seem to be random error, but are actually meaningful (e.g., Andersen & Thorpe, 2009; Fleeson & Noftle, 2009; Fournier, Moskowitz, & Zuroff, 2009; Furr, 2009b; Geukes et al., 2016; Sherman et al., 2010, 2015; Smith, Shoda, Cumming, & Smoll, 2009; Van Mechelen, 2009). The alternative reaction is to embrace the critique and embark on a mission to make situational influences visible (Ziegler & Horstmann, 2015) and useful for the prediction of behavior. To do so first requires an understanding of which situational variables will be useful in predicting behavior and how they can (or should be) measured. To get an overview of what has already been proposed, we will provide a brief review of situational taxonomies.

**Situational Taxonomies.** In Table 2.4.1, we list 26 prominent situational taxonomies that have been developed in temporal sequence. There are largely two clusters of published taxonomies. The first cluster appeared in the early 1970s, the second one in the first years of the new millennium. The first cluster (i. e., Battistich & Thompson, 1980; Forgas, 1976; King & Sorrentino, 1983; Krause, 1970; Magnusson, 1971; Moos, 1973; Nascimento-Schulze, 1981; Pervin, 1976; Price, 1974; Price & Blashfield, 1975; Van Heck, 1984) could be seen as a direct answer to Mischel's (1968) book. However, this surge of situational taxonomies did not end the person-situation debate. For example, Hogan (2009) argued that there were no effective strategies for measuring situations in a special issue of the *Journal of Research in Personality* concerning the person-situation debate that was published in honor of the 40th anniversary of Mischel's (1968) book. The entire abstract of his paper read *"The person–situation debate is an empty exercise because the perpetrators cannot define or measure situations"* (Hogan, 2009). Although several situational taxonomies had been proposed by then, no measurement tools were available at the time.

Table 2.4.1
*Existing Situation Taxonomies*

| **Work by** | **Summary** | **Number and type** | **Situational Information assessed** | **Tool** |
|---|---|---|---|---|
| Endler, Hunt, & Rosenstein (1962) | Rating of 11 anxiety-evoking situations, factor analyses | 3 clusters | Classes | Yes |
| Krause (1970) | Theoretical overview, definition of social situations | 7 clusters | Classes | No |
| Magnusson (1971)* | 36 situations sampled, similarity judgments of situations | 5 factors | Characteristics, Classes | No |
| Moos (1973) | Theoretical overview | 6 broad dimensions | Characteristics, Classes | No |
| Price (1974) | Ratings of fixed situations, cluster analysis | 4 clusters | Classes | No |
| Price & Blashfield (1975) | Ratings of 455 distinct settings of a town, factor and cluster analysis | 9 factors and 12 clusters of behavioral settings | Classes | No |
| Forgas (1976)* | 25 everyday situations sampled, similarity judgment of situations | 2 to 3 factors | Characteristics, Classes | No |
| Pervin (1976)* | Up to 29 everyday situations, judged on situational features | 6 factors | Classes | No |
| Battistich & Thompson (1980)* | 30 everyday situations, similarity judgments of situations | 4 factors | Characteristics, Classes | No |
| Nascimento-Schulze (1981) | Ratings of 12 descriptions of interpersonal situations, principal component analysis | 2 factors | Characteristics | No |
| King & Sorrentino (1983) | Rating of 40 situations, multidimensional scaling | 7 dimensions | Characteristics | No |
| Van Heck (1984, 1989)* | 248 nouns, judged on situational features | 10 factors | Characteristics, Classes | No |
| Eckes (1995)* | 30 everyday situations, judged on situational features | 9 clusters | Characteristics | No |
| Ten Berge & De Raad (2001)* | 132 situations generated from personality traits, rated trait expressions in these situations | 5 factors (of behavior in situations) | Classes | No |
| Ten Berge & De Raad (2002)* | 237 situations generated from personality traits, | 4 factors | Classes | No |

| | | | | |
|---|---|---|---|---|
| | ability of dealing with that situation | | | |
| Kelley at al. (2003) | Definition of interpersonal situations | 6 dimensions | Characteristics, Classes | No |
| Edwards & Templeton (2005) | Self-reported situations, rated on list of adjectives | 4 factors | Characteristics | No |
| Yang, Read, & Miller (2006) | Lexical approach to Chinese idioms, cluster analyses | 3 cluster | Characteristics, Classes | No |
| Saucier, Bel-Bahar, & Fernandez (2007) | 7000 randomly reported situations | 4 broad domains of variables | Cues | No |
| Fournier, Moskowitz & Zuroff (2008; 2009) | Event contingent recordings of behavior in situations, situations defined by behavior | 4 dimensions | Characteristics, Classes | No |
| Sherman, Nave, & Funder (2010) | Students' self-reported and randomly selected situations | NA | Classes | No |
| Rauthmann et al. (2014) | Rating of randomly selected situations of the previous day | 8 factors | Characteristics | Yes |
| Ziegler et al. (2014a; Ziegler et al., submitted) | Lexical approach, ratings of randomly selected situations, factor analysis | 5 factors | Characteristics | Yes |
| Brown, Neel, & Sherman (2015) | Ratings of self-reported situations from students, factor analysis | 7 factors of situational affordance | Characteristics | Yes |
| Parrigon, Woo, Tay, & Wang (2017) | Lexical approach, combined with ratings of situations and factor analyses | 7 factors | Characteristics | Yes |
| Gerpott, Balliet, & De Vries (2017) | Factor analyses of ratings on self-reported situations | 5 factors | Characteristics | Yes |

*Note.* Situational Information assessed is based on the definition of Cues, Characteristics, and Classes provided by Rauthmann (2015). Entries are sorted by date of publication in ascending order. *based on Table 1 from Yang, Read, and Miller (2006). Tool indicates if a measurement tool is available.

The second cluster of situational taxonomies (i.e., Brown, Neel, & Sherman, 2015; Eckes, 1995; Edwards & Templeton, 2005; Fournier, Moskowitz, & Zuroff, 2008; Fournier et al., 2009; Gerpott, Balliet, & de Vries, 2017; Kelley et al., 2003; Parrigon, Woo, Tay, & Wang, 2017; Rauthmann et al., 2014; Saucier, Bel-Bahar, & Fernandez, 2007; Sherman, Nave, & Funder, 2010; Ten Berge & De Raad, 2001, 2002; Yang, Read, & Miller, 2006) finally brought forth not only taxonomies, but some measurement tools as well (e.g., Brown et al., 2015; Gerpott et al., 2017; Parrigon et al., 2017; Rauthmann & Sherman, 2016a; Ziegler, 2014a). Thus, 80 years after Lewin initially proposed recognizing the influence of the situation, the existence of both situational taxonomies and measurement tools now provide the ability to directly quantify the role of the environment when it comes to the prediction of behavior. We will take a closer look at selected works from the second cluster of situational taxonomies throughout this chapter.

## The Theoretical Underpinnings of Situational Assessment

To assess situational information, we need to first establish whether we seek to measure the perceiver's interpretation of the situation's psychological characteristics (subjective approach) or the actual physical cues of the situation (objective approach).

### The Objective Approach

Elements that are physically present and constitute the situation are referred to as *situation cues* (Rauthmann, Sherman, & Funder, 2015b). Cues give the answer to five simple *W*-questions. *Who* is with you? *Which* objects are around you? *What* is happening? *Where* are you? *When* is this happening? Cues constitute the situation and do not need to be interpreted by a perceiver, that is, they are present whether or not someone observes the situation or takes part in it. This approach to situational assessment is called the objective approach (Furr & Funder, 2004; Rauthmann, Sherman, & Funder, 2015b), and situations assessed under the objective approach have no psychological meaning in and of themselves – they simply exist.

There are numerous examples of studies that used modifications of situational cues as their independent variable in an experimental design (Funder & Ozer, 1983; Furr & Funder, 2004; Horstmann & Ziegler, 2016; Krueger, 2009; F. D. Richard et al., 2003). Altering situational cues, sometimes only slightly, and examining the resultant change in behavior is probably *the* most widely used research design of social psychology.

In a meta-meta-analysis of social psychology, including over 8 million people and 25,000 studies, Richard and colleagues (2003) showed that the average effect found in social psychology was $r = .21$. Most of these studies focused on social interactions or other persons' effects on behavior. Thus, the topics identified by Richard et al. are mostly of a social nature (e.g., group processes, helping behavior, or social cognition; see Richard et al., 2003, Table 1). However,

without a taxonomy of situations or situational influences, a systematic integration of these effects is impossible and gains in cumulative knowledge may be limited.

The objective approach has significant shortcomings for the assessment of situational influences. First, it is simply impractical. One of us once asked a social psychologist, "What is it exactly that matters about the situation?" The psychologist's response: "All of it." Although this answer points to the notion that all situational cues might influence behavior, it also points to the impossibility of advancing our psychological understanding of situations, if such a view is taken seriously. Listing and quantifying all cues (e.g., temperature, background noise, ambient light, and the exact location, color, texture, smell, taste, sound, etc. made by every object) in a situation would take a tremendous amount of time and effort, if it could even be achieved. Further, if a single change in a single cue is intended to demarcate a different situation, we would quickly realize that different situations are hard or nearly impracticable to detect.

However, the strongest objection to the assessment of situations via situational cues is that (a) in the same situation, not every cue is perceived by every individual and (b) even if a cue may be perceived by all individuals, they may have different interpretations of that cue (e.g., Edwards & Templeton, 2005; Lewin, 1936; Rauthmann, 2012; Rauthmann, Sherman, & Funder, 2015b). This is why the subjective approach, detailed below, is gaining traction in recent literature.

**The Subjective Approach**

The subjective approach focuses on how a situation is perceived by an individual as experiences and perceptions of situations will determine what a person thinks, feels, wants, and how he/she acts within it (Rauthmann et al., 2015b). All recent situational taxonomies as well as their assessment tools were developed around this idea (N. A. Brown et al., 2015; Gerpott et al., 2017; Parrigon et al., 2017; Rauthmann et al., 2014; Ziegler, 2014a) and focus thus on the *psychologically important characteristics* of the situation. Situation characteristics are comparable to traits of persons as they describe the rather broad dimensions used to differentiate situations.

However, assessing situations via their perceived characteristics requires that perceivers rate situations on these characteristics. Variance of these ratings is thus not only due to the influence of the situation, but also due to the perceiver (Rauthmann, 2012). For example, most people would agree that sitting in a café and enjoying a drink with friends is more pleasant than cleaning one's house. Of course, some people may hold a different view on this, which needs to be explicitly considered when seeking to assess the situation in its completeness. In contrast to the objective approach where the assessment of the situation is in principle independent from any perceiver, the subjective approach does not assume this independence. This is reflected in the three principles of situation research proposed by Rauthmann and colleagues (2015b): the Processing

Principle, the Reality Principle, and the Circularity Principle. The Processing Principle states that psychological experiences of the situation matter and drive behavior. The Reality Principle states that situations have three different types of reality (physical: what actually exists; consensual: what people agree on; idiosyncratic: what only one person perceives). The Circularity Principle states that persons and situations are always conflated if the situation is assessed via the perception of a person. Each of these principles will be reviewed in more detail later in the chapter.

To exclusively measure situational influences in terms of situation characteristics, it is necessary to obtain ratings of a situation by more than one rater. Each rater will have his/her personal perception of the situation (idiosyncratic reality), but these perceptions will, to some extent, co-vary with others' perceptions (consensual reality), which is a good approximation of the socially agreed upon characteristic of the situation within a socio-culture. Additionally, these perceptions are usually based on actual cues in the environment (physical reality).

These principles of situation research can guide the assessment of situations and their influences. Recent situational taxonomies have implicitly or explicitly embraced these principles. Based on subjective ratings of situations, all recent taxonomies we present below focus on the assessment of situational characteristics. It should be noted that these taxonomies are the few and first that actually provided measurement tools and sought to devise a relatively integrated taxonomic system.

**Five Taxonomies of Situation Characteristics**

An overview of five recent situational taxonomies is provided in Table 2.4.2. These taxonomies are: CAPTION (Parrigon et al., 2017), DIAMONDS (Rauthmann et al., 2014), Situational Affordances and Adaptive Problems (SAAP; N. A. Brown et al., 2015), Social Interdependence Scale (SIS; Gerpott et al., 2017), and the Situation 5 (Ziegler, 2014a; Ziegler et al., submitted).

Table 2.4.2
*Comparison of Five Taxonomies of Situation Characteristics*

| **Taxonomy & Dimensions** | **Description** | **Sample Items** | **Tradition** |
|---|---|---|---|
| **CAPTION**<br>**7 dimensions** | by Parrigon, Woo, Tay, and Wang (2017) | | Lexical approach (USA) |
| Complexity | Describes how much a situation is "marked by learning, in-depth thought and investigative exploration" (Parrigon et al., 2017, p. 33). | *analytical, academic* | |
| Adversity | Describes to which extent situations are difficult, depleting and cost physical and psychological resources. | *stressful*<br>*fatiguing* | |
| Positive Valence | Describes to what extent a situation is positive or interpersonally warm, i.e., including also interpersonal love and affection. | *heartwarming*<br>*cherished* | |
| Typicality | Describes situations that are usual, regularly experienced and not new. | *typical*<br>*regular* | |
| Importance | Describes how well the situation is perceived to be important for the fulfillment of a certain, personal goal. | *effective*<br>*useful* | |
| Humor | Describes to what extend the situation is amusing or playful. | *wacky*<br>*mischievous* | |
| Negative Valence | Describes situations that are perceived to be menacing or threatening. | *repulsive*<br>*despicable* | |
| **DIAMONDS**<br>**8 dimensions** | by Rauthmann, Gallardo-Pujol, Guillaume, Todd, Nave, Sherman, Ziegler, Jones, and Funder (2014) | | Riverside Situational Q-Sort |
| Duty | Describes to what extent work is to be done and tasks need to be completed. | *A job needs to be done* | |
| Intellect | Describes to what extent a situation is perceived to require intellectual engagement or cognitive demands. | *Situation affords an opportunity to demonstrate intellectual capacity* | |
| Adversity | Describes that problems may arise, blaming and criticism are perceived. It describes a situation of threat. | *Being blamed for something* | |
| Mating | Describes to what extent people are present who are could be romantic or sexual partners. | *Potential partners are present* | |
| pOsitivity | Describes to what extent the situation is fun, clear, or pleasant. | *Situation is potentially enjoyable* | |
| Negativity | Describes to what extent a situation may lead to any negative feeling or anxiety. | *Situation is potentially anxiety-inducing* | |
| Deception | Describes to what a situation is perceived to contain mistrust, deception, lies and betrayal. | *It is possible to deceive someone* | |
| Sociality | Describes to what extent a situation is perceived to contain social interaction, other people and the possibility to communicate with them. | *Social interaction is possible* | |

| **SAAP**<br>**7 dimensions** | by Brown, Neel, and Sherman (2015) | | Evolutionary Theory |
|---|---|---|---|
| Self-protection | Describes to what extent the situation affords the need to show defensive and protective behavior. | *I need to protect myself* | |
| Disease Avoidance | Describes to what extent there is a need to avoid catching a disease in a certain situation. | *It is important to avoid visibly sick people* | |
| Affiliation | Describes to what extent the situation affords getting along with other people. | *Getting along with others is important* | |
| Status | Describes to what extent the situation is relevant to gaining status and respect from others. | *It is important to gain respect from others* | |
| Mate Seeking | Describes to what extent there is the need or opportunity to find a partner suitable for mating. | *There is an opportunity for a 'one night stand'* | |
| Mate Retention | Describes to what extent keeping and caring for ones mate is relevant in a situation. | *It is important to keep my romantic partner happy* | |
| Kin Care | Describes to what extent a situation affords caring about other genetically related others. | *It is important to help my child* | |
| **Situation 5**<br>**5 dimensions** | by Ziegler, Horstmann, and Ziegler ((submitted) | | Lexical Approach (Germany) |
| Outcome-Expectancy | Describes to what extent a person perceives a situation to be relevant for the achievement of his or her personal goals. | *I assess the situation as full of potential.* | |
| Briskness | Describes to what extent a situation is perceived to be stimulating and encourages enactment. | *I assess the situation as lively.* | |
| Psychological and Physical Load | Describes to what extent a situations is perceived to be mentally and physically challenging, stressful and burdening. | *I assess the situation as burdensome.* | |
| Lack of Stimuli | Describes to what extent a situations provides little or no information, is seen as having little importance and dull. | *I assess the situation as boring.* | |
| Cognitive Load | Describes to what extent the situations occupies mental capacity, requires thinking, engages problem-solving abilities, and is difficult to apprehend. | *I assess the situation as challenging.* | |

| **Social Interdependence Scale** | by Gerpott, Balliet, and De Vries (2017) | | Inter-dependence Theory |
|---|---|---|---|
| Interdependence | Describes to what extent each person's outcomes depend on other person's behavior. | *What each of us does in this situation affects the other.* | |
| Conflict | Describes to what extent a conflict of interests exists, good outcome for one person results in bad outcome for another person. | *The other prefers different outcomes than I do in this situation.* | |
| Power+ | Describes to what extent one individual person has the influence on the outcome while at the same time others do not have an influence. | *Who do you feel was most in control of what happens in the situation?*[1] | |
| Future Interdependence | Describes to what extent behavior in the current situation towards each other influences future behavior towards each other. | *How we behave now will have consequences for future outcomes.* | |
| Information Certainty | Describes to what extent a person knows other persons preferences in a situation and how much one person's result influences that of another person. | *We both lack knowledge about what the other wants.* | |

*Note.* + = these items are answered on a Likert-type response scale with the endpoints 1 = *Definitely the other* and 5 = *Definitely Myself*. All other items are answered on a standard Likert-type response scale (*Strongly Agree – Strongly Disagree*).

**Differences and Similarities between the Taxonomies**

The five situational taxonomies have different backgrounds and construction processes. The CAPTION and the Situation 5 models are based on the lexical approach. For the construction of both taxonomies, adjectives potentially useful for the description of situations were sampled. This corpus of adjectives was then reduced to a manageable number of adjectives by independent raters. Finally, participants rated self-reported situations on these adjectives, and these ratings were used in exploratory and confirmatory factor analyses to extract the final number of situation dimensions. The SIS and SAAP on the other hand are based on different theories, the first on interdependence theory (H. H. Kelley et al., 2003; H. H. Kelley & Thibaut, 1978), the second on evolutionary theory. In both cases, items were generated that could potentially measure the theoretically plausible dimensions. Participants were then required to rate self-reported situations on these items. Using exploratory and confirmatory factor analyses in independent samples, these ratings were condensed to the final dimensions of situational perception. The DIAMONDS are based on the first validated tool for the systematic assessment of situations, the Riverside Situational Q-Sort (Sherman et al., 2010; Wagerman & Funder, 2009). The RSQ is itself based on the California Adult Q-Sort, a measure to assess personality (Block, 1978). The DIAMONDS therefore assess situational characteristics that relate to particularly to personality expressions (Parrigon et al., 2017; Rauthmann, Sherman, & Funder, 2015a; Sherman et al., 2015).

Besides these differences in the approach taken (lexical vs. theoretical vs. atheoretical), there are also other notable differences between these taxonomies. One is the underlying sample that was used for the construction (e.g., US students vs. German representative sample). Another and more relevant difference can occur due to differences in the intended use of each measure or taxonomy (Horstmann, Ziegler, et al., 2018; Ziegler, 2014c). The DIAMONDS and the CAPTION models focus on the description of everyday, broad situations, whereas the SIS and SAAP focus more on social situations and situations that require or allow interaction. The Situation 5 focus on situations in a professional occupational setting and may exclude, for example, mating dimensions.

These five situational taxonomies could potentially be integrated into one broad situational taxonomy, comparable to the Big Five in personality trait research (John & Srivastava, 1999; Pervin & John, 1999). However, a problem to a well-integrated joint taxonomy is the current "jingle-jangle jungle" of situation research (Rauthmann, Sherman, & Funder, 2015b, p. 372).

**Jingle-Jangle Fallacies**

The jingle-jangle fallacies describe two unfortunate cases that can occur if a construct (or a scale) is labeled. The jingle fallacy occurs when different constructs receive the same label, whereas the jangle fallacy occurs when the same construct receives different labels. Concerning situation perception dimensions, both fallacies seem to have occurred so far. For example, it is likely that Adversity (CAPTION) and Negativity (DIAMONDS) actually are labels for the same construct (jingle fallacy), whereas Adversity (CAPTION) and Adversity (DIAMONDS) are the same label for the two different constructs even though they have the same label (jangle fallacy). Adversity (CAPTION) refers to stressful and tiresome situations (captured by Negativity in the DIAMONDS), whereas Adversity (DIAMONDS) refers to situations in which someone is blamed or criticized (captured by Negative Valence in CAPTION).

Another example for a jingle fallacy could be the dimensions Mating (DIAMONDS) and Mate Seeking (SAAP), both referring to the same situational characteristic. In the current state of research progress, it is an empirical question as to which dimensions best describe the underlying dimensions of situation perception characteristics, which dimensions describe actually the same characteristic, and which dimensions are labeled more or less the same but describe different characteristics of situational perception.

Further, the level of abstraction may be completely different for each of the different situational taxonomies, as it is also the case with personality dimensions and facets of personality (e.g., Revelle & Condon, 2015). Rauthmann and colleagues (2014) have already presented evidence that there are different levels of situational perception dimensions and that diverging numbers of dimensions across situational taxonomies are likely to occur. Facetted structures of situational perception taxonomies are therefore likely. The recent development of situational taxonomies and their broadened empirical application however make integration more feasible and likely.

**Possible Integration of Situational Taxonomies and Measures**

Each of the presented taxonomies has advantages and disadvantages. For example, the taxonomies have different contexts they seek to address. The DIAMONDS and CAPTION models focus on broad, everyday situations, the SAAP and SIS on situations that are relevant for achieving evolutionary goals, the SIS is particularly focused on social situations, and the Situation 5 is focused on occupational settings. Yet, there is large content overlap between the taxonomized dimensions which may allow for an integration into one joint situational taxonomy despite the differences in the construction process and aim for each taxonomy (Ziegler, Booth, & Bensch, 2013).

Figure 2.4.1 gives an overview over all five situational dimensions published since 2014 (i.e., CAPTION, DIAMONDS, SAAP, Situation 5, SIS) and their empirical or conceptual relation to each other. Particularly other measures' relations to the DIAMONDS seem to be well-examined so far with CAPTION (Parrigon et al., (2017), SAAP (Brown et al., (2015), SIS (Gerpott et al., 2017), and Situation 5 (Horstmann & Ziegler, 2018). All other postulated relations are, however, primarily based on item content and the construct definition of the underlying dimensions. Based on the available information, we identified at least six factors that are replicably found across *independent* research efforts: Threat, Stress, Tasks, Processing, Social Positive, and Mundane (see also Rauthmann & Sherman, 2017).[15]

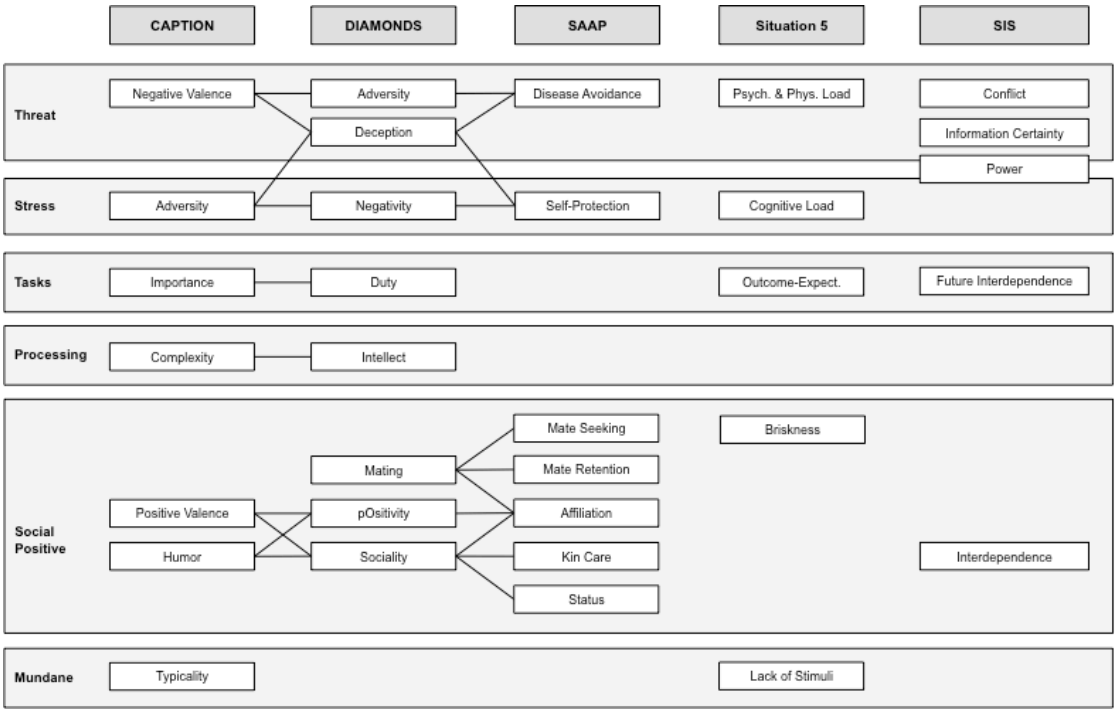



*Figure 2.4.1.* Situation Characteristics Taxonomies Overview, by Rauthmann & Horstmann (2017, licensed under CC-BY 4.0). For details, see text.

*Threat* describes situations that are perceived as threatening. The perceived threat primarily results from external sources, for example, other people or any form of danger. The perceived threat need not be physical in nature and may often be more psychological (e.g., criticism as a threat to self-esteem). Depending on the ability to cope with these external forces, the situation may or may not be perceived as negative or burdensome. Even though it is likely that most situations that are threatening are also perceived by most people as negative or stressful, this does

[15] Note that the first 5 of these are the same dimensions identified by Rauthmann and Sherman (2018a). However, here we provide psychologically meaningful labels as opposed to Domains I-V used by Rauthmann and Sherman.

not have to be the case: Some people may like or actively seek threatening situations to have a positive experience (e.g., sensation seeking), or they might see them as "challenging" in the sense of an opportunity to overcome something and grow by doing so (e.g., growth-mindset). Threat consists of the dimensions Negative Valence (CAPTION), Adversity, Deception (both DIAMONDS), Disease Avoidance (SAAP), Psychological and Physical Load (Situation 5), and Conflict and Information Certainty (SIS).

*Stress* describes stressful and tense situations. A person that characterizes situations in such a way already experiences stress and negative feelings. Resources may then be required to further cope with the stress itself. Whereas Threat refers to situations that are threatening and therefore *potentially* stressful, Stress refers to situations that are already interpreted as stressful. Unlike Threat, situations characterized by high Stress tend to be more internally focused (e.g., situation is potentially anxiety-inducing) than those characterized by Threat, which typically stem from something external. Stress consists of the dimensions Adversity (CAPTION), Negativity (DIAMONDS), Self-Protection (SAAP), and Cognitive Load (Situation 5).

*Tasks* refers to situations that have important outcomes, in which work needs to be done, and on which the future depends. Minor details are often important in such situations and the cost of making mistakes is often quite high. These situations just require the completion of a task, job, or duty. Tasks consists of the dimensions Importance (CAPTION), Duty (DIAMONDS), Outcome-Expectancy (Situation 5), Future Interdependence (SIS).

*Processing* refers to situations that require thinking, analysis, and intellectual engagement. These situations may range from complex problem solving to simple recollection of knowledge. When social in nature, situations characterized by processing may involve philosophical or political discussions. Processing consists of the two dimensions Complexity (CAPTION) and Intellect (DIAMONDS).

*Social Positive* is the by far broadest dimension, and primarily refers to social situations. With few exemptions (pOsitivity, positive Valence, and Humor, Briskness), all dimensions refer to the presence of other humans. As Brown and colleagues (N. A. Brown et al., 2015) note, most daily activities are social, and it is therefore reasonable to differentiate between numerous aspects of social situations. Some of the Social Positive dimensions can therefore be considered specific facets. For example, there are sexually relevant dimensions including Mate Seeking, Mate Retention, and Mating. All dimensions subsumed under Social Positive are furthermore positive in nature – other dimensions that are not subsumed under Social Positive (e.g., Deception) may clearly have a social component, but do not have to be positive.

One noteworthy dimension in Social Positive is Status (e.g., SAAP: "It is highly important to gain respect from others"). This dimension correlates highly with all of the DIAMONDS dimension (N. A. Brown et al., 2015, Table 4), and is therefore difficult to assign to any of the broader six dimensions presented in Figure 2.4.1. It has to be determined in the future if Status is a separate dimension or a higher order factor of many or even all situational perception dimensions. Social Positive consists of the dimensions Positive Valence and Humor (CAPTION), Mating, pOsitivity, and Sociality (DIAMONDS), Mate Seeking, Mate Retention, Affiliation, Kin Care, Status (all SAAP), Briskness (Situation 5), and Interdependence (SIS).

*Mundane* describes situations that have few stimuli and are characterized as boring, normal, known, or lacking the presence of input. It does not correlate highly with any DIAMONDS dimension (the highest correlation is with Duty, r = .12, see Parrigon et al., 2017, Table 7). Only two dimensions constitute Mundane so far: Typicality from CAPTION and Lack of Stimuli from Situation 5. It is noteworthy that these two situation taxonomies are grounded in the lexical approach and may thus be more open-ended in their construction process. Even though Typicality and Lack of Stimuli may be facets of the same dimension, they are certainly not the same construct: Lack of Stimuli describes situations that are dull and boring, yet none of such situations need to be likely, known, or typical. Typicality on the other hand may be not so much be a content dimension – such as Deception, Adversity, Outcome-Expectancy – but rather a valence dimension, such as positive or negative.

At this stage, no study has examined more than two taxonomies of situation characteristics at the same time, and the findings of the studies presented here have not been replicated. Nevertheless, this integration based on item content, theoretical reasoning, and those scarce correlations that are available is useful and crucial (e.g., for pre-registered studies). It is also notable how the dimensions found mirror the Big Five or HEXACO taxonomy of personality traits: Threats ≈ Dis(agreeableness) or Dis(honesty)/In(humility); Stress ≈ Neuroticism; Tasks ≈ Conscientiousness; Processing ≈ Intellect, Openness to new experiences; and Social Positive ≈ Extraversion. Only Mundane does not seem to fit in here.

**Future Recommendations for the Development of Situation Taxonomies**

The integration of existing taxonomies of situation perception is the next important step of situation research. Integration does not necessarily mean that a major situational taxonomy unifies all extant ones; it rather means that the structure of the constructs and the relationships of the constructs to each other is established. Cronbach and Meehl (Cronbach & Meehl, 1955) referred to this process as the establishment of the nomological net. It is likely that further situation taxonomies – which may not only focus on situational characteristics – will be developed and

thereby fill existing (and possibly unknown) gaps in the current nomological net of situation research.

Future developers of situation taxonomies should consider pre-registration of their expectations and beliefs (e.g., Asendorpf et al., 2013). With a glance at all the possible connections between dimensions of situational assessment (Figure 2.4.1), it is likely that correlations occur that have previously not been thought of or were not considered. Finding explanations for these associations can lead to post-hoc explanations and meaning-making where none is (Gelman & Loken, 2014; Kerr, 1998; Miguel et al., 2014; Simmons, Nelson, & Simonsohn, 2011; Ware & Munafò, 2015) and thus unreliable findings. Only a sound integration of situational measures will lead to acceptance and usage outside of the current field.

**Further Issues of Situational Measurement**

As noted above, a measurement tool is crucial to the widespread use of situational taxonomies and even more so for an integrated version of a situational taxonomy. Designing a measurement tool for any taxonomy is difficult and requires multiple studies to refine the tool and make it as reliable and valid as possible. The result of the construction process will not only depend on the theoretical underpinnings, but largely on the intended use of the measure (Ziegler, 2014c). Situational assessment tools can, for example, be used to assess a "fixed" situation where several participants are exposed to and rate the same situation (Serfass & Sherman, 2013). Another option is measuring (e.g., in ambulatory assessment) a "random" situation that will vary between participants and within each participant between measurement occasions (because different situations are sampled per person). This again requires the development of situational measures that are short but nevertheless reliable and valid (e.g., Rauthmann & Sherman, 2016c).

At the current stage, we recommend using the extant situational taxonomies based on the requirements of the study at hand. For any study that seeks to compare situations across cultures, the DIAMONDS as well as the SAAP are a good choice. The DIAMONDS taxonomy was developed using multiple samples from different countries, and the corresponding measurement tool (i.e., the RSQ) has been used in a cross-cultural study and comparisons (as well as translations) are thus available (Guillaume et al., 2016). The SAAP scales are based on evolutionary theory and therefore claim to be also applicable outside of the US population. However, translations of the items are still required. Note also that the SAAP dimensions are age-graded and will in some cases (e.g., Kin Care) only be relevant for certain populations.

For social and interdependent situations (especially within the context of social exchange games), the SIS will be the best current measure available. Conditional on the existence of

translations, the SIS should also be applicable in different cultures, since interdependence theory is also applicable across cultures.

For any occupational settings, the Situation 5 may be the best option to use. Especially its assessment tool and the resulting trait scores of situational perceptions can be helpful for the purpose of personnel selection. The Situation 5 are also well-equipped for the characterization of occupational settings and may thus be used to support analyses of job-requirements (Ziegler et al., submitted).

For brief and repeated assessment, for example in an experience sampling design or daily diary study, the DIAMONDS as well as the CAPTIONs can be used, since shorter versions of both taxonomies exist (Parrigon et al., 2017; Rauthmann & Sherman, 2016c). If, on the other hand, more time is available and a rich psychological assessment of the situation is desired, the RSQ is perhaps the best tool available. One major advantage of the RSQ is that it has already been translated into many languages.

**The Principles of Situation Research**

Situation research does not only require tools and taxonomies to assess and describe situations, but also a solid methodological approach. Rauthmann and colleagues (2015) have presented principles of situation research that give guidance to the assessment of situations: The Processing Principle, the Reality Principle, and the Circularity Principle. The Processing Principle claims that only the psychologically important aspects of a situation are relevant. The Reality Principle states that situations have multiple realities, which are their physical one (how the situation actually is), their consensual one (how people perceive a situation in the same way), and their idiosyncratic one (how people perceive a situation in a unique way). Finally, the Circularity Principle states that persons – defined as perceivers of a situation – and the situation itself are "conflated once a situation variable is defined by a person variable" (Rauthmann, Sherman, & Funder, 2015b, p. 367, Table 3).

Rauthmann and colleagues also present logical conclusions (corollaries) from these principles. From the Circularity Principle, it follows that situations cannot be assessed by the current mental states or activities of a person (State Corollary, e.g., I feed sad, therefore, the situation is a sad situation) or consequences (Consequences Corollary, e.g., Now I feel sad, therefore I must have experienced a sad situation previously). To obtain the best estimate of a psychological meaningful situation, the Approximation Corollary states that any situation is best assessed from three different angles or raters: Raters *in situ*, raters *juxta situm*, and raters *ex situ*. Each describes the position of a perceiver relative to the situation. *In situ* refers to a person that has firsthand experience of the situation. *Juxta situm* refers to a perceiver that is present in the

situation, but has no direct role in it. An example for an observer *juxta situm* could be a research assistant in a behavioral experiment. A rater is positioned *ex situ*, or outside of the situation, when he/she has not had direct contact with the situation but can only assess information about the situation. Raters *ex situ* can be raters of video recordings or rates of transcribed reports of the situation. The shared view of all perceivers is therefore the best and most accurate description of a situation.

For the most accurate assessment of any situation, the Circularity Principle demands multiple informants with different perspectives and ratings. Depending on the task at hand, this may be more or less easy to achieve. For a rather stable and "fixed" environment, such as a workplace or an experimental situation, numerous ratings from all three sources may be acquired. For others, such as daily experiences, even ex situ ratings are hard to come by, and juxta situm ratings may be practically impossible to obtain. Yet, newer technological developments, such as virtual reality, wearable cameras, and smart-phones with geo-position may help overcome these challenges in the near future. In turn, such tools may help provide better explanations and ultimately predictions of human behavior, feelings, or trait expressions.

## Conclusion

The assessment of situations has come a long way. After two waves of situational research in the last 50 years, consensus about the assessment and description of situations is starting to emerge. The next steps should involve the replication of existing findings, the validation of situational measures, and the application of situation measures in other fields of psychology. These advancements will allow other researchers to tackle questions that could previously not be addressed and extend knowledge about the person in his or her environment. Finally, the initial question that sparked all situation research – why personality does not perfectly predict behavior, and how the prediction of behavior can be improved – is ready to be faced with new insights from situational assessment.

## 2.5 Conclusion Part 2

Situation research has recently seen a surge in the number of published situational taxonomies (Rauthmann et al., 2014; Yang et al., 2006), and, notably, means of assessing situation characteristics in psychological research. This heightened interest in psychological situations comes at a time where the person-situation debate has ended and personality theories call explicitly for an integration of situations into personality research (see Part 1, section 1.3). Furthermore, new (or easier to use) psychological research methods allow sampling participants during their daily life and in natural situations that are not under the control of the scientist (Harari et al., 2016; Wrzus & Mehl, 2015). The ability to describe any unknown situation and assess its effects on behavior will thus benefit future research.

However, many researchers are currently using different measures of situation perception dimensions in their respective areas of research. The development of the nomological net of situation characteristics (and future tests thereof), will allow comparing results from different studies as well as identifying gaps in existing taxonomies and thus subsequent research. For example, researchers who are currently using the DIAMONDS taxonomy (Rauthmann et al., 2014) and its related measures (Rauthmann & Sherman, 2016a, 2016c) may consider including additional items from the CAPTION or the Situation Five framework to assess the dimensions Typicality or Lack of Stimuli, respectively.

If situations are measured repeatedly over time (e.g., in experience sampling or ambulatory assessment), one way to analyze this data is by using multi-level models. These models frequently require that the mean of all state assessments per person (i.e., measures at the situation level) is computed and included as a person variable (i.e., at the person level) to control for inter-individual differences on the respective variable (Enders & Tofighi, 2007). One could argue that the average of all situation perception assessments resembles that of a personality trait of situation perception (Augustine & Larsen, 2012; Fleeson & Gallagher, 2009; Rauthmann et al., 2018). However, previous research has not interpreted these variables in that manner (e.g., Sherman et al., 2015). Given the promising effects of trait situation perception (Horstmann & Ziegler, 2016; Ziegler et al., submitted), I would argue that the trait-component of situation perception should be acknowledged. At the same time, defining situations via the perception of persons confounds person and situation variance (circularity principle, Rauthmann, Sherman, & Funder, 2015b). The topic of Part 3 is precisely this confound, and more specifically, the overlap of situation perception with affect.

# 3 Part 3: Affect and Situation Perception

Operationalizing situations by how they are perceived is, on the one hand, an economical way of examining situational influences. Not only does it save a researcher from collecting all possibly relevant situation cues, it also eliminates the necessary burden of assigning psychological meanings to these cues. Although this may be achieved using modern algorithms (N. A. Brown et al., 2017), it is not clear if each situation cue has the same meaning for each perceiver. As an example, consider a party: Whereas one person may perceive the party as stimulating, another person may perceive the same party as boring (Rauthmann, 2012). On the other hand, this may easily confound person and situation effects (Rauthmann, Sherman, & Funder, 2015b). Selecting, perceiving, and interpreting situation cues may not only be correlated with personality traits (Funder, 2016; Ickes et al., 1997; Rauthmann, Sherman, Nave, et al., 2015), but also with momentary states of the person. In this part, I will examine one particularly important class of person states: affect and emotion.

**Distinction: Affect vs. Emotion.** Throughout this work, the term affect and emotion are used more or less interchangeably. Although correlated, affect and emotions are not entirely exchangeable. Most notably, emotions are assumed to be tied to a certain object, whereas this may not always be the case for affect (Gendolla, 2000). Further, affect is mostly experienced without awareness of its origin, whereas this usually is the case for emotions. Emotions are therefore object-related. However, affect is seen as "residuals of emotions" (Gendolla, 2000, p. 379), and can thus be understood as a consequence of emotions, albeit an indirect one. For the most part of the current work, I will therefore not distinguish between affect and emotions. This has to be kept in mind for the current section, and I will discuss this limitation of the current work in the next part.

## 3.1 Situation Perception and Appraisal Theories of Emotion

The CAPS model (Mischel & Shoda, 1995), or the explanatory part of traits in the WTT (Fleeson & Jayawickreme, 2015), but also the principles of situation research (Rauthmann, Sherman, & Funder, 2015b) all sought to bridge the gap between situation cues and their interpretation, which are situation characteristics. This interplay of events and their interpretation is at the center of appraisal theories of emotion (Kuppens, 2009; Kuppens & Van Mechelen, 2007; Sander et al., 2005). Kuppens (2009) therefore suggests that such theories may be of high value to the resolution, or "fundamental convergence" (Kuppens, 2009, p. 255), of the person-situation debate. One relevant model is the component process model (Scherer, 1999, 2001), which depicts

meticulously how the perception of a situation cue is evaluated and thus leads to emotions. Precisely, four steps are required such that an individual perceives an object *and* experiences an emotion (Scherer, 2001, p. 94). These steps are: (1) The relevance of the event (e.g., whether it affects the perceiver or other persons relevant to the perceiver), (2) the implications or consequences of the event, (3) the ability to cope with or adjust to the event, and (4) the normative significance of the event regarding the persons' self-concept and social norms and values. Note that, given the relation of emotion and affect, such an evaluation will not only lead to emotions, but also affect (Gendolla, 2000).

The four steps described by Scherer (2001) may lead to a certain emotion in reaction to an object or situation cue. However, it could also be argued that the exact same steps take place during the formation of situation perceptions. As an example, let us consider the situation of taking an exam at a university, at which students and supervisors are present. Depending on the task at hand (e.g., supervising the exam or taking it), these four steps may result not only in different emotions or affect (e.g., fear vs. contempt, high negative affect vs. medium negative affect), but also in different levels of situation characteristics. Whereas the situation may be highly relevant (1) for the student, this may not be the case for the supervisor. The exam may have different consequences (2) for the supervisor (he or she may have to sit there for 90 minutes and grade the exam afterwards), but repeatedly failing the exam may lead to expulsion from the university for the student. The supervisor surely has all resources (3) to cope with the situation, whereas the student may not be so sure about his or her abilities to correctly answer the questions. Finally, the supervisor is just doing his or her job, and it may not involve any particular expectations (4) about the future, but, depending on the result of the exam, the self-concept of the student may be severely changed.

In their perceptions, students might rate the situation on the DIAMONDS (Rauthmann et al., 2014) as high on Duty and high on Adversity, and the supervisor may rate the situation high on Duty but very low on Adversity. Given the possible similarity in the processes that form affect as well as situation perceptions, it is likely that both correlate. The examinations of this potential overlap was part of the study *Situational perception and affect: Barking up the wrong tree?* (Horstmann & Ziegler, 2018). Previous studies that had examined the connection of affect and situation perception (Parrigon et al., 2017; Sherman et al., 2015) interpreted this overlap as being indicative of predictive validity of the measures of dimensions of situation perception. In our study, we investigate if this overlap may not, in fact, be too high, thus threatening discriminant validity of situation perception dimensions. If this were the case, an extreme case of a jingle-jangle-fallacy had occurred.

### *3.1.1 Article Summary: Situational Perception and Affect*

**Background.** The simple process model of situation perception (Rauthmann, Sherman, & Funder, 2015b) describes how situation perceptions are formed. A person perceives a situation cue, which is a physical object, such as a chair or table, other persons, but also the current time and location. These physical cues are then interpreted. This interpretation depends on stable personality characteristics or traits (Funder, 2016; Rauthmann, Sherman, Nave, et al., 2015), but also on the social role of the perceiver, and last but not least, on the affect or mood in the situation. Thus, the interpretation of cues and current affect should be related. On the other hand, appraisal theories of emotion posit that emotional processes are triggered by situation cues: For example, if a person sees a dangerous animal, they might experience fear (Sander et al., 2005; Scherer, 2001). It is therefore plausible that situation perception and affect are related. Previous studies have reported a correlation of situation perception and affect (Edwards & Templeton, 2005; Parrigon et al., 2017; Sherman et al., 2015). However, this overlap has usually been interpreted as evidence for predictive validity of situation measures, in the sense that measures of situation perception dimensions allow explaining variance in affect or happiness. However, given that this overlap is sometimes quite substantial (in the range of $r$ = .60 to .70), we sought to examine this overlap in more detail.

**Study Design.** $N$ = 157 participants completed an online questionnaire. Participants were required to rate three situations from their previous day on measures of the Situational Eight DIAMONDS (Rauthmann et al., 2014) and the Situation Five (Ziegler et al., submitted). Additionally, participants rated their own state affect during the situation on the Positive and Negative Affect Schedule (Krohne, Egloff, Kohlmann, & Tausch, 1996; Watson, Clark, & Tellegen, 1988).

**Analyses.** We examined bivariate correlations between the Situational Eight DIAMONDS dimensions, the Situation Five, and Positive (PA) and Negative Affect (NA). Additionally, we used multiple regression to predict PA and NA with the Situation Five, the DIAMONDS, and finally both the Situation Five and the DIAMONDS combined. Due to the nested data structure (situations in persons), we computed robust standard errors to take the nested structure into account (McNeish, Stapleton, & Silverman, 2017).

**Results.** The bivariate correlations revealed that all situation perception dimensions were substantially correlated with either PA, NA, or both. The multiple regression analyses revealed that substantial proportions of variance in PA and NA could be explained with situation perception measures ($R^2$ between .35 and .62). This indicated a strong overlap of situation perception and PA and NA.

**Conclusion.** We were able to replicate the results presented by Parrigon and colleagues (2017) and show that affect and situation perception are related. This has several implications for an understanding of situation perception. First, the nomological network of situation perception dimensions should consider positive and negative affect as relevant constructs. Second, future developments but also applications of situation perception taxonomies must establish discriminant validity for measures of situation perception with regard to measures of affect. It could otherwise be possible that effects that are attributed to situation perception (for example effects on behavior) should have been attributed to affect. Thus, discriminant predictive validity needed to be examined in further studies. Finally, this research raised concerns that other person states, such as physical arousal, emotions, or even satiation may play a considerable role in the perception and interpretation of situations.

### *3.1.2 Article: Situational Perception and Affect*

**Situational Perception and Affect:**
**Barking up the wrong tree?**

Kai T. Horstmann & Matthias Ziegler
Humboldt-Universität zu Berlin, Germany

This article was published as:

Horstmann, K. T., & Ziegler, M. (2018). Situational perception and affect: Barking up the wrong tree? *Personality and Individual Differences*. doi.org/10.1016/j.paid.2018.01.020.

This is the final author copy before copy editing.

**Funding:**
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Conflict of Interest:**
There are no interests that might be interpreted as influencing the research.

**Abstract**

The current study was conducted to raise awareness to the possibility that perceptions of situations and *in situ* affective states might be highly correlated. To investigate this potential overlap, two recent taxonomies of situational perception, the Situational Eight DIAMONDS and the Situation 5, were assessed in a sample of $n = 157$ along with a measure of positive and negative affect. Participants provided accounts and ratings for all constructs for three self-selected situations. Overall, 383 situations could be analyzed using multiple regression while considering the nested data structure. Both the DIAMONDS and the Situation 5 scores showed considerable overlap with positive and negative affect scores. The study further advances the growing nomological net of situational perception dimensions and other constructs. Limitations such as the selective-reporting-bias and implications for future situation research are discussed.

## Introduction

The person-situation-debate, that is the question how internal person factors as opposed to situational forces influence human behavior in any given situation, has, more or less, come to an end (Fleeson & Noftle, 2009). A consensus has emerged that both internal, stable person factors (personality) as well as external situations can influence behavior (Mischel & Shoda, 1995; Rauthmann, 2012; Sherman et al., 2010, 2015), and one subsequent consequence was that the situation has received a lot of research interest (Horstmann, Rauthmann, et al., 2018). One of the main points of criticism that had to be initially addressed was that it was considered nearly impossible to measure situations (Hogan, 2009). Without a descriptive system and an accompanying measure for situations, the situation could not be considered as a predictor for human behavior. Recent research endeavors have fortunately produced such situation taxonomies and assessment tools that allow to measure interindividual differences in how situations are perceived (N. A. Brown et al., 2015; Gerpott et al., 2017; Horstmann, Ziegler, et al., 2018; Parrigon et al., 2017; Rauthmann et al., 2014).

Situation perception as a construct has since been successfully used to predict various outcomes, such as behavior, affect, well-being, or happiness (Parrigon et al., 2017; Sherman et al., 2015) cooperation (Gerpott et al., 2017) or goal related behaviors (N. A. Brown et al., 2015). Even though this looks very promising, there is cause for concern: Measures of situational perception, especially when used as *in situ* ratings of situations, have been shown to be related to affect (Parrigon et al., 2017; Sherman et al., 2015). The questions now are how strong this relation between affect and situation perception is and whether it occurs for different conceptualizations of situational perceptions. If affect was, in fact, strongly related to different situational perception measures, operationalizing different situational perception taxonomies, these two constructs might be indistinguishable in the worst case. Many of the effects attributed to situational perception reported so far could then actually be due to affect, and situation research would be barking up the wrong tree.

This paper therefore serves two purposes: First, we examine how situational perception and affect are conceptually related, and second, we show empirically the overlap between affect and situational perception for two different measures and taxonomies of situational perception.

**Situational perception**

According to Rauthmann and colleagues (Rauthmann, Sherman, & Funder, 2015b), psychological situations can be organized on three different levels: situational Cues, situational Characteristics and situational Classes. Cues refer to the physically present elements in a situation; for example, a table or a computer, but also time and space. Situational characteristics describe the psychological meaningful aspects of a situation on a broader level (Horstmann, Ziegler, et al., 2018; Horstmann & Ziegler, 2016; Ziegler et al., submitted). On the broadest level, situations can be clustered into classes or contexts, such as "work" or "home" (e.g., Geukes et al., 2016).

If a person perceives a situational cue, this cue has to be interpreted, and the result of this interpretation is the psychologically meaningful characteristic of a situation (the "Processing Principle", Rauthmann, Sherman, & Funder, 2015b; see also Ziegler & Horstmann, 2015). For example, based on knowledge about animals, a large snake may be interpreted as dangerous or beautiful. Different persons will thus have different perceptions of such a situation: Generally, most people would agree that this could be a dangerous situation (shared variance), but each person will also have a unique perception of this situation (idiosyncratic variance). This idiosyncratic variance depends on the unique perspective of the person, and therefore could be correlated with affect.

**Situational Perception and Affect**

How are situational perception and affect related? If we imagine a perceiver entering two identical situations (with respect to their situational cues, with the exception of time), it is possible that this perceiver will interpret the situation differently. If this were to happen, it could only be due to changes within the person since the cues within both situations are identical. One construct very likely to change within a person over time is affect. This would mean that affect influences the perception of a situation, and, following the process model of situational perception, consequently behavior (Rauthmann, Sherman, & Funder, 2015b). On the other hand, the perception of a situation could also influence affect. As described by Sander and colleagues, "emotional processes are elicited and dynamically patterned as the individual continuously and recursively appraises objects, behaviors, events, and situations"[16](Sander et al., 2005, p. 318). Thus, emotions or affect develop along with the person's perception of the situation as being relevant to well-being, behavior, goals, values, etc. For example, if a person perceives that a situation calls for a task to be done (e.g., learning for an exam), but simply is too tired, negative affect might result. Thus, not the fact that work needs to be done per se, but the interaction of this perception and the current general state of the person results in affect.

[16] We are very thankful to an anonymous reviewer for pointing this out.

If situation perception and affect were related, it would have to be examined how large this overlap actually is. There are many studies reporting results that might be indicative of the relation between situational perceptions and affect (Edwards & Templeton, 2005). However, only a few studies specifically used measures of situational perception and affect. For example, Sherman and colleagues (2015) first examined the relation of situational perception and happiness and found that some dimensions of situational perception were substantially related to happiness. However, they did not explicitly investigate the relation of *all* situation measures with happiness. Yet, Parrigon and colleagues examined the overlap of situational perception and affect in a cross-sectional design (Parrigon et al., 2017). Using positive (PA) and negative (NA) affect as dependent variables and measures of three different situational taxonomies as predictors, those authors showed that measures of situational perception explained between 51 – 69% of variance in PA, and between 67 – 83% of variance in NA. Furthermore, bivariate correlations between their measure of situational perception (CAPTION), ranged between -.03 and .53 (median $r = .27$) for PA, and between .02 and .77 (median $r = .47$) for NA. In other words, with the exception of the dimension Typicality, all other dimensions of situational perception correlated with either PA or NA above .46 (significant at $p < .05$, with $N = 522$). Parrigon and colleagues interpreted these findings in favor of their measure, demonstrating its ability to predict affect. However, if these results were robust and replicable across other measures and taxonomies of situational perception, we would like to offer an alternative interpretation based on the theoretical reasoning stated above, namely a lack of discriminant validity of these measures.

**Research Question**

Based on the previous findings and theoretical assumptions, positive and negative affect should be substantially correlated with measures of situational perception. In the present study, we will examine if these findings can be replicated using two different measures of situational perception representing two different taxonomies. This approach will also allow us to examine the convergent validity among these measures of situational perception and help to expand the nomological net of taxonomies of situational perception.

## Methods

### Sample

From initially 190 participants who started the study, 157 participants completed all relevant questionnaires. Of these, 84 % (=132) were female. The mean age was $M_{age}$ = 23.96 ($SD_{age}$ = 6.53, $median_{age}$ = 22). Most of the participants were undergraduate students in psychology, participating for course credit. The sample size was not determined a-priori, since the data reported here constitute the first part of a multi-wave assessment in preparation for a behavioral experiment. Data collection was stopped when the research program for the behavioral experiment was stopped due to resource-constraints.

### Measures and Procedure

Participants first had to answer a standard set of questions regarding their age, gender, educational status, job success and job satisfaction. They subsequently had to take two standardized trait measures, the German version of the Big Five Inventory (Lang, Lüdtke, & Asendorpf, 2001) as well as the German version of the Positive and Negative Affect Schedule (PANAS; Krohne et al., 1996; Watson et al., 1988). From these measures, only age, gender, and PANAS trait are reported in this study.

Subsequently, participants were required to report three situations from the previous day which were chosen based on time of day: 9 a.m., 2 p.m., and 7 p.m. Participants then had to briefly describe the situation (verbally), list present persons, describe their own behavior, what they were doing, what time it actually was (in case participants were sleeping, they could then select another situation), and how long the situation lasted. This procedure is commonplace in the construction and development of situation taxonomies (N. A. Brown et al., 2015; Gerpott et al., 2017; Parrigon et al., 2017; Rauthmann et al., 2014; Ziegler et al., submitted). Afterwards, they completed the following measures for each of the situations they reported.

**Situational Eight DIAMONDS.** The Situational Eight DIAMONDS were developed by Rauthmann and colleagues (Rauthmann et al., 2014) to describe situational characteristics. They consist of a set of statements that can describe a situation, such as "The situation is playful" (see Table 3.1.1 for sample items, which are similar to those provided by Rauthmann & Sherman, 2016). The participant then has to rate how much s/he agrees to this statement on a 6-point Likert-type scale. The dimensions assessed are Duty (the extent to which work has to be done), Intellect (the extent to which the situation requires deep thinking, or analysis), Adversity (to which extent the situation is eliciting stress or is dangerous), Mating (to which extent potential sexual partners are present), pOsitivity (to which extent the situation is positive), Negativity (to which extent the situation is negative), Deception (to which extent someone can be deceived), and Sociality (to

extent the situation is social, allows interaction and communication). Each dimension was measured with four items each.

**Situation 5.** The Situation 5 are a taxonomy that captures five dimensions of situational perception (Ziegler, 2014a; Ziegler et al., submitted). The Situation 5 were developed in the tradition of the lexical approach (Horstmann, Rauthmann, et al., 2018; see Parrigon, 2017 for an overview of the lexical approach for situations). Initially, roughly 15,000 adjectives were extracted from a German spelling dictionary, reduced to about 300, suitable to describe situations, by three independent raters. Participants were then asked to rate situations they encountered during the previous day (Horstmann, Ziegler, et al., 2018; Ziegler et al., submitted). The final five dimensions found and confirmed using factor-analytical techniques are Outcome-Expectancy (e.g., "full of potential"), Briskness (e.g., "lively"), Psychological and Physical Load (e.g., "burdensome"), Lack of Stimuli (e.g., "boring", "dull"), Cognitive Load (e.g., "challenging").

For the current study, participants had to rate the situations on three adjectives (for sample adjectives, see Table 3.1.1) for each of the Situation 5. A 6-point Likert-type scale was used.

**PANAS State.** The Positive and Negative Affect Schedule assesses positive (PA) and negative (NA) affect of a person. Depending on the instruction (e.g., "over the last 12 months" vs. "in the current situation"), the PANAS can be used to describe either trait affect or state affect. Positive Affect describes feelings such as positivity, activity, or enthusiasm. Negative Affect describes feelings such as negativity, sadness, or guilt. Participants were required to assess their feelings during their self-reported situation on the PANAS on 20 items, 10 for positive affect and 10 for negative affect, using a Likert-type scale ranging from 1-6. Descriptive statistics, reliabilities, and sample items of all scales are presented in Table 3.1.1. Item statistics for the trait-PANAS can be found in the OSM.

**Further items not used in the current study.** Participants also had to rate their physical arousal and feeling. The results of these ratings are not part of this study and have not been analyzed with respect to this research question. All items can be found in the codebook for this study provided in the OSM.

**Data analysis**

The data analyses took part in several steps. All materials to reproduce the analyses can be found in the OSM.

First, data from participants who had only clicked on the survey but not completed it were deleted ($n$ = 33). Due to an error in the study set-up, some responses were registered on a 7-12 format instead of a 1-6 format, which were therefore recoded. Second, all items that belonged to one scale were averaged to create the scale score (e. g. the score on Positive Affect for the situation

reported for 9 a.m.). Third, all situational reports of individual participants were collapsed, as if independent participants generated the reports. After deleting all reports that had more than one missing data point, 383 individual reports remained. The nested data structure will be addressed below. Fourth, bivariate correlations between the scales of *in situ*-ratings were computed and reported. Fifth, linear regressions were performed to predict either positive or negative affect with the Situation 5, the DIAMONDS, and the Situation 5 and the DIAMONDS combined to show the overall overlap in variance between the measures of situational perception and affect (this is similar to the procedure described by Parrigon et al., 2017).

Sixth, we took the nested structure of the data (i.e., situations nested in participants) into account by following recommendations by McNeish, Stapleton, and Silverman (2017) and computing cluster robust standard errors for the regression estimates. This changes the standard errors (and subsequently the confidence intervals and the *p*-values), but all other estimates (e. g. explained variance $R^2$ or beta-weights) remain unchanged. We report the results of the OLS regression with and without robust standard errors, as well as 95% confidence intervals for the estimates.

Additionally, further analyses were conducted:[17] We first computed Intra-Class Correlations (ICC) based on an unconditional random effects model for each of the situation variables and state-affect (Snijders & Bosker, 1999). We also used multi-level modelling to predict affect in situations by situational perception and trait affect. Independent state variables were centered within-cluster, and independent trait variables were grand-mean centered, as is recommended in multi-level analyses (Enders & Tofighi, 2007). The results of the multi-level analyses confirm the findings from the multiple-regression analyses and are therefore only displayed in the Online Supplemental Materials (OSM, section *Multi-Level Models*).

**Software.** We used R (R Core Team, 2016), version 3.3.0, as well as RStudio (RStudio Team, 2015), version 0.99.902. Furthermore, we used the following packages: apaTables (Stanley, 2015), clubSandwich (Pustejovsky, 2016), dplyr (Wickham & Francois, 2016), knitr (Xie, 2016), lm.beta (Behrendt, 2014), psych (Revelle, 2016), purrr (Wickham, 2016a), stringr (Wickham, 2016b), xlsx (Dragulescu, 2014), and their respective dependencies. For the estimation of robust standard errors, we adapted a solution provided in the package-vignette of the package clubSandwich (Pustejovsky, 2016). Further packages for additional analyses were used and are listed in the OSM.

[17] These are based on reviewer feedback and were added after the first version of this manuscript. We are thankful to the reviewers for suggesting these analyses.

Table 3.1.1
*Descriptive Statistics of the DIAMONDS, the Situation 5, and Positive and Negative Affect*

| DIAMONDS | Sample item/adjective[1] | *M* | *SD* | *med.* | *min* | *max* | *a* | *ICC* |
|---|---|---|---|---|---|---|---|---|
| Duty | Work needs to be done | 3.65 | 1.23 | 3.50 | 1 | 6 | .66 | .09 |
| Intellect | Deep thinking is required | 3.15 | 1.35 | 3.25 | 1 | 6 | .73 | .09 |
| Adversity | You are being criticized | 1.5 | 0.70 | 1.25 | 1 | 4.75 | .62 | .19 |
| Mating | Potential romantic partners are present | 2.45 | 1.33 | 2.25 | 1 | 5.75 | .60 | .11 |
| pOsitivity | Situation is playful | 3.78 | 1.04 | 3.75 | 1 | 6 | .37 | .12 |
| Negativity | Situation is frustrating | 2.47 | 1.15 | 2.25 | 1 | 6 | .76 | .08 |
| Deception | It is possible to deceive someone | 1.95 | 0.93 | 1.75 | 1 | 4.75 | .57 | .29 |
| Sociality | Social interactions are possible | 3.44 | 1.57 | 3.75 | 1 | 6 | .82 | .03 |
| Situation 5 | The situation is… | | | | | | | |
| Briskness | …lively | 3.76 | 1.36 | 4.00 | 1 | 6 | .89 | .08 |
| PP Load | …burdening | 2.79 | 1.33 | 2.67 | 1 | 6 | .86 | .03 |
| Out-Exp. | …potentially rewarding | 3.46 | 1.15 | 3.33 | 1 | 6 | .43 | .32 |
| Cognitive Load | …cognitively demanding | 3.07 | 1.49 | 3.00 | 1 | 6 | .90 | .09 |
| Lack of Stimuli | …boring | 2.14 | 1.19 | 1.67 | 1 | 6 | .81 | .03 |
| State-Affect | | | | | | | | |
| positive | I feel active | 3.54 | 1.00 | 3.6 | 1 | 6 | .84 | .26 |
| negative | I feel guilty | 1.97 | 0.90 | 1.8 | 1 | 5.1 | .84 | .22 |
| Trait-Affect | | | | | | | | |
| positive | I am frightened | 4.48 | 0.67 | 4.6 | 2.0 | 5.9 | .84 | |
| negative | I am confused | 3.13 | 0.87 | 3.1 | 1.2 | 5.2 | .87 | |

*Note.* *M* = mean; *SD* = standard deviation; *med.* = median; *min* = minimum; *max* = maximum; PP Load = Physical and Psychological Load; Out-Exp. = Outcome-Expectancy. *a* = Cronbach's alpha, for DIAMONDS, Situation 5, and State-Affect, nested alpha as recommended by Nezlek (2017) was computed. ICC = the intra-class correlation based on the unconditional effects model, with situations nested in persons.

## Results

### Bivariate Correlations

Table 3.1.2 displays the bivariate Pearson-correlations of all scales that were assessed for each situation as well as the significance levels. Note that 21 of the 26 correlations between Positive Affect and Negative Affect with the situational perception dimensions (DIAMONDS and Situation 5) were significantly different from zero (at $p < 0.05$). Some of the correlations between situational perception variables and affect were moderate to large. The averaged absolute correlation between PA and the situational perception dimensions was moderate $r = .34$ [.25; .43], $p < .001$, and the averaged absolute correlation between NA and the situational perception dimensions was also moderate $r = .36$ [.27; .44], $p < .001$.

Table 3.1.2

*Bivariate Pearson-correlations of the DIAMONDS, the Situation 5, and Positive and Negative Affect*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Duty | | | | | | | | | | | | | | |
| 2. Intellect | .387*** | | | | | | | | | | | | | |
| 3. Adversity | .311*** | .259*** | | | | | | | | | | | | |
| 4. Mating | -.011 | .186*** | .237*** | | | | | | | | | | | |
| 5. pOsitivity | -.063 | .339*** | -.039 | .297*** | | | | | | | | | | |
| 6. Negativity | .349*** | .073 | .479*** | -.069 | -.426*** | | | | | | | | | |
| 7. Deception | .270*** | .285*** | .602*** | .208*** | -.016 | .490*** | | | | | | | | |
| 8. Sociality | .073 | .453*** | .288*** | .575*** | .571*** | -.141** | .271*** | | | | | | | |
| 9. Briskness | -.019 | .233*** | -.059 | .272*** | .663*** | -.381*** | -.018 | .423*** | | | | | | |
| 10. PP Load | .399*** | .011 | .401*** | -.094 | -.499*** | .779*** | .334*** | -.202*** | -.456*** | | | | | |
| 11. Out-Exp. | .529*** | .332*** | .045 | .154** | .263*** | -.145** | .029 | .202*** | .390*** | -.093 | | | | |
| 12. Cog. Load | .565*** | .576*** | .226*** | .014 | -.041 | .359*** | .168*** | .081 | .015 | .402*** | .388*** | | | |
| 13. Lack of St. | .113* | -.153** | .110* | -.088 | -.437*** | .330*** | .172*** | -.291*** | -.488*** | .486*** | -.161** | -.034 | | |
| 14. positive Affect | .221*** | .350*** | .035 | .225*** | .515*** | -.211*** | .063 | .361*** | .684*** | -.319*** | .547*** | .276*** | -.465*** | |
| 15. negative Affect | .194*** | .039 | .468*** | -.044 | -.372*** | .735*** | .401*** | -.157** | -.349*** | .722*** | -.232*** | .234*** | .403*** | -.297*** |

*Note.* *$p \leq .05$; **$p \leq .01$; ***$p \leq .001$; variables 1-8 are the DIAMONDS, 9-13 the Situation 5, and 14-15 positive and negative Affect; PP Load = Psychological & Physical Load; Out-Exp. = Outcome-Expectancy; Lack of St. = Lack of Stimuli.

**Regressions**

We ran three separate analyses. Using ordinary least squares linear regression, we predicted NA and PA separately with the DIAMONDS and with the Situation 5. We then predicted NA and PA with the DIAMONDS and the Situation 5 simultaneously.[18]

**DIAMONDS.** Table 3.1.3 displays the results of the regression of positive and negative affect on the eight diamonds dimensions.

PA was significantly related to two of the eight DIAMONDS dimensions: Duty ($b$ = .20 [.123; .286], $p_{robust}$ < .001), and pOsitivity ($b$ = .41 [.288; .530], $p_{robust}$ < .001). Looking at the OLS regression, two other dimensions (Intellect and Negativity) also reached significance, yet their effect was no longer significant under robust conditions. All predictors combined shared 34% of the variance with positive affect ($R^2$ = .35, $R^2_{adj.}$ = .34, $F(8, 374)$ = 25.41, $p$ < .001).

NA was related to four out of eight DIAMONDS dimensions: Duty, ($b$ = -.08 [-.129; -.025], $p_{robust}$ < .01), Adversity ($b$ = .28 [.157; .400], $p_{robust}$ < .001), Negativity ($b$ = 0.49 [.411; .560], $p_{robust}$ < .001), and Sociality ($b$ = -.08 [-.136; -.017], $p_{robust}$ < .05). All predictors combined shared 57% of the variance with negative affect ($R^2$ = .58, $R^2_{adj.}$ = .57, $F(8, 374)$ = 65.07, $p$ < .001).

[18] The results of these analyses can be found in the OSM, section *Prediction of PA and NA with Situation 5 and DIAMONDS simultaneously.*

Table 3.1.3

*Regression Results: Positive and Negative Affect on the Situational Eight DIAMONDS*

PA

| | *b* | *SE* | *β* | $SE_{rob}$ | $p_{rob}$ | *lower CI* | *upper CI* | $sr^2$ | *r* |
|---|---|---|---|---|---|---|---|---|---|
| Intercept | 1.14 | .25 | | .317 | <.001*** | .514 | 1.76 | | |
| Duty | .20 | .04 | .25 | .041 | <.001*** | .123 | .286 | .05 | .22** |
| Intellect | .08 | .04 | .11 | .041 | .056 | -.002 | .161 | .01 | .35** |
| Adversity | -.05 | .08 | -.03 | .078 | .54 | -.202 | .106 | <.01 | .04 |
| Mating | .06 | .04 | .08 | .044 | .172 | -.026 | .149 | <.01 | .23** |
| pOsitivity | .41 | .06 | .42 | .061 | <.001*** | .288 | .53 | .09 | .51** |
| Negativity | -.11 | .05 | -.13 | .057 | .057 | -.223 | .003 | .01 | -.21** |
| Deception | .04 | .06 | .04 | .067 | .523 | -.089 | .176 | <.01 | .06 |
| Sociality | -.01 | .04 | -.01 | .047 | .863 | -.101 | .085 | <.01 | .36** |

multiple $R^2$ = .35, adjusted $R^2$ = .34; $F(8,374) = 25.41$, $p < .001$

NA

| | *b* | *SE* | *β* | $SE_{rob}$ | $p_{rob}$ | *lower CI* | *upper CI* | $sr^2$ | *r* |
|---|---|---|---|---|---|---|---|---|---|
| Intercept | 1.91 | .18 | | .213 | <.001*** | .486 | 1.325 | | |
| Duty | -.08 | .03 | -.10 | .026 | .004** | -.129 | -.025 | .01 | .19** |
| Intellect | .03 | .03 | .05 | .026 | .204 | -.018 | .085 | <.01 | .04 |
| Adversity | .28 | .06 | .21 | .062 | <.001*** | .157 | .4 | .02 | .47** |
| Mating | .02 | .03 | .03 | .026 | .478 | -.033 | .07 | <.01 | -.04 |
| pOsitivity | -.05 | .04 | -.06 | .044 | .273 | -.135 | .038 | <.01 | -.37** |
| Negativity | .49 | .04 | .62 | .038 | <.001*** | .411 | .56 | .19 | .73** |
| Deception | .01 | .04 | .01 | .049 | .802 | -.084 | .109 | <.01 | .40** |
| Sociality | -.08 | .03 | -.13 | .03 | .012* | -.136 | -.017 | .01 | -.16** |

multiple $R^2$ = .58, adjusted $R^2$ = .57; $F(8,374) = 65.07$, $p < .001$

*Note.* * $p \leq .05$; ** $p \leq .01$; ***$p \leq .001$; PA = positive affect; NA = negative Affect; *b* = unstandardized regression weight; *SE* = standard error; *β* = standardized regression weight; $SE_{rob}$ = robust standard error; $p_{rob}$ = *p*-value based on $SE_{rob}$, lower and upper CI = upper and lower bound of 95% confidence interval, based on $SE_{rob}$; $sr^2$ = semi-partial correlation squared; *r* = bivariate Pearson-correlation between predictor and PA or NA.

**Situation 5.** Table 3.1.4 shows the results for the same analyses but using the Situation 5 as predictor variables.

PA was related to four out of five Situation 5 simultaneously: Briskness ($b$ = .35 [.261; .433], $p_{robust}$ < .001), Outcome-Expectancy ($b$ = .22 [.143; .3], $p_{robust}$ < .001), Cognitive Load ($b$ = .14 [.075; .201], $p_{robust}$ < .001), and Lack of Stimuli (b = -.12 [-.208; -.032], $p_{robust}$ < .01). Psychological and Physical Load was no longer significantly related using the cluster robust standard errors (b = -.07 [-.145; .002], $p_{robust}$ = .057). All predictors combined shared 61% of the variance with positive affect ($R^2$ = .61, $R^2_{adj.}$ = .61, $F$(5, 377) = 119, $p$ < .001).

NA was related to two of the Situation 5 variables, Psychological and Physical Load ($b$ = .47 [.384; .552], $p_{robust}$ < .001) as well as Outcome-Expectancy ($b$ = -.16 [-.223; -.094], $p_{robust}$ < .001). All variables combined shared 55% of the variance with negative affect ($R^2$ = .56, $R^2_{adj.}$ = .55, $F$(5, 377) = 94.03, $p$ < .001).

Table 3.1.4

*Regression Results: Positive and Negative Affect on the Situation 5*

predictors for PA

| | *b* | *SE* | *β* | $SE_{rob}$ | $p_{rob}$ | *lower CI* | *upper CI* | $sr^2$ | *r* |
|---|---|---|---|---|---|---|---|---|---|
| (Intercept) | 1.51 | .18 | | .274 | <.001*** | .967 | 2.044 | | |
| Briskness | .35 | .03 | .47 | .044 | <.001*** | .261 | .433 | .13 | .68** |
| PP Load | -.07 | .03 | -.09 | .037 | .057 | -.145 | .002 | <.01 | -.32** |
| Out-Exp. | .22 | .03 | .25 | .04 | <.001*** | .143 | .30 | .04 | .55** |
| Cognitive Load | .14 | .03 | .20 | .032 | <.001*** | .075 | .201 | .03 | .28** |
| Lack of St. | -.12 | .03 | -.14 | .045 | .008** | -.208 | -.032 | .01 | -.47** |

multiple $R^2$ = .61, adjusted $R^2$ = .61; $F(5,377) = 119$, $p < .001$

predictors for NA

| | *b* | *SE* | *β* | $SE_{rob}$ | $p_{rob}$ | *lower CI* | *upper CI* | $sr^2$ | *r* |
|---|---|---|---|---|---|---|---|---|---|
| Intercept | .82 | .18 | | .199 | <.001*** | .431 | 1.215 | | |
| Briskness | .05 | .03 | .08 | .033 | .112 | -.012 | .117 | <.01 | -.35** |
| PP Load | .47 | .03 | .69 | .043 | <.001*** | .384 | .552 | .24 | .72** |
| Out-Exp. | -.16 | .03 | -.20 | .033 | <.001*** | -.223 | -.094 | .03 | -.23** |
| Cognitive Load | .02 | .03 | .04 | .031 | .468 | -.038 | .083 | <.01 | .23** |
| Lack of St. | .06 | .03 | .08 | .044 | .185 | -.028 | .143 | <.01 | .40** |

multiple $R^2$ = .55, adjusted $R^2$ = .55; $F(8,374) = 65.07$, $p < .001$

*Note.* *$p \leq .05$; **$p \leq .01$; ***$p \leq .001$; PA = positive affect; NA = negative Affect; PP Load = Psychological & Physical Load; Out-Exp. = Outcome-Expectancy; Lack of St. = Lack of Stimuli; *b* = unstandardized regression weight; *SE* = standard error; *β* = standardized regression weight; $SE_{rob}$ = robust standard error; $p_{rob}$ = *p*-value based on $SE_{rob}$; lower and upper CI = upper and lower bound of 95% confidence interval, based on $SE_{rob}$; $sr^2$ = semi-partial correlation squared; *r* = bivariate Pearson-correlation between predictor and PA or NA.

**Situation 5 and DIAMONDS.** In a combined model with all situational perception variables predicting positive affect, only the regression weights for the Situation 5 variables remained significant and $R^2$ still amounted to .63 ($R^2$ = .63, $R^2_{adj.}$ = .61, $F(13, 369) = 47.73$, $p < .001$), slightly more than was found for the Situation 5 alone. For Negative Affect, eight out of 12 predictors were significant and $R^2$ amounted to .64 ($R^2$ = .66, $R^2_{adj.}$ = .65, $F(13, 369) = 53.99$, $p < .001$). Thus, the increment compared to the regression based on only one set of situational perception predictors was much stronger than for positive affect.[19]

## Discussion

The results presented by Parrigon and colleagues (Parrigon et al., 2017) could be replicated. Both taxonomies share a significant and substantial amount of variance with positive and negative affect, and individual dimensions showed very high correlations with affect. Moreover, multiple regressions revealed that this overlap is stronger for some dimensions than for others.

### Situational Perception and Affect

**DIAMONDS and affect.** Almost all DIAMONDS are correlated substantially with PA (an exception being Adversity and Deception) and NA (except Intellect and Mating) on a bivariate level. This sheds some light on the DIAMONDS dimensions: Some are related to increased Negative Affect (Duty, Adversity, Negativity, Deception) or decreased Positive Affect (Negativity). For others, this is the opposite, such as Sociality, Positivity, and Duty (related to increased PA). Duty is an exceptional case, since it is related positively to PA ($r$ = .22) and NA ($r$ = .19) at the same time, in the multiple regression, the effect of Duty on NA is reversed, meaning that participants experience more PA and less NA when Duty was increased and all other DIAMONDS are controlled for.[20] In a regression all of them combined explain a significant amount of variance in affect, either positive or negative. However, looking at the regression weights revealed that the overlap between situational perception and affect might be driven by specific dimensions of situational perception. Maybe not surprising, pOsitivity was strongly related with positive affect and Negativity with negative affect. Above we had argued that a strong overlap might question discriminant validity of situational perception scores. The results now show that this is true but might be especially the case for pOsitivity and Negativity. However, this does not mean that the other DIAMONDS dimensions are free of affect as evidenced by the bivariate correlations. We will discuss recommendations below.

---

[19] The results of these analyses can be found in the OSM, section *Prediction of PA and NA with Situation 5 and DIAMONDS simultaneously*.

[20] This effect is driven by Negativity, meaning that, after controlling for the negative aspects of a situation, perceiving Duty leads to positive affect. See OSM, *section Additional exploratory analyses, suppression effect of Duty*.

**Situation 5 and affect.** All Situation 5 correlate significantly and substantially with PA and NA. PA increased with increased Briskness, increased Outcome-Expectancy, increased Cognitive Load and decreased Lack of Stimuli. On the other hand, NA is positively related to Physical and Psychological Load and negatively to Outcome–Expectancy. Similar to Duty, Cognitive Load is positively associated with both PA and NA. Since these two affect dimensions are more or less independent from each other, a situation that is perceived as high on Cognitive Load can go along with positive and negative affect at the same time. For example, learning for an exam can lead to positive affect (feeling that one increased his or her knowledge) as well as negative affect (feeling the strain of constant learning efforts). Within the regression analyses, similar results to the DIAMONDS occurred with briskness (PA) and PP load (NA) revealing specifically strong overlaps with affect.

**Situation 5, DIAMONDS, and Affect.** To investigate which dimensions share more variance with affect, we included the Situation Five and the DIAMONDS simultaneously in a model to predict PA and NA. In this model, only the Situation 5 predicted PA, whereas the results for NA were more mixed: Duty, Intellect, Adversity, Negativity, Sociality, Psychological and Physical Load, Outcome-Expectancy, and Lack of Stimuli significantly predicted NA.

**Nomological Net of Situation Perception**

As Horstmann and colleagues have suggested (Horstmann, Rauthmann, et al., 2018), it is very likely that across different situational taxonomies similar or even redundant dimensions can be found. The establishment of a nomological net of situational perception dimensions *within* situational perception dimensions as well as *across* different constructs (e. g. situational perception, mood, personality) is crucial to an understanding of the construct itself (Cronbach & Meehl, 1955; Horstmann, Rauthmann, et al., 2018; Ziegler, 2014c; Ziegler et al., 2013).

The bivariate correlations provided in Table 3.1.2 give further insight into the nomological net of situational perception. Some dimensions across the two taxonomies correlate rather highly (Duty - Cognitive Load ($r$ = .57), Duty - Outcome-Expectancy ($r$ = .53), pOsitivity - Briskness ($r$ = .66), Negativity - Physical and Psychological Load ($r$ = .78)). Other dimensions fully expand and enrich the nomological net, for example Mating, Sociality, Adversity, or Lack of Stimuli. This is the first empirical evidence for convergent validity of the Situation 5 and DIAMONDS scores.

Note however, that all of the dimensions correlate significantly either with PA or NA, or both. This could be problematic for the development of a nomological net of situational perception. During the construction process of all recently published situational perception taxonomies (i.e., N. A. Brown et al., 2015; Gerpott et al., 2017; Horstmann, Ziegler, et al., 2018; Parrigon et al., 2017; Rauthmann & Sherman, 2016a) participants described and rated situations

experienced the previous day. It is reasonable to assume that those experiences are colored with affect. In that way, the reports of these situations can also not be neutral. In fact, this notion is very much in line with the general conceptualization of situational perception as the result of an individual information processing (Ziegler & Horstmann, 2015). Clearly, such an information processing is also driven by affect (Wyer, Clore, & Isbell, 1999). This, however, has consequences for the interpretation of the current findings. First, this "contamination" of affect and situational measures during the construction process might lead to affect being incorporated in measures of situational perception. Second, given that ratings of situational perception are correlated with affect in a substantial way, the correlations *among* dimensions of situational perception are very likely inflated. This hinders the clear understanding of the structure and nomological net of situational perception measures.

There are two possible solutions to this problem: First, To disentangle personal affect and situational perception, multiple ratings from raters *ex situ* could be obtained (Rauthmann, Sherman, & Funder, 2015b), for example in an experimental setting. The shared variance of *ex situ* raters is more likely to catch the true characteristic of the situation since idiosyncratic variance of raters is cancelled out. This approach is similar to other-ratings of personality vs. self-ratings of personality (e.g., Hofstee, 1994), which means that this it entails similar problems such as knowledge asymmetry (Simine Vazire, 2010). Thus, such approaches might introduce other problems. Second, positive and negative affect could be considered as a method factor when examining the structure of situation dimensions, similar to common method variance (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003) or socially desirable responding (Ziegler & Buehner, 2009). Modelling the influence of PA and NA would thus increase the specificity of situational perception ratings.

Finally, if positive affect and negative affect are relevant for or correlated with the description of situational perceptions, one could ask if other, more narrow emotions (e.g., fear or disgust) would also be important for the description of situations. Some of these (e.g., fear) may already be incorporated in situational measures (Adversity from the DIAMONDS). This has to be investigated in a more thorough and systematic way.[21]

[21] We thank the Editor for pointing this out.

**Barking up the wrong tree?**

Affect, positive and negative, are tied to a number of outcomes, especially behavior (Lench et al., 2011). Since nearly all situational perception dimensions across taxonomies are related to affect, it is possible that effects that are currently attributed to situational perception (or, the influence of the situation) could, in fact, be explained with positive and negative affect. Even though this may not be the case for all dimensions (e.g., Mating [DIAMONDS], or Lack of Stimuli [Situation Five], or Typicality [CAPTION]), it could certainly be the case for some dimensions, especially pOsitivity, Negativity, or Briskness. It remains an empirical question whether participants are able to differentiate between the assessment of a situation and the assessment of their own affect within this situation. Unless affect is considered explicitly in situation research, it might be the case that we are currently barking up the wrong tree. Future research needs to address these issues and establish discriminant validity for situational perception and affect, as well as predictive validity for situation perception over affect.

## Limitations

When sampling situations, both the sample of situation as well as the sample of participants has to be considered (Horstmann, Ziegler, et al., 2018). Participants and the situations they report are not independent (see situational selection and construal; Rauthmann, Sherman, Nave, et al., 2015). For example, certain participants will encounter and report only certain situations. This sample included mainly female students and variability of situations is likely restricted. Some dimensions were never reported in their extreme manifestations (Adversity, Deception, Mating, NA, see Table 3.1.1), and it is unlikely (and undesirable) that participants will never experience extreme situations. This *report-bias* reduces the variance in situations, which can have two effects: It may either decrease the covariance of situational dimension, due to reduced variance; or it may inflate the covariance of two dimensions. Duty and Intellect for example are highly correlated, which is likely due to the student sample, whose duty is usually intellectual. However, it is almost certain that a lot of situations exist that are high on Duty, but not Intellect. Addressing this bias may be possible in longitudinal designs, using heterogeneous samples, or experimental settings that force participants into certain situations (e. g. in a speed-dating study, Asendorpf, Penke, & Back, 2011). Further, the reliability estimates of some measures (indicated by Cronbach's alpha, Table 3.1.1) was rather low. pOsitivity and Outcome Expectancy for example had comparatively low alphas, which might indicate unique yet systematic variance attributable to the specific situation. However, additional preliminary multi-level analyses that took the nested structure into account did not alter the results or their interpretation.

The current cross-sectional data furthermore do not allow making any statements about directionality of effects or causality. Even though we predicted affect in the multiple regression models, it could also have been the antecedent to situational perception, as discussed earlier. This could only be thoroughly examined using an experimental design. However, the current data do also not reject any of the theories presented earlier, and additional research will most likely be successful in discovering valuable connections between affect and situations.

## Conclusions and Outlook

Using two recently developed taxonomies of situational perception we could show that affect and situational perception are correlated. Multiple regressions revealed that the DIAMONDS and the Situation 5 explain variance in positive and negative affect. These results call for a more thorough investigation of the nomological net of taxonomies of situational perception: Some dimensions, even though not all, clearly lack discriminant validity. Furthermore, person-states may not only include affect, but also motivation (Rauthmann, 2016) or even physical states such as satiation, thirst, or exhaustion. Linking these person variables to measures of situational perception would clearly be beneficial.

The ultimate goal remains the explanation of past and the prediction of future behavior. The development of integrated situation taxonomies and one-dimensional test scores for different situational perceptions will advance the understanding of the dynamic person in his or her ever-changing environment.

## 3.2 Affect and Situation Perception – Independent Predictors of Behavior?

The correlation between positive affect, negative affect, and dimensions of situation perception from two different situational taxonomies (the DIAMONDS and the Situation Five) have implications for understanding the phenomenon of situation perception (Horstmann & Ziegler, 2018). Although it is likely that affect and situation perception are correlated (see above, as well as Rauthmann et al., 2015), the question has to be asked if this overlap can be problematic for the interpretation of effects that are attributed to situation perception. Affect (and emotion) were shown to have strong effects on behavior (Lench et al., 2011; Wilson et al., 2017; Wilt & Revelle, 2017), and this effect could "swallow" the effect of situation perception.

Let us consider the following example from above (section 3.1): A student taking an exam reports that the situation was adverse. This perceived adversity may possibly predict dishonest behavior such as cheating during the exam. At the same time, the affect of the student was not examined. It could be possible that, in this case, not the perception of the situation was responsible for behavior, but the accompanying affect was.

This is important for two reasons: First, it matters for understanding the effects of situations. Affect is a well-researched phenomenon – if situation perception and affect were indistinguishable, many resources would be invested in the research on situation perception without gaining new insights. Second, scales to assess affect and emotions have been used and validated (e.g., Krohne et al., 1996; but see also Weidman, Steckler, & Tracy, 2017 for problems related to emotion assessment). It would be a fruitless exercise to design new scales for a new phenomenon (i.e., situation perception) if old scales (for affect) might do the job just as well.

Despite the overlap of situation perception and affect, the litmus test for both phenomena remained to be realized, especially given that the correlation of affect and situation perception was substantial, but not perfect: A direct, head-to-head comparison in predicting relevant outcomes. If, after the inclusion of affect, situation perception predicted no additional variance in an outcome, this would diminish the need for further investing in scales and taxonomies for situation characteristics. It was thus the aim of our study *Unveiling an Exclusive Link: Predicting Behavior with Personality, Situation Perception, and Affect in a Pre-Registered Experience Sampling Study* (Horstmann, Rauthmann, Sherman, & Ziegler, submitted) to shed more light on the relation of personality, affect, situation perception, and behavior.

### *3.2.1 Article Summary: Unveiling an Exclusive Link*

**Background.** Sherman and colleagues (2015) examined how personality and the perception of situations predicted behavior and happiness in daily life. Using an experience sampling design, the authors required $N$ = 210 participants to first rate their personality traits on the six HEXACO dimensions (Ashton & Lee, 2007, 2009; Lee & Ashton, 2008). Participants were then invited several times daily to report their behavior on the six HEACO dimensions, their subjective happiness, and the perception of the current situation on the Situational Eight DIAMONDS (Rauthmann & Sherman, 2016c).

Sherman and colleagues used multi-level modelling to examine the relation of behaviors and situation perception. Specifically, the authors assumed that only certain situation perception dimensions were predictive of certain behaviors. For example, the perception of Duty should lead to more conscientious behavior but not to more honest behavior. Overall, Sherman and colleagues could show that both personality and situation perception were independent predictors of behavior and happiness in daily life.

**Study 1.** Based on the assumption that affect and situation perception overlap (Kuppens, 2009; Rauthmann, Sherman, & Funder, 2015b; Sander et al., 2005), and supporting empirical evidence (Edwards & Templeton, 2005; Horstmann & Ziegler, 2018), we re-analyzed the data by Sherman and colleagues. We included happiness in a situation not as an outcome-variable in the multi-level models but used it as a further predictor. Additionally, and to examine if only certain behaviors are linked to specific dimensions of situation perception, we crossed all eight DIAMONDS dimensions with all six HEXACO dimensions.

**Study 1 – Analyses.** Multi-level models were used to re-analyze the data and predict behavior. Personality (corresponding to the behavior, e.g., trait conscientiousness for the prediction of conscientiousness behavior), trait subjective happiness, and average situation perception were included as a predictor at person level. Situation perception and state happiness were included at the occasion/situation level. First, we crossed all DIAMONDS with all HEXACO dimensions (e.g., predicting conscientiousness behavior with all DIAMONDS dimensions separately). Second, we extended the model and included happiness at trait and state level as predictors in all models.

**Study 1 – Results.** The re-analyses revealed two remarkable findings: First, nearly all DIAMONDS dimensions predicted nearly all dimensions of behavior. Second, the inclusion of happiness at trait and state level eliminated some of the effects of the situation perception dimensions. Specifically, only those dimensions were no longer significant that were also predicted not to be significant by Sherman and colleagues (Rauthmann, Jones, & Sherman, 2016; Sherman et al., 2015).

**Study 1 – Conclusion.** Although an overlap between happiness and situation perception occurred, this overlap did not threaten the predictive validity of the situation perception dimensions. Quite the contrary, including happiness unveiled the exclusive links that were expected by Sherman and colleagues (2015) in the first place. However, the analyses were exploratory in nature. We therefore decided to pre-register our findings and expectations and replicate Study 1 in an independent sample.

**Pre-registration.** Between Study 1 and Study 2, we pre-registered findings from Study 1 as well as the hypotheses, the methods (determining the sample size of $N$ = 250), and the procedure on the Open Science Framework. Thus, Study 2 can be thought of as strictly confirmatory.

**Study 2.** In Study 2, we aimed at replicating and extending the results from Study 1. More specifically, we were interested three questions: 1) Can we replicate the finding that, without controlling for happiness, nearly all dimensions of situation perception predict nearly all behaviors? 2) After including

happiness at trait and state level, do only logically coherent effects of DIAMONDS on HEXACO behavior persist? 3) If better measures of positive and negative affect were included as predictors (both at trait and state level), would the DIAMONDS still explain within-person variance in behavior? All expectations, the study design, and analyses were pre-registered on the Open Science Framework. Furthermore, we changed some of the hypotheses from Sherman and colleagues (2015) to reflect findings from Study 1. For example, Sherman and colleagues (2015) assumed that the perception of Duty and open behavior were not related; however, as we were to collect a student sample which was likely to experience Duty in situations where open behavior is required (e.g., learning), we updated our expectations accordingly.

**Study 2 – Study Design.** We aimed at replicating Sherman and colleagues' (2015) original study as closely as possible and used the same measures. Additionally, we included measures of positive (PA) and negative affect (NA) at state level (Hampel, 1977). $N = 274$ participants provided sufficient data for subsequent analyses.

**Study 2 – Analyses.** We used multi-level models to predict participants' behavior with the corresponding personality trait (e.g., we predicted conscientious behavior with conscientiousness at trait level), situation perception at state and trait level (i.e., at trait level, situation perception was the average of all situation perception assessments in one dimension across all measurement occasions during experience sampling), as well as happiness, PA, and NA (all at state and trait level). This model was extended in a step-wise manner to reflect the questions 1-3: For 1) only personality and the DIAMONDS were included as predictors, for 2) happiness was included at trait and state level, and for 3) PA and NA were included at trait and state level.

**Study 2 – Results.** Concerning the three questions, 1) we could show that nearly all dimensions of situation perception were indeed again substantially related to nearly all dimensions of behavior; 2) after including happiness at trait and state level, mainly the logically coherent DIAMONDS predicted variance in behavior; and 3) including PA and NA at trait and state level did not change the effects of the DIAMONDS on behavior substantially compared to including happiness at trait and state level only.

**Study 2 – Conclusion.** Study 2 revealed that the effects examined in Study 1 were in fact robust and could be replicated. Several very specific predictions were made for the DIAMONDS at person and occasion level. Although the predictions for the DIAMONDS on occasion level (i.e., during experience sampling) were mainly accurate, those on person level (i.e., aggregated DIAMONDS dimensions) were not very accurate. Thus, whereas we were confident that we understood the effects of the DIAMONDS at occasion level, this could not be said with confidence for the effects of the DIAMONDS at person level. Furthermore, although PA and NA predicted variance in behavior, even above and beyond variance explained by happiness, it did not alter the effects of the DIAMONDS. Thus, controlling for happiness in future applications of the DIAMONDS – or situation perception in general – could be sufficient.

**General Conclusion.** Across Study 1 and Study 2, we could show that a specific, logically coherent link between the DIAMONDS and behavior existed, but that this link was only unveiled after controlling for happiness or affect. As we previously discussed (Horstmann & Ziegler, 2018), the overlap of affect and situation perception did threaten discriminant validity of situation perception – however, this study revealed that the DIAMONDS show good predictive discriminant validity and provide a useful tool in a personality researcher's toolbox.

### *3.2.2 Article: Unveiling an Exclusive Link*

**Unveiling an Exclusive Link: Predicting Behavior with Personality, Situation Perception, and Affect in a Pre-Registered Experience Sampling Study**

Kai T. Horstmann[1], John F. Rauthmann[2], Ryne A. Sherman[3], Matthias Ziegler[1]

[1]Humboldt-Universität zu Berlin (Germany)

[2]Wake Forest University (USA)

[3]Hogan Assessment Systems, Tulsa (USA)

This article has been submitted for publication in an academic journal and is currently under review.

**Abstract**

Affect and situation perception are intertwined in any given situation, but the extent to which both predict behavior jointly and uniquely has not yet been examined so far. Using two studies with experience sampling methodology, we examine how trait-like variables (Big Six, trait affect, general situation experience) and state-like variables (momentary affect, situation perceptions of the DIAMONDS) account for variance in behavioral states of the Big Six. In Study 1, we re-analyzed data from Sherman, Rauthmann, Brown, Serfass, and Jones (2015) and found that situation perception explained variance in behavior in logically coherent ways, but only after considering happiness as an additional predictor. These results were replicated in pre-registered Study 2, in which positive and negative affect were additionally assessed as distinct variables. Based on both studies, we conclude that personality, affect, and situation perception are independent predictors of behavior in daily life. Importantly, situation perceptions and affect do overlap, but are not the same or redundant to each other. Indeed, theoretically justified and logically coherent links between situation perceptions and behavioral states remain intact once affect is controlled for, while the other links not predicted by theory disappear. These results have implications for personality theories as well as appraisal theories of emotion.

# Unveiling an Exclusive Link: Predicting Behavior with Personality, Situation Perception, and Affect in a Pre-Registered Experience Sampling Study

"Different [wo]men may be differently affected by the same object, and the same [wo]man may be differently affected at different times by the same object."
– Benedictus de Spinoza (1677). Ethics, Part 3: *On the Origin and Nature of Emotions*

The idea that people behave consistently in their differences from others, yet vary in their own behavior across situations and time is hardly a new one. However, given that people's reactions to the same object may be different at different times, and different people may react differently to the same object, how can behavior be reliably explained? Kurt Lewin (1936) proposed that a person's momentary behavior B can be predicted by their current mental states P and the actual environment E they currently are in: $B = f(P, E)$. The person variables P have later been understood in terms of personality traits, and the environmental variables E as parts of the psychological situation (e.g., Funder, 2006). In an attempt to unravel how these variables functioned together, Sherman and colleagues (2015) examined in a large experience sampling study how personality traits and psychological situation characteristics predicted momentary personality expressions (states) in daily life. They successfully demonstrated that both traits and situation characteristics independently accounted for variance in behavior. However, Sherman and colleagues' results are difficult to interpret until affect, an important and theoretically relevant as well as potentially competing predictor of behavior in a situation (Lench et al., 2011), has been integrated into the analyses. In the present article, we therefore examine the role of affect and its interplay with other, established predictors of behavior – personality and situations. We thus are able to not only replicate and corroborate, but also extend prior research on the interplay between persons and situations.

## Background

### Explaining Behavior

Before engaging in a discussion on how personality, affect, and situations together shape behavior, we first define our terms. Behavior, broadly defined, includes "actions, cognition, motivation, and emotions" (Fleeson & Noftle, 2008, p. 1358). A more narrow definition focuses on actions by a person that are potentially available to careful observers using normal sensory processes (Furr, 2009a, p. 372), which would exclude thoughts and feelings as long as these are not explicitly expressed. For the current article, we operationalize behavior as the self-reported

account of current actions in a given situation as often done in current experience sampling practices (Sherman et al., 2015; Wilson et al., 2017; Wilt & Revelle, 2017). In terms of the content of such behavioral states, we use the Big Six framework which has already been used in numerous other studies effectively (e.g., Fleeson, 2001; Fleeson & Gallagher, 2009). This is consistent with personality states being defined as "having the same affective, behavioral, and cognitive content as a corresponding trait (…), but as applying for a shorter duration" (Fleeson & Jayawickreme, 2015, p. 84), or, as Baumert and colleagues (2017, p. 528) put it "quantitative dimension[s] describing the degree/extent/level of coherent behaviours, thoughts, and feelings at a particular time." In this work, we focus on three broad categories of predictors of such momentary behavior: personality, situations and their perceptions, and affect.

**Personality.** A personality trait, in its broadest definition, is a stable characteristic of a person (Funder, 2001). More narrowly defined, traits have been understood in taxonomic systems housing five to six broad factors, such as the Big Five or the HEXACO model (Ashton & Lee, 2007; Funder, 2001; John & Srivastava, 1999; Lee & Ashton, 2008). These stable traits allow accounting for variance in behavior and predicting important outcomes (Ozer & Benet-Martínez, 2006; Roberts et al., 2007). At the same time, traits – operationalized usually as a fixed value assigned to a person – are not well suited to explain variations in behavior. A person usually characterized as extraverted may behave introvertedly, and this seemingly "inconsistent" behavior cannot be fully accounted for by the trait (Fleeson, 2004; Mischel & Shoda, 1995; Sherman et al., 2010; Shoda & Mischel, 2000). As Sherman and colleagues (2015) summarized, recent personality theories acknowledge that manifestations of personality, so-called states, can be different from the corresponding personality trait (Whole Trait Theory: Fleeson & Jayawickreme, 2015; Five Factor Theory: McCrae & Costa, 2008; Cognitive-Affective Processing Systems: Mischel & Shoda, 1995; Trait Activation Theory: Tett & Guterman, 2000).

A state should be understood as occasion-specific (Geiser et al., 2015; Steyer et al., 2015, 1999). In most models this means that states are not just error, but actually contain variance that is both systematic and potentially explainable by external situational influences and other occasion-specific characteristic of the person (e.g., ongoing affect). These points are important because explanations of states should not only be sought in characteristics of the external or current situation (Sherman et al., 2015), but also in internal specifics of the person *in the situation* (Cattell, 1963; Horstmann, Ziegler, et al., 2018; Read et al., 2017).

**Situations**

The situation plays an important role in the explanation and prediction of behavior, which is supported by ample empirical evidence (Deinzer et al., 1995; Fleeson, 2004; Furr & Funder,

2004; Mischel & Peake, 1982; Sherman et al., 2010, 2015; Shoda et al., 1994). However, to examine the role of situations more generally, it is first necessary to describe situations systematically and second to provide measures for situational information.

**Situation taxonomies and measures.** Rauthmann and colleagues (Rauthmann, 2015; Rauthmann et al., 2014; 2015b) proposed that there are three kinds of situational information: cues (physical elements in an environment), characteristics (psychological meaning of elements), and classes (groups of situations similar in cues and/or characteristics). Rauthmann and colleagues (2015b) argued that one of the most fruitful levels to study and understand situations is characteristics as situations can be described and compared on a set of continuous dimensions (Horstmann, Rauthmann, et al., 2018). The question then is which dimensions should be used?

In the last 50 years, several situational taxonomies have been developed, and most taxonomies focused on the description of situation characteristics (Horstmann, Rauthmann, et al., 2018). However, only few and recently published taxonomies also provided psychometrically validated assessment tools to measure characteristics (i.e., N. A. Brown et al., 2015; Gerpott et al., 2017; Parrigon et al., 2017; Rauthmann et al., 2014; Ziegler et al., submitted). The most widely used taxonomy so far is the DIAMONDS taxonomy (Rauthmann et al., 2014; Rauthmann & Sherman, 2016a, 2016c). According to this taxonomy, situations can be described on eight dimensions: Duty (e.g., "work needs to be done"), Intellect (e.g., "deep thinking is required"), Adversity (e.g., "someone is being threatened"), Mating (e.g., "potential romantic partners can be attracted"), pOsitivity (e.g., "the situation is positive, playful"), Negativity (e.g., "the situation could be tainted by negative feelings"), Deception (e.g., "someone can be deceived"), and Sociality (e.g., "meaningful social interaction is possible or required"). Correlations of the DIAMONDS with other situation measures suggest that several proposed taxonomies converge on a broader level to at least six general dimensions (Horstmann, Rauthmann, et al., 2018; Rauthmann & Sherman, 2018a). To assess situation characteristics, participants are usually asked how they perceive or interpret a given situation of interest. These subjective situation judgments can then be used to explain behavior of the person in that situation (Rauthmann et al., 2014; Sherman et al., 2015).

**Situation perception and behavior.** Several studies have examined the powerful and robust effect that situations can have on behavior. In fact, the manipulation of situation cues represents one of the central experimental designs of social psychology (Horstmann, Rauthmann, et al., 2018; Krueger, 2009). Further, situation characteristics have been shown to predict the expression of certain, related behaviors (Fleeson, 2007; T. a Judge, Simon, Hurst, & Kelley, 2014), and this is also true for situation classes (e.g., Geukes et al., 2017; Oud et al., 2017). However, as Sherman and colleagues (2015) point out, most of the situation measures used in these studies

were defined ad hoc rather than systematically. With the advent of situational taxonomies and measures, effects of situations have been examined more systematically. Situation characteristics that were assessed with standardized and validated measures allowed theoretically meaningful predictions of behavior in several instances (N. A. Brown et al., 2015; Gerpott et al., 2017; Parrigon et al., 2017; Rauthmann et al., 2014; Ziegler et al., submitted). For example, Gerpott and colleagues (2017) showed that descriptions of social interdependence in a situation explained substantive amounts of variance in cooperation. Further, several situation dimensions have been explicitly linked to behavioral manifestations of personality traits (de Vries, Tybur, Pollet, & van Vugt, 2016; Parrigon et al., 2017; Rauthmann et al., 2016; Sherman et al., 2015).

More specifically, the DIAMONDS dimensions were argued to have exclusive ties to the HEXACO dimensions of personality (Rauthmann et al., 2016; Sherman et al., 2015). For example, perceiving Duty in a situation should predict conscientious behavior, but not honest behavior. Sherman and colleagues predicted and examined several specific combinations of situation characteristics and personality state behavior (see upper third of Table 3.2.1). In terms of "discriminant associations," it is important here that the perception of a situation on a specific characteristic dimension is linked to a specific behavior (e.g., perceiving Duty leads to more conscientious behavior, but not to more agreeable behavior). If this were not the case, and any characteristic dimension was linked to any behavior, at least two explanations would be possible. First, methodological issues could drive undifferentiated associations. If this were the case, we would need to question if discriminantly valid assessments of characteristics and behavior can be achieved at all in real-life or if participants' responses follow an artificial positive-manifold pattern (similar to a common-method factor or socially desirable responding) which is unrelated to the (specifics of the) current situation. Such a methods artifact amongst all data collected during experience sampling could in fact explain the successful predictions of behavior with situation perception presented by Sherman and colleagues (2015). However, such a positive manifold would not reflect specific relations between situation perceptions and behavior.

Second, participants may evaluate real-life situations on a broader, more general level and do indeed not thoroughly discriminate nuances of situation characteristics (i.e., such indiscrimination is not just a product of invalid measures or special response styles or sets). This effect could be especially present in repeated, intense experience sampling where participants may evaluate situations on a much broader level (e.g., approach – avoidance, positive – negative, etc.). Such "general factor" evaluations could be in line with appraisal theories of affect where the affect of a perceiver may be driving how the situation is perceived (Rauthmann & Sherman, 2018b; Rauthmann, Sherman, & Funder, 2015b). This would severely question the validity of in situ

measures of situation perception. It is therefore not only important to examine the specificity of the situation-behavior links proposed in Table 3.2.1, but also to gauge the role of other evaluative constructs, such as affect.

Table 3.2.1

*Predicted Effects for Situation Perception in Study 1 and Study 2*

| States | D | | I | | A | | M | | O | | N | | De | | S | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | L1 | L2 | L1 | L2 | L1 | L2 | L1 | L2 | L1 | L2 | L1 | L2 | L1 | L2 | L1 | L2 |
| *Sherman and colleagues (2015) original predictions* | | | | | | | | | | | | | | | | |
| H | | | | | | | | | | | | | - | | | |
| E | | | | | + | | | | | | + | | | | | |
| X | | | | | | | | | | | | | | | + | |
| A | | | | | | | +f | | + | | | | - | | + | |
| C | + | | | | | | | | | | | | | | | |
| O | | | + | | | | | | | | | | | | | |
| D | | | | | + | | +m | | | | | | | | | |
| *Study 2: Prediction for the Replication of the Re-analyses from Study 1* | | | | | | | | | | | | | | | | |
| H | - | | - | | - | - | | | + | + | - | - | - | - | + | |
| E | + | | + | + | + | + | | | - | - | + | + | + | + | - | |
| X | - | | - | - | - | - | + | | + | + | - | - | - | - | + | + |
| A | - | | - | - | - | - | + | | + | + | - | - | - | - | + | + |
| C | + | | + | | - | - | | | | | | - | - | - | + | |
| O | + | | + | | - | - | + | | + | + | - | - | - | - | + | |
| *Study 2: Predictions for the Extension, including happiness or/and PA and NA* | | | | | | | | | | | | | | | | |
| H | | | | | - | - | | | | + | | - | - | - | | |
| E | | | + | + | + | + | | | - | - | + | + | + | + | | |
| X | | | | - | | | + | | + | + | - | - | | - | + | + |
| A | | | | - | - | - | | - | + | + | - | - | - | - | + | |
| C | + | | + | | - | - | | | - | | + | - | | - | + | |
| O | + | | + | | - | - | | - | | + | | - | - | - | + | |

*Note.* L1 = expected effects at level 1; L2 = expected effects on level 2;
Model 1: Predicting behavior from situation perception at level 1 and level 2.
Model 2: Predicting behavior from situation perception at level 2 and level 2, and Happiness at level 1 and level 2.
Model 3: Predicting behavior from situation perception at level 2 and level 2, Happiness on level 1 and level 2, and PA (positive affect) and NA (negative affect) at level 1 and level 2.
+ = a positive effect was expected; - = a negative effect was expected; empty cells indicate that no effect was expected. +m and +f means that the effect was initially (Sherman et al., 2015) predicted to be significant only for males and females, respectively.
Situation perception (columns): D = Duty; I = Intellect; A = Adversity; M = Mating; O = pOsitivity; N = Negativity; De = Deception; S = Sociality.
Behavior (rows): H = Honesty/Humility; E = Emotionality; X = eXtraversion (Sociability aspect); A = Agreeableness; O = Openness; D = Dominance. D was not included in Study 2, as we overlooked it in the re-analysis in Study 1.

**Affect**

Affect (or mood; Watson & Gray, 2007) is a diffuse, evaluative but "consciously accessible" (Russell, 2003, p. 147) state that is generally either positive or negative and activated or deactivated (Russell, 2003; Watson & Tellegen, 1985; Winkielman, Knutson, Paulus, & Trujillo, 2007; Yik, Russell, & Barrett, 1999). As such it can be "felt [as a] tendency to approach or avoid" (Larsen, 2000, p. 130). Affect can be distinguished from emotion primarily because emotions are generally directed towards and tied to an object (e.g., Gendolla, 2000; Larsen, 2000), whereas affect can have many, unknown causes (Russell & Barrett, 1999) and is more unspecific (Siemer, 2009). Further, emotions last seconds to minutes, whereas affect lasts between hours and days (Watson & Gray, 2007), but both are subject to intra-individual change (Kuppens, 2015).

Still, affect and emotion are closely tied to one another and often co-occur (Russell & Barrett, 1999; Siemer, 2009). Indeed, both are also strong predictors of cognition and behavior (Gendolla, 2000; Lench et al., 2011). However, due to its short durations, affect is especially suited as a predictor of behavior in situ. At the trait level, it has been shown that positive affect is correlated with extraversion (Fleeson, Malanos, & Achille, 2002) and negative affect with neuroticism (Costa & McCrae, 1980; Diener, Oishi, & Lucas, 2003; Richard E Lucas & Fujita, 2000). Correspondingly at the state level, acting neurotic or extraverted can lead to negative or positive affect, respectively (McNiel & Fleeson, 2006). Positive and negative affect were further related to expressions of all Big Five personality traits in daily life (Wilson et al., 2017; Wilt et al., 2017; Wilt & Revelle, 2017).[22]

**Affect and situation perception.** Affect is related to information processing (Forgas, 2000; Wyer et al., 1999). For example, positive affect was found to be associated with less careful evaluation of information of persuasive communications (Bless, Bohner, Schwarz, & Strack, 1990; Petty, Fabrigar, & Wegener, 2003), and participants with happier moods evaluated stimuli more superfluously than participants in a negative mood (Gasper & Clore, 2002). As such, affect may exert an influence on the interpretation and evaluation of situational information (Rauthmann, 2012, 2016; Rauthmann, Sherman, & Funder, 2015b).

On the other hand, appraisal theories of emotions posit that persons evaluate their environment based on their current "concerns, goals, and competencies […] and that the outcome of this appraisal process is associated with specific emotional experiences" (Kuppens, 2009, p. 255). The subjective perception of the individual's situation thus elicits (rather than is elicited by) certain affective responses (Griner & Smith, 2000; Kuppens, 2009; Kuppens & Van Mechelen,

[22] These studies support the claims and assumptions that we had when we first examined the data by Sherman and colleagues (2015). However, these studies were not published at the time and therefore did not inform the current research or any specific hypotheses in Study 1 or Study 2.

2007; Moors, 2014; e.g., Sander et al., 2005; Scherer, 2001). As Horstmann and Ziegler (2018) concluded, the perception of a certain situation characteristic (e.g., Duty as a demand to work) and the current state of the perceiver (e.g., the missing competency to cope with this demand) could result in negative affect as a downstream consequence.

Both strands of theory and research – either positing affect governing situation perception or situation perceptions governing affect – together suggest that perceived situation characteristics and affect should in general be tied to one another. However, the exact causal relation and direction(s) remain problematic to examine thus far. It may stand to reason that situation perceptions and affect influence each other or are even inextricably intertwined. In the latter case, situation perceptions may have no incremental value above and beyond affect (and vice versa) as both would be contaminated by each other or form an amalgamation. This would mean that once controlling for affect, situation perception might not be tied to behavior anymore. Regrettably, though, so far only few studies have empirically examined relations between situation perception and positive or negative affect or related constructs (i.e., Edwards & Templeton, 2005; Gerpott et al., 2017; Horstmann & Ziegler, 2018; Parrigon et al., 2017; Serfass & Sherman, 2013; Sherman, Nave, & Funder, 2012; Sherman et al., 2015). Horstmann and Ziegler (2018) examined the relations of self-reported affect and situation perception assessed with the Situation Five (Ziegler et al., submitted) and the DIAMONDS (Rauthmann & Sherman, 2016a). All situation perception dimensions were substantially correlated with either positive or negative affect (or both), and situation measures explained substantial proportions of variance in positive and negative affect ($R^2$ = .63 for positive affect and $R^2$ = .66 for negative affect). These results are very similar to the findings presented by Parrigon and colleagues (2017) and Gerpott and colleagues (2017). Finally, Sherman and colleagues (2015) showed in experience sampling data that certain situation characteristics explained variance in state happiness in daily life.

As Horstmann and Ziegler (2018) argued, such links between situation perceptions and affect could be problematic for two related reasons. First, a lack of discriminant validity of measures of situation characteristics and affect would mean both are virtually indistinguishable once measured and thus hinder a clear understanding of the nomological networks of situation perception dimensions. For example, effects that are attributed to situation perception – and thus the influence of the situation – might have to be attributed to the in situ affect of the person in the situation. Second and related to the first point, measures of situation perceptions may harbor no incremental ability above measures of affect. If they do not predict more or other variance in behavior, it is then unclear why we should use them.

Taken together, this state of affairs calls for a more thorough investigation of the relations between situation perceptions and affect in daily life, especially when both concurrently predict behavior in situ. As Wilson and colleagues (2017) suggested, variance in behavior that could not be predicted with affect could possibly be predicted with information about the situation. This is only possible, however, if situation perception measures capture something distinct from affect measures. Given the strong overlap of situation perception and affect as well as their individual ability to predict in situ behavior, we examine how both *simultaneously* predict behavior.

## The Current Studies

In this work, we examine the interplay of situation perception, affect, and behavior in daily life across two independent studies. The first study, exploratory in nature, is a re-analysis of the data from Sherman et al. (2015). Sherman and colleagues used happiness as an outcome variable which was predicted by situation perception and personality traits. We re-examined the data and employed happiness as a predictor of behavior vis-à-vis situation perception. The results of these analyses were used to pre-register an extended replication in Study 2. Whereas no specific hypotheses for the effects existed for Study 1, specific predictions were made for the combined effects of affect and situation perception on behavior in Study 2. An overview of the study details and which elements were pre-registered can be found in Table 3.2.2. All materials, analysis scripts, data sets, as well as additional results can be found in the Online Supplemental Materials (OSM: osf.io/zctv4)

Table 3.2.2

*Overview of Research Questions, Analyses, and Results*

| Research Question | Study 1 | | Study 2 | | | | |
|---|---|---|---|---|---|---|---|
| | Results in | Type | Results in | Type | PR | Sens. | Spec. |
| Sources of Behavior (Trait Expression) | *Sherman 2015, Table 3.2.3* | - | Table 3.2.4 | Conf. | Yes | - | - |
| Sources of Situation Perception (Situation Experiences) | *Sherman 2015, Table 3.2.3* | - | Table 3.2.4 | Conf. | Yes | - | - |
| Association between Personality Traits and Situation Perception (Situation Experiences) | *Sherman 2015, Table 3.2.4* | - | Table 3.2.11 | Conf. | Yes | .86 | - |
| *Specific Effects of Situation Perception on Behavior* | | | | | | | |
| Specific Effects of Situation Perception (Level 2) on Behavior | Table 3.2.6, upper half | Exp. | Table 3.2.13 | Conf. | Yes | .67 | .57 |
| Specific Effects of Situation Perception (Level 1) on Behavior | Table 3.2.7, upper half | Exp. | Table 3.2.12 | Conf. | Yes | .93 | .20 |
| Specific Effects of Situation Perception (Level 2) on Behavior, after controlling for Happiness (on Level 1 and Level 2) | Table 3.2.6, lower half | Exp. | Table 3.2.13 | Conf. | Yes | .61 | .50 |
| Specific Effects of Situation Perception (Level 1) on Behavior, after controlling for Happiness (on Level 1 and Level 2) | Table 3.2.7, lower half | Exp. | Table 3.2.12 | Conf. | Yes | .93 | .38 |
| Specific Effects of Situation Perception (Level 2) on Behavior, after controlling for Happiness, PA, and NA (on Level 1 and Level 2) | - | - | Table 3.2.13 | Conf. | Yes | .36 | .29 |
| Specific Effects of Situation Perception (Level 1) on Behavior, after controlling for Happiness, PA, and NA (on Level 1 and Level 2) | - | - | Table 3.2.12 | Conf. | Yes | .81 | .52 |
| *Predicting Behavior with Situation Perception and Affect* | | | | Conf. | | | |
| Predicting Behavior with Personality Traits and Situation Perception | *Sherman 2015, Table 3.2.5* | - | Table 3.2.8 | Conf. | Yes | .87 | .50 |
| Predicting Behavior with Personality Traits and Situation Perception and Happiness | Table 3.2.8 | Exp. | Table 3.2.8 | Conf. | No | .83 | .40 |
| Predicting Behavior with Personality Traits and Situation Perception and Happiness, PA, and NA | - | - | Table 3.2.8 | Conf. | No | .83 | .33 |

*Note.* If Type = Conf. (Confirmatory), all predicted effects are evaluated on a one-tailed test (one-tailed 95% confidence interval, critical *t*-value of 1.64), whereas assumed null-effects are evaluated on a two-tailed 90% bootstrapped confidence interval and a critical *t*-value of 1.64. If Type = Exp. (Exploratory), all effects are evaluated on a two-tailed 95% bootstrapped confidence interval and a *t*-value of 1.96. Tables in *italics* can be found in Sherman et al., 2015.

PR = Indicates if the predicted effects were pre-registered or not; Sens. = Sensitivity; Spec. = Specificity. If Sensitivity and Specificity are not indicated (-), it was not possible or sensible to compute these values. Sherman (2015) = the original study from Sherman et al. (2015).

## Study 1

Study 1 is a re-analysis of the data provided by Sherman and colleagues (2015). The authors assessed the HEXACO personality traits (Ashton & Lee, 2007, 2009; Lee & Ashton, 2008) and their state expressions (i.e., behavior), subjective happiness at the trait and state level (Lyubomirsky & Lepper, 1999), and the Situation Eight DIAMONDS in daily life (Rauthmann & Sherman, 2016c). The authors reported logically coherent predictions of daily behavior and happiness with personality traits and situation perception (see Table 5, Sherman et al., 2015, p. 881). The authors *only* tested those predictions in their analysis. As a result, the degree to which other relationships may exist among personality traits, situation characteristics, and daily experiences is unknown. The authors concluded that "[u]ltimately, both personality traits and experienced situation characteristics appear to independently predict behavior" (Sherman et al., 2015, p. 886). We re-analyzed the data to examine which role affect/happiness played in relations between traits, states, and situation perceptions.

### Research Questions

We examined two main research questions in an exploratory manner. First, are the exclusive links between situation perception and behavior in fact exclusive (see Table 3.2.1) or are there other, so far not considered substantial links? As noted earlier (Table 3.2.1), the perception of specific characteristics of the situation should only predict certain behaviors and not all kinds of behaviors. Second, can situation perception explain variance in behavior after controlling for affect? This second analysis was conducted for all combinations of situation perception and affect, but also for the specific models examined by Sherman and colleagues (2015) that predicted behavior with multiple situation perception dimensions at the same time (see Sherman et al., Table 5, p. 881). We did not have any specific a priori hypotheses about the magnitude or direction of the examined effects. Studies that examined the interplay of affect and behavior (Wilson et al., 2017; Wilt et al., 2017; Wilt & Revelle, 2017) were not published at the time we assembled and examined these questions.

### Method

**Participants.** Data from 209 participants were analyzed. Participants were undergraduate students from Florida Atlantic University who took part for course credit. The sample size for the original analyses were based on the assumption that 200 participants suffice to get a 95% confidence interval (CI) on the $z$-scale smaller than |.15|. The mean age of the 209 participants was 18.61 ($SD$ = 1.78), and 136 participants were male and 73 female.

**Procedure.** The study was conducted in two different phases. First, participants were invited to the laboratory and completed several personality assessments (see *Trait measures*),

including self-reports on traits and an interview. Starting one day after the initial assessment, participants received eight text messages between 9 a.m. and 11 p.m. for the next seven days. It was made sure that no messages were sent less than one hour apart. Each text message included an invitation to an online survey assessing happiness, the Situation Eight DIAMONDS, and a self-rating of current HEXACO behavior (plus authenticity and self-esteem, which are not analyzed in the current study). All measures that were used and re-analyzed in the current study are displayed in Table 3.2.3.[23]

**Trait measures.** Several measures were used to assess characteristics of the participants.

***Personality Traits.*** Participants rated themselves on the HEXACO-60 (Ashton & Lee, 2009) using a 5-point Likert-type scale (1 = "strongly agree", 5 = "strongly agree"). All items of each scale were averaged to form the scale-score for each dimension. Sample items and descriptive statistics are reported in Table 3.2.3.

***Subjective happiness.*** Participants rated their general happiness on the Subjective Happiness Scale (SHS, Lyubomirsky & Lepper, 1999) using a seven-point Likert-type scale (e.g., "In general, I consider myself: 1 = "not a very happy person" to 7 = "a very happy person"). The four items are averaged to form the subjective happiness score, after reverse coding the fourth item. Sample items and descriptive statistics for the measures are displayed in Table 3.2.3.

[23] Sherman and colleagues also assessed further measures not used in the current study. A full list of the other measures used is presented by Sherman and colleagues (2015, p. 876). However, none of the other measures were analyzed or examined for the present study.

Table 3.2.3
*Descriptive Statistics of Scales and Items at Person level (trait, mean state) in Study 1 and Study 2*

| | | Study 1 | | | | | Study 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Variables | Sample Item/Adjective | *n* | *M* | *SD* | range | α | *n* | *M* | *SD* | range | α |
| **Trait Variables** | | | | | | | | | | | |
| **Personality** | | | | | | | | | | | |
| Honesty/Humility | *Having a lot of money is not especially important to me.* | 209 | 3.33 | 0.55 | 2.00 - 5.00 | .63 | 274 | 4.11 | 0.76 | 1.00 - 5.90 | .70 |
| Emotionality | *I sometimes can't help worrying about little things.* | 209 | 3.27 | 0.67 | 1.30 - 5.00 | .76 | 274 | 3.86 | 0.84 | 1.10 - 5.80 | .79 |
| Extraversion | *In social situations, I'm usually the one who makes the first move.* | 209 | 3.57 | 0.62 | 1.60 - 4.80 | .79 | 274 | 3.80 | 0.85 | 1.30 - 5.60 | .80 |
| Agreeableness | *Most people tend to get angry more quickly than I do* | 209 | 3.31 | 0.63 | 1.70 - 4.70 | .75 | 274 | 3.75 | 0.74 | 1.60 - 5.40 | .75 |
| Conscientiousness | *I often push myself very hard when trying to achieve a goal.* | 209 | 3.59 | 0.57 | 1.60 - 4.90 | .75 | 274 | 4.15 | 0.83 | 1.80 - 5.90 | .83 |
| Openness | *People have often told me that I have a good imagination.* | 209 | 3.20 | 0.66 | 1.30 - 4.60 | .74 | 274 | 4.42 | 0.77 | 1.90 - 6.00 | .73 |
| **Happiness** | | | | | | | | | | | |
| Subjective Happiness | *In general, I consider myself: not a happy person – a very happy person.* | 209 | 5.30 | 1.20 | 1.75 – 7.00 | .82 | 274 | 5.15 | 1.37 | 1.75 - 8.00 | .86 |
| **Mean State Variables** | | | | | | | | | | | |
| **Behavior** | | | | | | | | | | | |
| Honesty/Humility | *humble, honest—arrogant, dishonest* | 210 | 5.66 | 1.07 | 5.90 - 5.75 | - | 244 | 6.15 | 0.79 | 4.00 - 8.00 | - |
| Emotionality | *nervous, emotional— calm, unemotional* | 210 | 3.51 | 1.17 | 3.51 - 3.49 | - | 245 | 4.07 | 1.08 | 1.33 - 8.00 | - |
| Sociability | *outgoing, sociable—reserved, quiet* | 210 | 4.70 | 1.13 | 4.64 - 4.69 | - | 241 | 5.2 | 0.88 | 2.00 - 7.55 | - |
| Dominance | *dominant, assertive—submissive, unassertive* | 210 | 4.44 | 1.08 | 4.24 - 4.39 | - | 244 | 5.13 | 0.78 | 2.50 - 7.33 | - |
| Agreeableness | *warm, agreeable— cold, quarrelsome* | 210 | 5.39 | 1.04 | 5.41 - 5.42 | - | 243 | 6.16 | 0.80 | 3.33 - 8.00 | - |
| Conscientiousness | *organized, hardworking— disorganized, lazy* | 210 | 4.85 | 1.05 | 4.71 - 4.84 | - | 244 | 5.16 | 0.99 | 1.67 - 7.46 | - |
| Openness | *intelligent, creative— unintelligent, uncreative* | 210 | 5.12 | 1.04 | 5.06 - 5.12 | - | 244 | 5.34 | 0.91 | 2.60 - 8.00 | - |
| **Situational Perception** | | | | | | | | | | | |
| Duty | *Work has to be done.* | 210 | 4.19 | 1.16 | 4.23 - 4.19 | - | 262 | 4.49 | 1.40 | 1.00 - 8.00 | - |
| Intellect | *Deep thinking is required.* | 210 | 3.35 | 1.08 | 3.28 - 3.29 | - | 254 | 3.80 | 1.25 | 1.10 - 8.00 | - |
| Adversity | *Someone is being threatened, blamed, or criticized.* | 210 | 1.69 | 0.84 | 1.38 - 1.53 | - | 264 | 1.59 | 0.69 | 1.00 - 5.58 | - |
| Mating | *Potential romantic partners are present.* | 210 | 2.53 | 1.21 | 2.33 - 2.39 | - | 261 | 2.55 | 1.43 | 1.00 - 8.00 | - |
| pOsitivity | *Situation is enjoyable.* | 210 | 4.44 | 1.01 | 4.30 - 4.45 | - | 256 | 4.48 | 1.03 | 1.00 - 6.80 | - |
| Negativity | *Situation includes negative feelings (e.g., stress, anxiety, guilt).* | 210 | 2.46 | 1.02 | 2.35 - 2.38 | - | 261 | 3.20 | 1.30 | 1.00 - 8.00 | - |
| Deception | *Someone is being deceived.* | 210 | 1.68 | 0.82 | 1.36 - 1.52 | - | 252 | 1.40 | 0.67 | 1.00 - 6.00 | |
| Sociality | *Social interaction is possible or required.* | 210 | 4.04 | 1.04 | 4.07 - 4.06 | - | 259 | 4.21 | 1.28 | 1.00 - 7.79 | - |
| **Happiness** | | | | | | | | | | | |
| Subjective Happiness | *happy, positive—sad, negative* | 210 | 5.34 | 1.11 | 5.44 - 5.40 | - | 246 | 5.61 | 1.02 | 1.88 - 8.00 | - |
| **Affect** | | | | | | | | | | | |
| Positive Affect | *lustig [funny] - Study 2 only* | | - | - | - | - | 270 | 4.16 | 0.98 | 1.71 - 6.94 | .78 |
| Negative Affect | *gereizt [irrtated] - Study 2 only* | | - | - | - | - | 270 | 2.39 | 0.94 | 1.00 - 7.00 | .83 |

*Note.* Results from Study 1 are from Sherman and colleagues (2015). α = Cronbach's alpha for the trait-scales and nested alpha (Nezlek, 2017) for the state measures of affect. Alpha could not be computed for scales that were assessed with one item only during state assessment.

**State measures.** During the experience sampling phase, Sherman and colleagues (2015) assessed situation perception, behavior (trait-expressions, personality states), and happiness.

***Situation perceptions.*** Participants rated their current situation on eight DIAMONDS items from the S8-I (Rauthmann & Sherman, 2016c) using a 7-point Likert-type scale asking how characteristic an item is for a given situation (1 = "extremely uncharacteristic", 7 = "extremely characteristic"). Items were randomized each time the survey was taken. Items and descriptive statistics are displayed in Table 3.2.4.

***Behavior.*** For behavior or state-expressions, participants self-reported their current in situ behavior using a 7-point Likert-type bipolar-rating scale (for end-points, see Table 3.2.4). The construction of these items was based on items presented by Fleeson (2007) and Denissen, Geenen, Selfhout, and van Aken (2008). Altogether, 10 items were presented at each occasion. Seven of these items assessed behavior associated with each of the HEXACO personality dimensions. However, eXtraversion was assessed with two items, one assessing its dominance component and one assessing its sociality component. The three remaining items assessed self-esteem, authenticity, and happiness. Self-esteem and authenticity were not examined in the current study. Behavioral rating scales and their descriptive statistics are presented in Table 3.2.4.

***Happiness.*** For subjective happiness, participants rated their current in situ mood on a 7-point Likert-type- bipolar rating scale. The item and its descriptive statistics are presented in Table 3.2.4.

**Processing of experience sampling data.** The processing of experience sampling data was described by Sherman and colleagues (2015, p. 878).[24] Altogether, the participants completed 9,753 reports ($M$ = 46.44, $SD$ = 9.61). All reports that were not completed within one hour after the invitation was sent were removed. Thus, 8,318 reports could be analyzed in the present study.

**Software.** We used R (R Core Team, 2016) and RStudio (RStudio Team, 2015) for all data analyses. Additionally, we used the following packages: *broom* (Robinson, 2016), *lme4* (Bates, Maechler, Bolker, & Walker, 2016), *multicon* (Sherman, 2015), *MuMIn* (Bartoń, 2016), *psych* (Revelle, 2014), *purrr* (Wickham, 2016a), and *piecewiseSEM* (Lefcheck, 2016).

[24] Sherman and colleagues (2015) shared the R code for the preprocessing and analysis of the data. The code was adapted and extended for the current study.

Table 3.2.4

*Results from Random-Intercept Only Models for Behavior, Situation Perception, Affect and Happiness in Study 1 and Study 2*

| | Study 1 | | | | | Study 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Scales | $\tau_{00}$ | $\sigma$ | ICC | Intercept | *n* | $\tau_{00}$ | $\sigma$ | ICC | intercept | *n* |
| **Behavior** | | | | | | | | | | |
| Honesty/Humility | 1.08 | 1.27 | .46 | 5.67 | 8295 | 0.62 | 1.29 | .32 | 6.16 | 7677 |
| Emotionality | 1.27 | 2.65 | .32 | 3.51 | 8264 | 0.87 | 2.81 | .24 | 4.05 | 7693 |
| eXtraversion | | | | | | | | | | |
| Sociability | 1.14 | 2.76 | .29 | 4.71 | 8284 | 0.56 | 2.59 | .18 | 5.25 | 7669 |
| Dominance | 1.12 | 1.80 | .38 | 4.44 | 8289 | 0.53 | 1.45 | .27 | 5.15 | 7671 |
| Agreeableness | 1.01 | 1.69 | .37 | 5.39 | 8281 | 0.53 | 1.79 | .23 | 6.16 | 7668 |
| Conscientiousness | 1.02 | 2.35 | .30 | 4.87 | 8296 | 0.71 | 2.70 | .21 | 5.19 | 7674 |
| Openness | 1.00 | 1.63 | .38 | 5.13 | 8286 | 0.68 | 1.98 | .26 | 5.36 | 7674 |
| **Situational Perception** | | | | | | | | | | |
| Duty | 1.23 | 4.12 | .23 | 4.19 | 8290 | 1.39 | 5.68 | .20 | 4.45 | 7711 |
| Intellect | 1.02 | 3.79 | .21 | 3.34 | 8286 | 1.29 | 4.41 | .23 | 3.82 | 7699 |
| Adversity | 0.68 | 1.18 | .36 | 1.69 | 8284 | 0.34 | 1.48 | .19 | 1.57 | 7714 |
| Mating | 1.29 | 3.19 | .29 | 2.51 | 8302 | 1.67 | 4.53 | .27 | 2.55 | 7707 |
| pOsitivity | 0.92 | 3.30 | .22 | 4.43 | 8285 | 0.74 | 2.98 | .20 | 4.55 | 7705 |
| Negativity | 0.96 | 2.46 | .28 | 2.45 | 8298 | 1.30 | 3.16 | .29 | 3.13 | 7710 |
| Deception | 0.63 | 1.09 | .37 | 1.68 | 8295 | 0.33 | 0.86 | .28 | 1.37 | 7696 |
| Sociality | 0.90 | 4.43 | .17 | 4.03 | 8297 | 1.10 | 5.90 | .16 | 4.26 | 7709 |
| **Happiness** | | | | | | | | | | |
| Subjective Happiness | 1.10 | 1.97 | .36 | 5.35 | 8286 | 0.73 | 2.29 | .24 | 5.65 | 7689 |
| **Affect** | | | | | | | | | | |
| Positive Affect | - | - | - | | - | 0.81 | 2.17 | .27 | 4.18 | 7720 |
| Negative Affect | - | - | - | | - | 0.65 | 1.70 | .28 | 2.34 | 7720 |

*Note.* *N* (Study 1) = 210, *N* (Study 2) = 241 to 270. Results from Study 1 are taken from Sherman et al., (2015). $\tau_{00}$ = Variance between intercepts (i.e., between-person variance in random intercept only model); $\sigma$ = Variance around intercepts (within-person variance). ICC = intra-class correlation, proportion of variance between persons per total variance (i.e., the higher the value, the more is the variance attributable to the person), Intercept = fixed effect for intercept in random-intercept only model, corresponds to the across-person mean).

**Data Analysis**

Sherman and colleagues (2015) addressed four different questions in their study, and these analyses and results (with the exception of the descriptive statistics) are not repeated here. For the current analyses, all items that belonged to one scale were first recoded if necessary and then averaged to form the score for each domain. Similar to the data preparation performed by Sherman and colleagues, the means for variables at the person level (level 2) were grand mean centered. Predictor variables at the situation level (level 1, repeated assessment of situation perception and happiness) were group-mean centered (i.e., at the person level). The outcome variable, behavior, was not transformed. These procedures are in line with general recommendations regarding centering of variables in multi-level models (Enders & Tofighi, 2007).

To test our two main research questions, we performed several multilevel analyses (Hox, 2010; Snijders & Bosker, 1999). For all analyses reported here, the person is located at level 2, and the situations are nested in persons, thus located at level 1. First, we addressed the proposed specific links of situation perception and behavior (Table 3.2.1) by examining all 48 possible combinations of situation perception and behavior (8 DIAMONDS by 6 HEXACO). For example, instead of predicting conscientious behavior just from the perception of Duty in a situation, we also predicted it in seven other models from the remaining DIAMONDS dimensions. Additionally, the corresponding personality trait assessed with the HEXACO-60 (e.g., trait conscientiousness for the prediction of conscientious behavior) and the mean situation perception across all situations sampled (e.g., the mean of perceived Duty across all occasions) were included as predictors at level 2.

Second, in a new set of analyses, we examined the additional role of happiness as a predictor of behavior. We therefore extended the models described previously and included subjective happiness (assessed at level 2 with the subjective happiness scale) and subjective happiness during each situation (included at level 1 as a single item).

We ran four different models for each of the 48 combinations between situation characteristics and behavior, yielding 192 multilevel models analyzed in total. First, behavior was predicted with personality and situation perception at level 1 and level 2, with random slopes for situation perception (Model 1). Note that situation perception at level 2 can be understood as the person's general experience of situations (Horstmann & Ziegler, 2016; Ziegler & Horstmann, 2015). Second, subjective happiness was only included at level 2 and not at level 1 (Model 2). Third, subjective happiness was only included at level 1 and not at level 2 (Model 3). Lastly, subjective happiness was included both at level 1 (random slopes) and level 2 (Model 4). Table 3.2.5 gives an

overview of the models analyzed in Study 1 and Study 2. We included subjective happiness at level 1 and level 2 separately to examine their separate effects on behavior.

Finally, we also examined to what extent the affect and several situation characteristics combined predicted behavior simultaneously. We therefore examined if the specific effects of several situation characteristics presented by Sherman and colleagues (2015) would remain significant and sizable after the inclusion of happiness on level 1 and level 2. This analysis essentially tests if each situation characteristic explains unique shares of variance that cannot be explained by happiness.

Throughout Study 1, an effect was considered significant when the corresponding absolute *t*-value was larger than 1.96 (Wald $z$-test; Wald, 1943) and the bootstrapped 95% confidence interval excluded zero. Even though several models were examined, no correction for multiple testing was applied due to the exploratory nature of Study 1. This needs to be kept in mind when interpreting the results. Additionally, we computed marginal ($R_m$) and conditional ($R_c$) multiple $R$s for all models (Nakagawa & Schielzeth, 2013; Ozer, 1985). These values range from 0 to 1, and a larger value indicates better model fit. In the case of $R_m$, they can be understood as the model fit for the fixed effects only, and $R_c$ can be interpreted as the overall fit of the model.

Table 3.2.5

*Simple Notation of Multi-Level Models Presented in Study 1 and Study 2*

| Research Questions | Model | Study 1 | Study 2 |
|---|---|---|---|
| - Can behavior be predicted with situational perception and personality?<br>- Are the links between situational perception and behavior exclusive?<br>- Does personality predict behavior? | $b_{L1} = P_{L2} + S_{L2} + s_{L1}$ | Model 1 | Model 1 |
| - Are the links between situational perception and behavior exclusive after controlling for happiness at level 2?<br>- Does trait happiness predict behavior? | $b_{L1} = P_{L2} + S_{L2} + s_{L1} + SHS_{L2}$ | Model 2 | – |
| - Are the links between situational perception and behavior exclusive after controlling for happiness at level 1?<br>- Does state happiness predict behavior? | $b_{L1} = P_{L2} + S_{L2} + s_{L1} + h_{L1}$ | Model 3 | – |
| - Are the links between situational perception and behavior exclusive after controlling for happiness at level 1and level 2? | $b_{L1} = P_{L2} + S_{L2} + s_{L1} + SHS_{L2} + h_{L1}$ | Model 4 | Model 2 |
| - Are the links between situational perception and behavior exclusive after controlling for happiness, PA, and NA at level 1 and level 2?<br>- Do PA and NA, on person and situation level, predict behavior? | $b_{L1} = P_{L2} + S_{L2} + s_{L1} + SHS_{L2} + h_{L1} + NA_{L2} + PA_{L2} + na_{L1} + na_{L1}$ | – | Model 3 |

*Note.* $b$ = behavior; $P$ = personality trait; $S$ = mean situational perception, $s$ = situation perception; $SHS$ = trait subjective happiness; $h$ = state happiness; $NA$ = mean negative affect, $PA$ = mean positive affect, $na$ = state negative affect, $pa$ = state positive affect; $L1$ = variable included at level 1, the state, occasion, or situation level; $L2$ = predictor included at level 2, the person level. A detailed mathematical description of the multilevel-models, following Snijders' and Bosker's (1999) notation, can be found in the OSM C. A dash indicates that the model was not analyzed. We assumed random intercepts and random slopes for all models tested.

**Results**

Similar to the results reported by Sherman and colleagues (2015), personality predicted self-reported behavior in all models, regardless of the covariates included. These effects were already reported by Sherman and colleagues (see Sherman et al., 2015, Table, 5, p. 881) and will therefore not be reported in detail. These results can be found in the online supplemental materials (OSM A).

**Specific effects of situation perception and behavior.**[25] We first predicted behavior with situation perception, both at level 1 and level 2 (Model 1). Table 3.2.6 presents the results for the level 2 effects of mean situation perception. As can be seen, behavior was predicted mainly from the average perception of Intellect, Adversity, pOsitivity, Negativity, and Deception, and Sociality (upper half, Table 3.2.6, bold effects). The effects of situation perception at level 1 on behavior are presented in the upper half of Table 3.2.7 (Model 1). With only five exceptions, situation perceptions in terms of the DIAMONDS significantly predicted behavior in terms of HEXACO states in all models examined, that is, in 45 out of 48 combinations.

$R_m$s ranged from .12 to .44, with an average of .23 across all 48 combinations (excluding the results for dominance). $R_c$s ranged from .54 to .71, with an average of .62. These results are displayed in the OSM A. Thus, situation perception, both *in* the situation as well as measured *across* several situations, seemed to have explained substantial amounts of variance in behavior. These results replicate central findings from Sherman and colleagues (2015) as well as Rauthmann, Sherman, and colleagues (2016). However, they also reveal that situation characteristics are not exclusively tied to certain behaviors but rather exert a general effect on behavior.

[25] Sherman and colleagues (2015) initially analyzed two facets of extraversion, sociability and dominance. Throughout Study 1 and Study 2, we focused solely on sociability. Therefore, results of sensitivity and specificity (see below) do not include the results from dominance. Further, the results and hypotheses for dominance were not pre-registered (with the exception of the direct replications).

Table 3.2.6

*Level 2 Fixed Effects of Situation Perception on Behavior in Study 1*

| Sit.-Perc. | Duty | | Intellect | | Adversity | | Mating | | pOsitivity | | Negativity | | Deception | | Sociality | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Behavior | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* |
| *Model 1: Predicting Behavior from Situation Perception at level 1 and level 2* | | | | | | | | | | | | | | | | |
| H | -0.11 [-0.25; 0.03] | -1.51 | -0.53 [-0.69; -0.37] | **-6.21** | -0.11 [-0.23; 0.01] | -1.80 | 0.26 [0.11; 0.41] | **3.58** | -0.44 [-0.57; -0.32] | **-6.62** | -0.60 [-0.76; -0.44] | **-7.12** | 0.09 [-0.06; 0.23] | 1.20 | -0.11 [-0.25; 0.03] | -1.51 |
| E | 0.15 [0.03; 0.31] | **2.07** | 0.36 [0.19; 0.54] | **4.05** | 0.04 [-0.09; 0.19] | 0.55 | -0.41 [-0.57; -0.26] | **-5.16** | 0.30 [0.15; 0.44] | **4.16** | 0.30 [0.11; 0.47] | **3.16** | -0.04 [-0.21; 0.13] | -0.46 | 0.15 [0.03; 0.31] | **2.07** |
| X | -0.19 [-0.32; -0.06] | **-2.75** | -0.21 [-0.38; -0.05] | **-2.53** | 0.00 [-0.13; 0.12] | -0.06 | 0.24 [0.08; 0.39] | **3.15** | -0.27 [-0.41; -0.12] | **-3.67** | -0.24 [-0.4; -0.07] | **-2.81** | 0.31 [0.16; 0.46] | 4.10 | -0.19 [-0.32; -0.06] | **-2.75** |
| A | -0.2 [-0.33; -0.07] | **-2.98** | -0.49 [-0.62; -0.34] | **-6.43** | -0.10 [-0.21; 0.01] | -1.68 | 0.42 [0.28; 0.55] | **5.93** | -0.54 [-0.66; -0.41] | **-9.12** | -0.57 [-0.73; -0.42] | **-7.23** | 0.15 [0.01; 0.29] | 2.02 | -0.20 [-0.33; -0.07] | **-2.98** |
| C | -0.01 [-0.14; 0.12] | -0.18 | -0.24 [-0.4; -0.07] | **-2.82** | -0.02 [-0.13; 0.1] | -0.36 | 0.13 [0.00; 0.26] | 1.75 | -0.39 [-0.52; -0.25] | **-5.90** | -0.25 [-0.42; -0.09] | **-2.93** | 0.09 [-0.05; 0.24] | 1.24 | -0.01 [-0.14; 0.12] | -0.18 |
| O | -0.07 [-0.2; 0.05] | -1.08 | -0.34 [-0.51; -0.19] | **-4.11** | -0.05 [-0.18; 0.06] | -0.86 | 0.22 [0.08; 0.35] | **2.91** | -0.46 [-0.6; -0.33] | **-7.19** | -0.38 [-0.55; -0.22] | **-4.44** | 0.05 [-0.1; 0.21] | 0.70 | -0.07 [-0.2; 0.05] | -1.08 |
| D* | -0.02 [-0.16; 0.13] | -0.3 | -0.2 [-0.37; -0.03] | **-2.32** | 0.00 [-0.13; 0.13] | 0.04 | -0.14 [-0.27; 0.01] | -1.76 | -0.19 [-0.33; -0.04] | **-2.55** | -0.21 [-0.4; -0.03] | **-2.37** | 0.01 [-0.15; 0.17] | 0.07 | -0.02 [-0.16; 0.13] | -0.30 |
| *Model 4: Predicting Behavior from Situation Perception at level 1 and level 2, and Happiness at level 1 and level 2* | | | | | | | | | | | | | | | | |
| H | -0.12 [-0.25; 0.01] | -1.81 | -0.51 [-0.68; -0.37] | **-6.17** | -0.13 [-0.26; -0.02] | **-2.26** | 0.21 [0.06; 0.36] | **2.91** | -0.41 [-0.55; -0.28] | **-5.98** | -0.58 [-0.74; -0.41] | **-7.00** | 0.07 [-0.06; 0.21] | 0.91 | -0.12 [-0.25; 0.01] | -1.81 |
| E | 0.17 [0.03; 0.29] | **2.42** | 0.29 [0.13; 0.45] | **3.37** | 0.05 [-0.09; 0.17] | 0.73 | -0.4 [-0.54; -0.24] | **-5.38** | 0.33 [0.19; 0.47] | **4.58** | 0.22 [0.01; 0.38] | **2.41** | 0.00 [-0.15; 0.15] | 0.01 | 0.17 [0.03; 0.29] | **2.42** |
| X | -0.14 [-0.26; -0.01] | **-2.22** | -0.14 [-0.31; 0.01] | -1.83 | -0.01 [-0.12; 0.12] | -0.16 | 0.22 [0.07; 0.38] | **3.09** | -0.22 [-0.36; -0.08] | **-3.20** | -0.19 [-0.35; -0.04] | **-2.41** | 0.23 [0.08; 0.35] | **3.14** | -0.14 [-0.26; -0.01] | **-2.22** |
| A | -0.19 [-0.31; -0.07] | **-3.11** | -0.48 [-0.63; -0.33] | **-6.82** | -0.14 [-0.27; -0.02] | **-2.49** | 0.35 [0.22; 0.48] | **5.12** | -0.49 [-0.62; -0.37] | **-8.26** | -0.53 [-0.68; -0.38] | **-7.16** | 0.10 [-0.04; 0.26] | 1.47 | -0.19 [-0.31; -0.07] | **-3.11** |
| C | 0.03 [-0.09; 0.14] | 0.53 | -0.21 [-0.37; -0.05] | **-2.64** | -0.07 [-0.18; 0.03] | -1.37 | 0.08 [-0.05; 0.23] | 1.26 | -0.33 [-0.46; -0.21] | **-5.24** | -0.20 [-0.36; -0.05] | **-2.57** | 0.07 [-0.06; 0.22] | 1.08 | 0.03 [-0.09; 0.14] | 0.53 |
| O | -0.03 [-0.16; 0.08] | -0.53 | -0.28 [-0.43; -0.13] | **-3.56** | -0.11 [-0.22; -0.01] | **-2.11** | 0.16 [0.01; 0.28] | **2.32** | -0.38 [-0.52; -0.27] | **-6.07** | -0.28 [-0.46; -0.13] | **-3.50** | -0.01 [-0.16; 0.12] | -0.10 | -0.03 [-0.16; 0.08] | -0.53 |
| D* | -0.03 [-0.15; 0.09] | -0.46 | -0.18 [-0.34; -0.02] | **-2.30** | -0.04 [-0.14; 0.06] | -0.74 | -0.08 [-0.2; 0.05] | -1.11 | -0.18 [-0.32; -0.05] | -2.69 | -0.21 [-0.38; -0.06] | **-2.77** | 0.00 [-0.13; 0.16] | 0.04 | -0.03 [-0.15; 0.09] | -0.46 |

*Note.* Sit.-Perc. = situation perception (predictor level 2): *b* = estimate (unstandardized) effect; LL = lower limit of 95% CI, UL = upper limit of 95% CI; *t* = *t*-value of fixed effect. Predicted behaviors: H = Honesty/Humility; E = Emotionality; X = eXtraversion; A = Agreeableness; C = Conscientiousness; O = Openness; Bold effects are significant. *The effects for Dominance were not included in the computation of sensitivity and specificity and the hypotheses were not pre-registered for Study 2.

Table 3.2.7

*Level 1 Fixed Effects of Situation Perception on Behavior in Study 1*

| Sit.-Perc. | Duty | | Intellect | | Adversity | | Mating | | pOsitivity | | Negativity | | Deception | | Sociality | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Behavior | *b* [LL, UL] | *t* | b [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* |
| *Model 1: Predicting Behavior from Situation Perception at level 1 and level 2* | | | | | | | | | | | | | | | | |
| H | -0.03 [-0.04; -0.01] | **-3.14** | -0.02 [-0.04; -0.01] | **-2.50** | -0.10 [-0.15; -0.05] | **-4.34** | 0.01 [0.00; 0.03] | 1.36 | 0.09 [0.06; 0.11] | **6.98** | -0.11 [-0.14; -0.08] | **-7.51** | *-0.16 [-0.21; -0.11]* | ***-6.67*** | 0.02 [0.01; 0.03] | **2.30** |
| E | 0.05 [0.02; 0.07] | **3.23** | 0.09 [0.06; 0.12] | **5.85** | *0.14 [0.09; 0.20]* | ***5.11*** | 0.00 [-0.04; 0.03] | -0.24 | -0.21 [-0.25; -0.17] | **-11.67** | *0.23 [0.18; 0.28]* | ***9.64*** | 0.16 [0.10; 0.21] | **5.83** | -0.04 [-0.07; -0.02] | **-3.51** |
| X | -0.05 [-0.07; -0.02] | **-3.75** | -0.06 [-0.09; -0.04] | **-4.40** | -0.08 [-0.14; -0.02] | **-2.63** | 0.12 [0.10; 0.15] | **8.35** | *0.22 [0.18; 0.26]* | ***11.24*** | -0.23 [-0.27; -0.18] | **-10.66** | *-0.08 [-0.15; -0.03]* | ***-2.96*** | *0.25 [0.21; 0.28]* | ***14.02*** |
| A | -0.06 [-0.08; -0.04] | **-6.60** | -0.08 [-0.10; -0.06] | **-7.65** | -0.20 [-0.26; -0.15] | **-7.61** | *0.04 [0.02; 0.06]* | ***3.93*** | *0.20 [0.16; 0.23]* | ***11.43*** | -0.27 [-0.3; -0.23] | **-15.19** | *-0.20 [-0.25; -0.14]* | ***-7.33*** | *0.07 [0.05; 0.09]* | ***6.75*** |
| C | *0.14 [0.11; 0.17]* | ***8.58*** | 0.13 [0.10; 0.16] | **7.70** | -0.07 [-0.11; -0.02] | **-2.99** | 0.00 [-0.03; 0.02] | -0.19 | -0.03 [-0.07; 0.00] | -1.91 | 0.00 [-0.04; 0.03] | -0.13 | -0.09 [-0.13; -0.03] | **-3.43** | 0.04 [0.02; 0.07] | **3.93** |
| O | 0.04 [0.02; 0.07] | **3.84** | *0.04 [0.01; 0.06]* | ***2.96*** | -0.10 [-0.14; -0.05] | **-4.59** | 0.03 [0.01; 0.04] | **2.90** | 0.07 [0.04; 0.10] | **4.98** | -0.12 [-0.16; -0.09] | **-7.14** | **-0.11 [-0.16; -0.07]** | **-4.89** | 0.05 [0.03; 0.07] | **5.41** |
| D* | 0.02 [0.00; 0.04] | 1.72 | 0.01 [-0.01; 0.04] | 1.13 | 0.05 [-0.01; 0.11] | 1.82 | *0.01 [-0.01; 0.04]* | *1.29* | 0.00 [-0.03; 0.03] | 0.14 | -0.02 [-0.05; 0.01] | -1.23 | 0.01 [-0.04; 0.06] | 0.26 | *0.04 [0.01; 0.06]* | ***3.00*** |
| *Model 4: Predicting Behavior from Situational Perception at level 1 and level 2, and Happiness a level 1 and level 2* | | | | | | | | | | | | | | | | |
| H | 0.01 [-0.01; 0.02] | 0.90 | 0.01 [0.00; 0.03] | **1.96** | -0.04 [-0.08; 0.01] | -1.80 | 0.00 [-0.01; 0.02] | 0.15 | 0.01 [-0.01; 0.02] | 0.68 | -0.01 [-0.02; 0.01] | -0.68 | *-0.09 [-0.13; -0.05]* | ***-4.60*** | 0.00 [-0.02; 0.01] | -0.53 |
| E | 0.01 [-0.01; 0.04] | 1.04 | 0.05 [0.02; 0.07] | **3.75** | *0.08 [0.03; 0.13]* | ***3.28*** | 0.02 [-0.01; 0.05] | 1.07 | -0.12 [-0.15; -0.09] | **-7.68** | *0.13 [0.1; 0.17]* | ***6.74*** | 0.10 [0.05; 0.14] | **4.35** | -0.01 [-0.03; 0.01] | -0.93 |
| X | 0.00 [-0.02; 0.02] | 0.28 | 0 [-0.02; 0.02] | -0.09 | 0.03 [0.00; 0.07] | 1.89 | 0.09 [0.07; 0.12] | **7.35** | *0.09 [0.05; 0.11]* | ***5.65*** | -0.04 [-0.07; -0.01] | **-2.43** | *0.02 [-0.02; 0.06]* | *0.95* | *0.21 [0.18; 0.24]* | ***13.23*** |
| A | -0.01 [-0.02; 0.01] | -1.18 | -0.01 [-0.03; 0.00] | -1.52 | -0.1 [-0.13; -0.05] | **-4.74** | *0.01 [0; 0.03]* | *1.79* | *0.05 [0.03; 0.07]* | ***4.55*** | -0.09 [-0.12; -0.07] | **-7.29** | *-0.08 [-0.11; -0.05]* | ***-4.69*** | *0.02 [0.01; 0.03]* | ***2.88*** |
| C | *0.16 [0.13; 0.2]* | ***10.22*** | 0.16 [0.13; 0.19] | **10.12** | -0.03 [-0.07; 0.01] | -1.42 | -0.01 [-0.04; 0.01] | -0.84 | -0.11 [-0.14; -0.08] | **-6.42** | 0.08 [0.05; 0.11] | **4.61** | -0.04 [-0.08; 0.01] | -1.68 | 0.03 [0.01; 0.05] | **2.60** |
| O | 0.08 [0.05; 0.1] | **6.96** | *0.08 [0.05; 0.11]* | ***6.59*** | -0.03 [-0.07; 0] | -1.92 | 0.01 [0; 0.03] | 1.39 | -0.02 [-0.04; 0.00] | **-2.06** | -0.01 [-0.04; 0.02] | -0.65 | -0.04 [-0.08; -0.01] | **-2.30** | 0.03 [0.01; 0.04] | **4.28** |
| D* | 0.03 [0.01; 0.05] | **2.68** | 0.03 [0.01; 0.05] | **2.67** | *0.07 [0.02; 0.12]* | ***3.00*** | *0.01 [-0.01; 0.03]* | *1.07* | -0.03 [-0.05; -0.01] | **-2.48** | 0.03 [0.01; 0.06] | **2.20** | 0.03 [-0.01; 0.07] | 1.54 | 0.03 [0.01; 0.05] | **2.38** |

*Note.* Sit.-Perc. = Situation Perception (Predictor level 1): *b* = estimate (unstandardized) effect; LL = lower limit of 95% CI, UL = upper limit of 95% CI; *t* = *t*-value of fixed effect.
Predicted behaviors: H = Honesty/Humility; E = Emotionality; X = eXtraversion; A = Agreeableness; C = Conscientiousness; O = Openness;
Bold effects are significant. Effects in *italics* were predicted to be significant. No effect was significant in the opposite than the predicted direction.
*The effects for Dominance were not included in the computation of Sensitivity and Specificity and the results not pre-registered.

**Effects of happiness.** We next examined the effects of happiness on behavior. Happiness was included both at level 2 and level 1 (Model 4). Trait happiness at level 2 predicted honest/humble, agreeable, conscientious, and open behavior across all models. For emotionality, trait happiness was only a significant predictor in the case where it was included in the model with Negativity. However, this effect was only "marginally significant" ($b$ = 0.12, 95% CI = [0.00, 0.24], $t$ = 1.99), indicating that it might not replicate (Benjamin et al., 2018). Including happiness at level 1 and level 2 revealed that the fixed effect for happiness at level 1 was significant in *all of the 48 models examined*. Compared to the previous model, conditional and marginal $R$s substantially increased. $R_m$s ranged from .25 to .60, with an average of .39. $R_c$s ranged from .70 to .80, with an average of .71. These results of these analyses can be found in the OSM A. These findings suggest that happiness as a trait, and especially as a state, explained substantial amounts of variation in behavior. Furthermore, happiness explained variance in behavior above and beyond measures of situation perception.

**Effects of happiness and situation perception.** The inclusion of happiness did not only reveal that it is a valuable predictor of behavior, but also that when including happiness at level 1 and level 2, several characteristics were no longer significant predictors of behavior. This effect occurred only for level 1 situation perception, which represents the perception of the current situation in daily life. Whereas the regression weights of mean situation perception at the person level were not altered substantially (Table 3.2.6, lower half compared to upper half), several regression weights of in situ situation perception were substantially reduced (Table 3.2.7, lower half compared to upper half). Furthermore, when comparing the significant effects of situation perception to the effects that were hypothesized by Sherman et al. (2015, see Table 3.2.1, upper third), the inclusion of happiness primarily eliminated those effects that were *not* supposed to be significant in the first place. For example, perceiving Adversity should be uniquely associated with decreased agreeable behavior and no other HEXACO behaviors. However, our first reanalysis revealed that perceiving Adversity explained variance in all HEXACO behaviors. After including happiness and thus controlling for its effect, Adversity no longer explained variance all HEXACO behaviors, but only in agreeable ($b$ = -0.10, 95% CI: [-0.13, -0.05], $t$ = -4.74) and emotional behavior ($b$ = 0.08, 95% CI = [0.03, 0.13], $t$ = 3.28).

To make the specific effect of the inclusion of happiness on in situ situation perception more apparent, we computed sensitivity and specificity indices for these results. We counted hits (i.e., confirmed hypotheses), correct rejections (i.e., hypotheses correctly rejected), false rejections (i.e., hypotheses that should be confirmed but were not), and false alarms (i.e., hypotheses that should be rejected but were not) based on Table 3.2.1. The sensitivity in the first analysis (situation

perception only) was .93, and the specificity was .26. Including happiness as a predictor reduced sensitivity to .78, but increased specificity to .50. Note that a value of .50 is similar to tossing a coin; however, this should not be taken as the baseline. Rather, .26 should serve as the baseline for the evaluation of future results as this was the accuracy in the first analyses without happiness. In other words, the theoretical predictions are more accurate and meaningful if happiness is included as a predictor. To examine if this effect was due to the inclusion of happiness at level 1 or level 2, we computed two additional models (Model 2 and Model 3, see above) and included both variables separately. These analyses revealed that the effect was due to the inclusion of happiness on level 1 and not on level 2. The results of these analyses can be found in the OSM A.

**Effect of happiness and multiple dimensions of situation perception.** Sherman and colleagues (2015) examined how several situation characteristics combined predicted in situ behavior (Sherman et al, 2015, Table 3.2.5, p. 881). For example, they predicted emotionality behavior with perceived Adversity and perceived Negativity at the same time. Based on the above results, we extended these specific models by Sherman and colleagues and added happiness as a trait and state as additional predictors. The results of these analyses are presented in Table 3.2.8 (Study 1, re-analysis). Overall, seven models were examined. First, state happiness was a significant predictor of behavior in all seven and trait happiness in five models. This is reflected in a substantial increase in explained variance. Furthermore, and most importantly, the effects of situation perception on behavior remain largely unchanged (compare column "*t* original effect Sherman et al." with *t*-values from Study 1, reanalysis, in Table 3.2.8). Even though they are somewhat reduced in magnitude after inclusion of happiness on level 1 and level 2, only two parameters that were statistically significant before including happiness were no longer significant.

Table 3.2.8

*Multi-Level Models of Interest: Predicting Behavior with Multiple Situation Dimensions, Affect, and Happiness*

| Variable | *t* Sherman et al., 2015 | Study 1, re-analysis | | | | | | PE | Study 2, replication | | | | | | Study 2, replication of reanalysis | | | | | | Study 2, extension | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *b* | LL | UL | *t* | $R_m$ | $R_c$ | | *b* | LL | UL | *t* | $R_m$ | $R_c$ | *b* | LL | UL | *t* | $R_m$ | $R_c$ | *b* | LL | UL | *t* | $R_m$ | $R_c$ |
| **Honest/Humble Behavior** | | 5.66 | 5.53 | 5.79 | 86.45 | .49 | .79 | | 6.16 | 6.09 | 6.25 | 123.48 | .25 | .61 | 6.16 | 6.09 | 6.24 | 123.46 | .30 | .65 | 6.16 | 6.07 | 6.23 | 124.05 | .33 | .68 |
| Trait Honesty | 2.30 | 0.23 | -0.01 | 0.47 | 1.88 | | | + | 0.31 | 0.21 | 0.42 | 4.72 | | | 0.30 | 0.19 | 0.42 | 4.54 | | | 0.28 | 0.18 | 0.40 | 4.19 | | |
| Mean Situation Deception | -7.10 | -0.58 | -0.74 | -0.41 | -6.98 | | | - | -0.15 | -0.27 | -0.02 | -1.91 | | | -0.15 | -0.29 | 0.00 | -1.89 | | | -0.08 | -0.23 | 0.07 | -0.91 | | |
| Situation Deception | -7.24 | -0.10 | -0.15 | -0.06 | -5.14 | | | - | -0.20 | -0.25 | -0.15 | -6.76 | | | -0.16 | -0.20 | -0.11 | -5.50 | | | -0.13 | -0.18 | -0.09 | -4.75 | | |
| Int. Honesty*Deception | -2.71 | -0.10 | -0.18 | -0.03 | -2.74 | | | - | 0.03 | -0.03 | 0.09 | 0.95 | | | 0.03 | -0.02 | 0.09 | 1.00 | | | 0.02 | -0.04 | 0.07 | 0.49 | | |
| Trait Happiness | | 0.19 | 0.08 | 0.30 | 3.44 | | | | | | | | | | 0.01 | -0.08 | 0.08 | 0.25 | | | -0.04 | -0.11 | 0.03 | -0.88 | | |
| State Happiness | | 0.31 | 0.27 | 0.35 | 15.56 | | | | | | | | | | 0.16 | 0.13 | 0.19 | 10.17 | | | 0.08 | 0.05 | 0.11 | 4.67 | | |
| Mean Positive Affect | | | | | | | | | | | | | | | | | | | | | 0.02 | -0.10 | 0.14 | 0.33 | | |
| Mean Negative Affect | | | | | | | | | | | | | | | | | | | | | -0.17 | -0.30 | -0.03 | -2.46 | | |
| State Positive Affect | | | | | | | | | | | | | | | | | | | | | -0.02 | -0.05 | 0.01 | -1.27 | | |
| State Negative Affect | | | | | | | | | | | | | | | | | | | | | -0.18 | -0.22 | -0.13 | -8.46 | | |
| **Emotionality Behavior** | | 3.50 | 3.32 | 3.67 | 41.61 | .30 | .66 | | 4.06 | 3.96 | 4.16 | 75.49 | .44 | .62 | 4.07 | 3.98 | 4.15 | 75.56 | .44 | .64 | 4.07 | 3.98 | 4.15 | 76.33 | .46 | .66 |
| Trait Emotionality | 2.16 | 0.18 | -0.06 | 0.40 | 1.52 | | | + | 0.35 | 0.25 | 0.47 | 5.25 | | | 0.35 | 0.24 | 0.47 | 5.21 | | | 0.35 | 0.23 | 0.46 | 5.27 | | |
| Mean Situation Adversity | 0.13 | -0.09 | -0.34 | 0.18 | -0.71 | | | 0 | -0.04 | -0.20 | 0.10 | -0.48 | | | -0.04 | -0.19 | 0.12 | -0.39 | | | -0.12 | -0.28 | 0.01 | -1.29 | | |
| Mean Situation Negativity | 2.59 | 0.39 | 0.17 | 0.59 | 3.45 | | | + | 0.35 | 0.27 | 0.44 | 6.99 | | | 0.37 | 0.28 | 0.46 | 6.82 | | | 0.28 | 0.19 | 0.38 | 4.72 | | |
| Situation Adversity | 2.18 | 0.05 | 0.00 | 0.09 | 2.07 | | | +, a | 0.15 | 0.12 | 0.18 | 7.89 | | | 0.14 | 0.11 | 0.18 | 7.45 | | | 0.10 | 0.07 | 0.14 | 5.39 | | |
| Situation Negativity | 8.55 | 0.12 | 0.08 | 0.16 | 5.67 | | | + | 0.32 | 0.29 | 0.35 | 17.75 | | | 0.29 | 0.26 | 0.32 | 15.67 | | | 0.30 | 0.27 | 0.33 | 15.55 | | |
| Trait Happiness | | 0.14 | 0.03 | 0.25 | 2.17 | | | | | | | | | | 0.03 | -0.06 | 0.13 | 0.72 | | | 0.04 | -0.06 | 0.14 | 0.80 | | |
| State Happiness | | -0.24 | -0.28 | -0.19 | -9.60 | | | | | | | | | | -0.06 | -0.10 | -0.01 | -2.60 | | | *removed* | | | | | |
| Mean Positive Affect | | | | | | | | | | | | | | | | | | | | | 0.05 | -0.07 | 0.18 | 0.85 | | |
| Mean Negative Affect | | | | | | | | | | | | | | | | | | | | | 0.24 | 0.08 | 0.40 | 2.81 | | |
| State Positive Affect | | | | | | | | | | | | | | | | | | | | | 0.15 | 0.10 | 0.19 | 6.26 | | |
| State Negative Affect | | | | | | | | | | | | | | | | | | | | | 0.22 | 0.17 | 0.26 | 8.89 | | |
| **Sociability Behavior** | | 4.71 | 4.56 | 4.87 | 62.76 | .48 | .75 | | 5.23 | 5.15 | 5.31 | 115.36 | .47 | .62 | 5.22 | 5.14 | 5.29 | 116.26 | .64 | .77 | 5.22 | 5.15 | 5.29 | 121.24 | .66 | .78 |
| Trait Extraversion | 5.28 | 0.4 | 0.16 | 0.69 | 2.87 | | | + | 0.42 | 0.33 | 0.51 | 7.70 | | | 0.29 | 0.16 | 0.41 | 4.15 | | | 0.27 | 0.17 | 0.38 | 3.99 | | |
| Mean Situation Sociality | 4.10 | 0.23 | 0.06 | 0.37 | 3.15 | | | + | 0.16 | 0.10 | 0.23 | 4.38 | | | 0.17 | 0.12 | 0.24 | 4.84 | | | 0.17 | 0.10 | 0.22 | 4.61 | | |
| Situation Sociality | 14.44 | 0.21 | 0.18 | 0.24 | 13.65 | | | + | 0.30 | 0.27 | 0.32 | 24.02 | | | 0.22 | 0.20 | 0.24 | 20.91 | | | 0.21 | 0.19 | 0.22 | 19.69 | | |
| Int. Extraversion*Sociality | -2.69 | -0.06 | -0.11 | -0.01 | -2.41 | | | - | 0.04 | 0.01 | 0.06 | 2.53 | | | 0.03 | 0.01 | 0.05 | 2.26 | | | 0.02 | 0.00 | 0.04 | 1.94 | | |
| Trait Happiness | | 0.12 | -0.03 | 0.24 | 1.59 | | | | | | | | | | 0.13 | 0.04 | 0.21 | 2.97 | | | 0.02 | -0.06 | 0.09 | 0.40 | | |
| State Happiness | | 0.48 | 0.43 | 0.52 | 20.81 | | | | | | | | | | 0.53 | 0.50 | 0.56 | 35.67 | | | 0.37 | 0.34 | 0.40 | 19.84 | | |
| Mean Positive Affect | | | | | | | | | | | | | | | | | | | | | 0.16 | 0.05 | 0.26 | 3.19 | | |
| Mean Negative Affect | | | | | | | | | | | | | | | | | | | | | -0.21 | -0.31 | -0.11 | -3.97 | | |
| State Positive Affect | | | | | | | | | | | | | | | | | | | | | 0.19 | 0.15 | 0.23 | 9.90 | | |
| State Negative Affect | | | | | | | | | | | | | | | | | | | | | -0.05 | -0.09 | -0.02 | -3.18 | | |

Table 3.2.8 (continued)

*Multi-Level Models of Interest: Predicting Behavior with Multiple Situation Dimensions, Affect, and Happiness*

| | *t* Sherman et al., 2015 | Study 1, re-analysis | | | | | | | Study 2, replication | | | | | | Study 2, replication of reanalysis | | | | | | Study 2, extension | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | | *b* | LL | UL | *t* | $R_m$ | $R_c$ | | *b* | LL | UL | *t* | $R_m$ | $R_c$ | *b* | LL | UL | *t* | $R_m$ | $R_c$ | *b* | LL | UL | *t* | $R_m$ | $R_c$ |
| **Agreeable Behavior** | | 5.43 | 5.31 | 5.57 | 82.36 | .60 | .81 | | 6.17 | 6.10 | 6.25 | 136.89 | .46 | .67 | 6.17 | 6.10 | 6.24 | 137.27 | .56 | .75 | 6.17 | 6.11 | 6.24 | 140.45 | .55 | .75 |
| Trait Agreeableness | 1.65 | 0.06 | -0.14 | 0.30 | 0.54 | | | 0 | 0.21 | 0.12 | 0.30 | 3.73 | | | 0.18 | 0.08 | 0.28 | 3.21 | | | 0.17 | 0.09 | 0.26 | 3.26 | | |
| Mean Situation Sociality | 0.40 | -0.01 | -0.19 | 0.16 | -0.17 | | | 0 | 0.06 | -0.01 | 0.12 | 1.47 | | | 0.04 | -0.02 | 0.10 | 1.10 | | | 0.05 | -0.02 | 0.11 | 1.39 | | |
| Mean Situation pOsitivity | 3.02 | 0.19 | 0.02 | 0.34 | 2.35 | | | + | 0.19 | 0.11 | 0.27 | 4.03 | | | 0.17 | 0.09 | 0.24 | 3.42 | | | 0.17 | 0.06 | 0.27 | 2.52 | | |
| Mean Situation Deception | -6.74 | -0.55 | -0.73 | -0.38 | -6.32 | | | - | -0.31 | -0.42 | -0.21 | -4.58 | | | -0.28 | -0.39 | -0.15 | -4.04 | | | -0.16 | -0.30 | -0.03 | -2.17 | | |
| Situation Sociality | 2.11 | | | *removed* | | | | + | 0.05 | 0.03 | 0.06 | 5.11 | | | 0.03 | 0.02 | 0.04 | 3.49 | | | 0.05 | 0.04 | 0.07 | 6.13 | | |
| Situation Positivity | 10.36 | 0.05 | 0.03 | 0.07 | 4.39 | | | + | 0.31 | 0.28 | 0.34 | 18.11 | | | 0.08 | 0.06 | 0.11 | 5.99 | | | 0.11 | 0.09 | 0.13 | 8.71 | | |
| Situation Deception | -6.29 | -0.08 | -0.12 | -0.04 | -4.00 | | | - | -0.20 | -0.25 | -0.16 | -6.79 | | | -0.13 | -0.17 | -0.08 | -4.81 | | | -0.11 | -0.15 | -0.06 | -3.78 | | |
| Trait Happiness | | 0.26 | 0.14 | 0.38 | 4.28 | | | | | | | | | | 0.07 | 0.01 | 0.14 | 2.14 | | | 0.04 | -0.03 | 0.11 | 1.20 | | |
| State Happiness | | 0.49 | 0.44 | 0.54 | 22.49 | | | | | | | | | | 0.42 | 0.38 | 0.46 | 21.26 | | | | *removed* | | | | |
| Mean Positive Affect | | | | | | | | | | | | | | | | | | | | | -0.08 | -0.20 | 0.05 | -1.21 | | |
| Mean Negative Affect | | | | | | | | | | | | | | | | | | | | | -0.23 | -0.34 | -0.12 | -4.04 | | |
| State Positive Affect | | | | | | | | | | | | | | | | | | | | | 0.12 | 0.08 | 0.15 | 6.70 | | |
| State Negative Affect | | | | | | | | | | | | | | | | | | | | | -0.34 | -0.38 | -0.30 | -14.82 | | |
| **Conscientious Behavior** | | 4.88 | 4.73 | 5.03 | 69.66 | .33 | .66 | | 5.18 | 5.09 | 5.28 | 95.10 | .37 | .60 | 5.17 | 5.09 | 5.26 | 98.88 | .46 | .67 | 5.17 | 5.08 | 5.24 | 98.67 | .46 | .68 |
| Trait Conscientiousness | 5.12 | 0.54 | 0.34 | 0.75 | 4.79 | | | + | 0.35 | 0.23 | 0.46 | 5.35 | | | 0.36 | 0.24 | 0.46 | 5.70 | | | 0.36 | 0.25 | 0.46 | 5.65 | | |
| Mean Situation Duty | -0.28 | 0.00 | -0.11 | 0.11 | -0.07 | | | 0 | 0.15 | 0.08 | 0.22 | 3.74 | | | 0.16 | 0.10 | 0.22 | 4.10 | | | 0.18 | 0.11 | 0.24 | 4.54 | | |
| Situation Duty | 8.56 | 0.16 | 0.13 | 0.19 | 10.22 | | | + | 0.25 | 0.23 | 0.27 | 17.94 | | | 0.27 | 0.24 | 0.29 | 19.84 | | | 0.27 | 0.24 | 0.29 | 19.97 | | |
| Trait Happiness | | 0.21 | 0.08 | 0.32 | 3.73 | | | | | | | | | | 0.18 | 0.11 | 0.25 | 4.89 | | | 0.11 | 0.01 | 0.20 | 2.26 | | |
| State Happiness | | 0.25 | 0.20 | 0.30 | 10.70 | | | | | | | | | | 0.30 | 0.27 | 0.33 | 18.06 | | | 0.23 | 0.19 | 0.28 | 10.50 | | |
| Mean Positive Affect | | | | | | | | | | | | | | | | | | | | | 0.09 | -0.02 | 0.21 | 1.59 | | |
| Mean Negative Affect | | | | | | | | | | | | | | | | | | | | | -0.12 | -0.24 | 0.02 | -1.83 | | |
| State Positive Affect | | | | | | | | | | | | | | | | | | | | | 0.10 | 0.06 | 0.14 | 4.57 | | |
| State Negative Affect | | | | | | | | | | | | | | | | | | | | | 0.00 | -0.04 | 0.04 | -0.02 | | |
| **Openness Behavior** | | 5.13 | 5.00 | 5.27 | 74.03 | .38 | .73 | | 5.34 | 5.25 | 5.44 | 99.12 | .27 | .60 | 5.34 | 5.26 | 5.43 | 101.97 | .47 | .72 | 5.34 | 5.26 | 5.42 | 105.53 | .49 | .73 |
| Trait Openness | 3.29 | 0.33 | 0.12 | 0.53 | 3.21 | | | + | 0.25 | 0.15 | 0.35 | 3.64 | | | 0.20 | 0.08 | 0.30 | 2.99 | | | 0.17 | 0.08 | 0.27 | 2.80 | | |
| Mean Situation Intellect | -1.08 | -0.03 | -0.16 | 0.09 | -0.53 | | | 0 | 0.15 | 0.07 | 0.22 | 3.32 | | | 0.18 | 0.11 | 0.25 | 4.21 | | | 0.20 | 0.14 | 0.28 | 4.83 | | |
| Situation Intellect | 2.90 | 0.08 | 0.06 | 0.10 | 6.59 | | | + | 0.17 | 0.14 | 0.19 | 11.44 | | | 0.19 | 0.17 | 0.21 | 14.62 | | | 0.19 | 0.17 | 0.21 | 14.50 | | |
| Trait Happiness | | 0.29 | 0.19 | 0.40 | 5.23 | | | | | | | | | | 0.17 | 0.09 | 0.24 | 4.29 | | | 0.04 | -0.05 | 0.13 | 0.79 | | |
| State Happiness | | 0.33 | 0.30 | 0.38 | 16.41 | | | | | | | | | | 0.39 | 0.36 | 0.42 | 26.17 | | | 0.30 | 0.26 | 0.33 | 15.83 | | |
| Mean Positive Affect | | | | | | | | | | | | | | | | | | | | | 0.15 | 0.04 | 0.26 | 2.57 | | |
| Mean Negative Affect | | | | | | | | | | | | | | | | | | | | | -0.24 | -0.35 | -0.12 | -3.86 | | |
| State Positive Affect | | | | | | | | | | | | | | | | | | | | | 0.14 | 0.10 | 0.18 | 7.57 | | |
| State Negative Affect | | | | | | | | | | | | | | | | | | | | | 0.00 | -0.03 | 0.03 | -0.11 | | |

Table 3.2.8 (continued)

*Multi-Level Models of Interest: Predicting Behavior with Multiple Situation Dimensions, Affect, and Happiness*

| | *t* Sherman | Study 1, re-analysis | | | | | | | Study 2, replication | | | | | | Study 2, replication of reanalysis | | | | | | Study 2, extension | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | et al. 2015 | *b* | LL | UL | *t* | $R_c$ | $R_m$ | | *b* | LL | UL | *t* | $R_c$ | $R_m$ | *b* | LL | UL | *t* | $R_c$ | $R_m$ | *b* | LL | UL | *t* | $R_c$ | $R_m$ |
| **Dominance Behavior** | | 4.43 | 4.27 | 4.59 | 55.52 | .21 | .69 | | 5.15 | 5.08 | 5.23 | 116.10 | .24 | .54 | 5.15 | 5.08 | 5.22 | 116.81 | .35 | .63 | 5.15 | 5.08 | 5.22 | 119.57 | .38 | .65 |
| Trait Extraversion | 3.22 | 0.40 | 0.15 | 0.67 | 3.00 | | | + | 0.38 | 0.30 | 0.47 | 7.47 | | | 0.31 | 0.20 | 0.44 | 4.62 | | | 0.29 | 0.18 | 0.40 | 4.26 | | |
| Mean Situation Adversity | -2.32 | -0.18 | -0.35 | -0.02 | -2.30 | | | - | -0.08 | -0.19 | 0.02 | -1.24 | | | -0.06 | -0.16 | 0.04 | -0.87 | | | -0.01 | -0.14 | 0.10 | -0.18 | | |
| Situation Adversity | 1.73 | 0.07 | 0.03 | 0.12 | 3.00 | | | 0 | -0.03 | -0.06 | 0.00 | -1.53 | | | 0.06 | 0.03 | 0.09 | 3.47 | | | 0.03 | 0.00 | 0.06 | 1.69 | | |
| Trait Happiness | | 0.00 | -0.14 | 0.15 | -0.03 | | | | | | | | | | 0.06 | -0.03 | 0.15 | 1.38 | | | -0.02 | -0.09 | 0.05 | -0.40 | | |
| State Happiness | | 0.15 | 0.10 | 0.19 | 6.13 | | | | | | | | | | 0.25 | 0.22 | 0.28 | 16.25 | | | 0.23 | 0.20 | 0.26 | 12.93 | | |
| Mean Positive Affect | | | | | | | | | | | | | | | | | | | | | 0.13 | 0.02 | 0.24 | 2.60 | | |
| Mean Negative Affect | | | | | | | | | | | | | | | | | | | | | -0.14 | -0.27 | -0.02 | -2.38 | | |
| State Positive Affect | | | | | | | | | | | | | | | | | | | | | 0.11 | 0.08 | 0.14 | 6.86 | | |
| State Negative Affect | | | | | | | | | | | | | | | | | | | | | 0.11 | 0.07 | 0.15 | 6.01 | | |

*Note.* PE = the predicted effect, based on Study 1can be found in Table 3.2.3 in the pre-registration. Note that we erroneously pre-registered a negative effect for Situation Adversity on Emotionality Behavior, whereas Sherman and colleagues reported and predicted a positive effect. Furthermore, in the pre-registration, we stated that Sherman and colleagues had found a negative effect. The expected effect therefore should have been positive, which has to be considered an error in the pre-registration. Int. = Interaction.

*t* Sherman et al., 2015: The original *t*-value reported by Sherman and colleagues 2015); Study 1, re-analysis: The results from Study 1, re-analysis including happiness at level 1 and level 2; Pred. Effect: The effects that were predicted for Study 2 based on the results and interpretation from Study 1; Study 2, replication: The results from Study 2, replication the original results from Sherman and colleagues (2015); Study 2, replication of re-analysis: The direct replication of the results presented in the column Study 1, re-analysis; Study 2, extension: Results from Study 2, predicting behavior with situation perception, positive and negative affect (both level 1 and level 2), and happiness (level 1 and level 2). Parameters: *b*: regression coefficient, unstandardized, LL and UL: Lower and Upper limit of the 90 or 95% confidence interval. Whenever an effect was predicted in one direction, the one-tailed 95% CI (= one side of the 90% CI) was used to determine significance. Otherwise, a 95% CI is reported. *t*: *t*-value of the fixed effect; *removed*: the predictor had to be removed, otherwise the models did not converge.

**Discussion**

The re-analyses of Sherman and colleagues' (2015) data revealed two important insights. First, happiness, both as a trait and as a state, was a predictor of everyday behavior. This is in line with other recently reported findings (Wilson et al., 2017; Wilt & Revelle, 2017) and highlights the importance of considering in situ person states, such as affect, as predictors of behavior. Including happiness at the state level substantially increased the explained variance in behavior ($\Delta R_{m(\text{Model1, Model3})} = .13$, and $\Delta R_{c(\text{Model1, Model3})} = .09$)[26]. It was further important to distinguish between effects at the person/trait level of happiness and the effect of happiness at the situation/state level. For example, trait subjective happiness did not predict extraverted behavior ($t = 1.60$), but state happiness did ($t = 20.82$)[27].

The re-analyses further revealed that in situ happiness and situation perception overlapped (see also Horstmann & Ziegler, 2018). Horstmann and Ziegler suggested that this overlap could be problematic for situation research: Effects attributed to situation perception might better be attributed to affect (or vice versa). As Table 3.2.7 shows, this is true for many of the effects reported in Model 1 (upper half of Table 3.2.7), but not for all (as seen in the results of Model 4, lower half of Table 3.2.7). Most importantly, the effects initially predicted by Sherman and colleagues (2015, see Table 3.2.1, *Sherman and colleagues (2015), original predictions*) remained significant even after controlling for happiness, thus supporting the exclusivity of the effects of certain situation characteristics on certain behaviors. For example, a person that is happy in a situation acts more extraverted in this situation, but persons that perceive the situation as social are likely to act *even more* extraverted (regression weight of in situ perceived Sociality on extraverted behavior: $b = 0.21$, 95% CI = [0.18, 0.24], $t = 13.65$, see Table 3.2.8). Taken together, the theoretical model that assumes specific effects of situation perception on certain behaviors is thus more accurate when including happiness as an additional predictor.

If these findings were to be found replicable, they would support the validity of situation perception measures: Participants' in situ perceptions of situation characteristics can be distinguished from their current affect with the measurement instruments we used. Moreover, each of the constructs uniquely contributed to the prediction of behavior in a theoretically sensible way.

---

[26] To compute the difference, we took the average $R_c$ and $R_m$ from Model 1 (i.e., situation perception only) and subtracted it from the average $R_c$ and $R_m$ from Model 3, respectively. Note that the effects of happiness between models with different situation perception variables are not independent. State happiness (Model 3) predicted the same behavior in eight models in which only situation perception is changed. Nevertheless, these differences give a good picture of the relative importance of happiness as a predictor. For a more accurate display of effects of level 1 and level 2 happiness, see Table 8.

[27] Note that in the case of trait happiness, one can really speak of a *prediction* of future behavior as the previously assessed trait assessment predicted later behavior during the experience sampling part.

The perception of a situation was thus more than just current affect/internal person states, and this additional variance in situation perception was systematic and tied to the behavior of the person in that situation. We will engage in broader discussion of these findings after presenting the results from Study 2 which extends the exploratory re-analyses of Study 1.

## Study 2

There are some shortcomings of the re-analyses presented in Study 1, thus limiting the robustness of our initial findings and making a second study necessary, which we report here. First, the analyses reported in Study 1 were strictly exploratory. We did not have specific hypotheses other than the general research question of happiness as a predictor of behavior. We also did not expect to find the particular pattern of exclusive effects of situation perception on behavior after controlling for happiness. Furthermore, we also tested different models (e.g., quadratic terms or cross-level-interactions); however, the results of these studies could not be interpreted coherently and were therefore abandoned. Performing many exploratory analyses results in an inflation of Type-I errors (Ioannidis, 2005; Simmons et al., 2011), which may result in detecting and placing value in potentially unreliable findings that ultimately cannot be replicated (see Open Science Collaboration, 2015). Thus, we needed to replicate our patterns of results.

Second, the measure of in situ happiness was suboptimal as only one item with a bipolar rating scale was used, created ad hoc for the assessment of happiness during experience sampling. Although this procedure is common in emotion assessment, it poses a threat for the validity and interpretation of the resulting affect score (Weidman et al., 2017). Further, affect has at least two dimensions, positive and negative, and could also be activating or deactivating (Russell, 2003; Watson & Tellegen, 1985; Winkielman et al., 2007; Yik et al., 1999). These dimensions may be correlated, but they are nevertheless distinct (Dejonckheere et al., 2018). This distinction and heterogeneity is lost when using only one item for the assessment of affect. Based on the findings reported above, it could therefore be possible that a better, more differentiated, and more reliable measure of affect would "swallow" the remaining effects of situation perception on behavior. If this were the case, then the re-analyses in Study 1 would have painted an overly optimistic picture in the exclusive links between certain situation characteristic and certain behaviors (Table 3.2.1). Thus, a second study needed to employ a more differentiated measure of affect.

Due to these two limitations, we aimed to replicate and extend the study by Sherman and colleagues (2015) as well as our re-analysis findings of Study 1 in an independent sample. We pre-registered this replication on the Open Science Framework (OSF) in November 2016 and collected the data in 2016 and early 2017.

### Research Questions

All hypotheses for the current study were pre-registered on November 21st 2016 on the OSF: osf.io/bzv5q. Instead of listing every hypothesis here, we will indicate in the results section if an effect turned out as expected. We deem a replication successful (a) if an effect that was previously shown to be significant in Study 1 is also significant in the analyses of the new data in Study 2 (see Tackett et al., 2017, p. 743) and (b) if an effect that was previously *not* significant is also *not* significant in the replication.

We pre-registered that an effect would be considered significant if its corresponding *p*-value is smaller than .05. We further pre-registered to test all directed effects (i.e., positive or negative) in a one-tailed fashion, and all other effects in a two-tailed one. However, we will deviate from this pre-registration such that we test effects that were pre-registered not to be significant now with an alpha level of .10, thereby accepting a larger Type-I error rate but also lowering the acceptance of the Type II error rate. The test of the hypotheses that an effect is not present is therefore stronger. This means that pre-registered positive or negative effects will be considered significant if the 90% bootstrapped confidence interval (one-tailed) excludes zero and the corresponding *t*-value is larger than $t = 1.64$ (Wald $z$-test). Similarly, hypotheses that an effect is zero will be rejected if the 90% bootstrapped confidence interval (two-tailed) includes zero and if the *t*-value of that effect is smaller than $t = 1.64$. Effects that were not pre-registered or hypothesized to be in any particular direction will be evaluated on an alpha level of .05, two-tailed (e.g., effects of affect in the extended replication of Study 1). This approach is in line with current recommendations to justify the alpha level (Lakens et al., 2018) instead of applying a one-size-fits-all approach in form of one general level of significance (Benjamin et al., 2018).

**Replication: Original study by Sherman et al., 2015.** The first part of the study is a direct replication of the findings provided by Sherman and colleagues. Sherman and colleagues' results referred to four research questions which will also be tested here in the same way as a replication. These four questions concerned (1) the amount of variability of behavior, (2) the amount of variability in situation perceptions, 3) the relation of personality traits to situation perception, and (4) the relation of person and state characteristics to behavior. We expected to replicate the original findings in the new sample, and such replication is important so that we can deem the findings robust. The expected effects were pre-registered in Table OSM B.

**Replication: The specific effect of situation perception (Study 1).** In the second part, we examined if the effect of situation perception on nearly all state expressions could be replicated from Study 1. The replication of this finding was a pre-requisite to support the general hypothesis that effects of situation perception are only specific once controlled for happiness or affect. The

expected effects are displayed in Table 3.2.1, *Study 2: Prediction for the Replication of the Re-analyses from Study 1*.

**Replication: The effect of happiness and situation perception (Study 1).** The third research question addresses to what extent the inclusion of happiness leads to better, more accurate predictions of the effect of situation perception. We expected that, after controlling for happiness, only specific, theoretically relevant situation characteristics would explain behavior. We made specific predictions for each effect that we expected (see Table 3.2.1, *Study 2: Extension, including happiness or/and PA and NA*). These predictions were not only based on the results obtained by Sherman and colleagues (2015), but also on the results of our re-analysis (see Study 1) and subsequent theoretical considerations. For example, we expected that perceiving Intellect in a situation would be associated with more open behavior, but also with more conscientious behavior. At the same time, some effects were "marginally significant" in the re-analysis; based on their low theoretical plausibility and the small magnitude of the effects, we expected that they would *not* replicate. For example, in Study 1 (Model 4), Adversity at level 1 predicted open behavior ($b$ = -0.10, 95% CI = [-0.05, -0.14], $t$ = 1.92). Although the 95% confidence interval indicates significance, we think that this effect would not replicate as it is not theoretically plausible. We have made all of our expectations explicit in the OSM A.

**Extension: Including (better) measures of affect.** The first three questions above addressed the first limitation of our Study 1, namely the exploratory nature of our initial data analysis. This question addresses the second limitation, namely the parsimonious measurement of happiness in Study 1. We expected that even after the inclusion of broader, more reliable measures of affect, the specific and theoretically plausible links between situation perception and behavior would remain significant and sizable. As we were not primarily invested in the predictive power of measures of affect predicting behavior, we did not formulate specific hypotheses for affective variables as predictors.

**Replication *and* extension: Effect of happiness, affect, and multiple dimensions of situation perception.** Sherman and colleagues (2015) did not examine each combination of situation perception and behavior independently but combined several dimensions of situation perception to predict one particular behavior. As was shown in the re-analyses of the data in Study 1, these effects remained significant after controlling for happiness. In Study 2, we further aimed to replicate these results when including measures of happiness and affect simultaneously as level 1 and level 2 predictors for behavior.[28] That is, we expected that behavior could be predicted with

[28] These particular analyses were not pre-registered. However, we consider these analyses strictly confirmatory with respect to the effects of situation perception.

several situation perception dimensions at the same time, and that these effects remained significant after controlling for happiness and affect at level 1 and level 2.

**Method**

**Pre-registration.** In addition to the hypotheses, we explicitly pre-registered all variables collected, the number of participants (see below), and the data-analytical plan. For the data analyses, we pre-registered (1) how each variable would be collected, (2) how each scale would be scored, and (3) how each variable would be transformed prior to analyses in a multilevel-model. All items assessed as well as the transformation for each item can be found in the OSM C. We will explicitly mention which elements in the data analyses (see below) and data processing (see below) were pre-registered and where we deviated from the pre-registration.

**Power analysis and determination of sample size.** Based on Study 1, an a priori Monte-Carlo simulation power analysis was performed to determine the sample size. The power to detect an effect in a multi-level model depends on the number of level 1 units (i.e., number of measurement occasions obtained), the number of level 2 units (i.e., participants sampled), as well as the size of the effect (Mathieu, Aguinis, Culpepper, & Chen, 2012). As many different models had to be analyzed, we took the smallest effect of interest in the final model of Study 1 (i.e., when happiness was included both at the trait and state level in the model predicting agreeable behavior with Sociality) and examined the sample size necessary to replicate this effect with 80% power. For a one-tailed *t*-test (Wald, 1943) at alpha = .05, 60 measurement occasions sampled across 250 participant yielded only 66% power. If the number of measurement occasions increased to 80, 76% power would be obtained; 120 measurement occasions yielded a power > 80%. On the other hand, larger effects (e.g., predicting Conscientiousness with Duty yielded a power > 99.99%). Based on these analyses, we decided to sample $N$ = 250 participants, aiming for $n$ = 45 measurement occasions per person. This will allow detecting smallest effects of interest (e.g., *b*s ~ .10) with sufficient power. At the same time, we used mainly one-tailed tests in Study 2, which further increases test power.

**Sample and procedure.** The data were collected using the open-source platform formr.org (Arslan & Tata, 2015). Participants were first informed about the study and required to submit their e-mail address to participate. Subsequently, participants received an e-mail with an invitation to respond to demographic questions and complete personality measures (see *Trait Measures*). After completing the initial assessment, participants received an e-mail every three hours with an invitation to respond to state measures (see below). The measures were presented in the same order each time. After each assessment, participants could opt out by unchecking a box. Either after opting out or after completing 50 measurement occasions, participants were directed

to a website with a debriefing and a general invitation to participate in further studies. Participants then received personalized feedback via e-mail.

In total, $N$ = 1,128 participants clicked on the link to the study, 367 submitted their e-mail, and 341 participants finalized the initial personality assessment. Of these, 274 provided sufficient data in the experience sampling phase of the study to be included in further analyses. Participants were on average 24.22 years old ($SD$ = 6.35); 14.96% percent were male, 84.31% were female, and 0.73% did not indicate their gender. Most participants were enrolled in a university program (90.51%). The completion of the initial trait assessment took about 19 minutes to complete (maximum number of 80 items), and each experience sampling assessment on average about three minutes to complete. All materials required to replicate the study as well as a detailed procedure can be found on the OSF.

**Trait measures.** The same personality measures as in Study 1were utilized. However, we did not apply measures that were not relevant for the replication of the study.

***Personality Traits.*** The German version of the HEXACO-60 was used (Moshagen, Hilbig, & Zettler, 2014). The six HEXACO personality dimensions were assessed with 10 items each. Items were answered on a 6-point Likert-type scale (1 = "not accurate at all" [*trifft überhaupt nicht zu*], 6 = "completely accurate" [*trifft völlig zu*]). Descriptive statistics, sample items, and internal consistencies can be found in Table 3.2.3.

***Subjective happiness.*** We used the German version of the subjective happiness scale (Lyubomirsky & Lepper, 1999; Swami et al., 2009). It assesses subjective happiness with four items. Instead of using a 7-point Likert-type-scale, we changed this to an 8-point Likert-type scale in order to avoid a middle category (Kulas & Stachowski, 2009). Descriptive statistics, a sample item, and internal consistency are reported in Table 3.2.3.

**State measures.** We used the same measures as Sherman and colleagues for the assessment during experience sampling. Sample items, descriptive statistics, and reliabilities (if applicable) for all state measures are presented in Table 3.2.4. State measures were presented on three different pages. Participants first rated their current affect, then the current situation, and then their behavior in the situation.

***Situation perceptions.*** For the assessment of situation perception, we used the Situation Eight DIAMONDS ultra-brief measure in its original German version (Rauthmann & Sherman, 2016c). Participants responded on an 8-point Likert-type scale how well each item describes the given, current situation.

***Behavior.*** Similar to Sherman and colleagues (2015), we used a bipolar rating scale for the assessment of current, in situ behavior. In total, seven different behaviors were assessed with one

item each. Each item consisted of an 8-point Likert-type-scale, with two adjectives marking each endpoint. The items were translated from Sherman and colleagues' (2015) original study by a bilingual native speaker.

***Happiness.*** Happiness was assessed with the same item as in Sherman et al. (2015). Participants rated their current happiness on a bipolar rating Eight-point Likert-type scale with the end-points 1 = "happy, positive" and 8 = "sad, negative". This item was translated by a bilingual native speaker.

***Affect.*** To assess affect, we used adjective scales developed by Hampel (1977). Hampel originally developed six scales to assess six different aspects of affect or emotion. However, as we aimed to reduce participant burden and keep the number of items minimal during ESM assessment, we decided to assess only two scales, positive affect (PA, German *gehobene Stimmung*, e.g., cheerful, happy, elated) and negative affect (NA, German *Missstimmung*, e.g., angry, edgy, bad-tempered), which are closest to positive and negative affect. Hampel originally developed two parallel seven-adjective scales for each dimension. From the PA scale, we selected the three items with the highest factor loading presented in the original study by Hampel, 1977. From the NA scale, we selected three items that were still widely in use (e.g., *brummig* [grumpy], was not selected), or items that we deemed not to be too extreme for an application on everyday context (e.g., *zornig* [furious]). We decided to use a narrower instead of a broader operationalization of affect in order to obtain more reliable measures and potentially unidimensional scores, allowing better interpretation of the resulting effects. To further reduce participant burden, the activating – deactivating components of affect were also ignored. At each assessment, participants were instructed to rate how they feel during the current situation on an eight-point Likert-type scale (1 = "not accurate at all" [*trifft überhaupt nicht zu*], 8 = "completely accurate" [*trifft völlig zu*]). The scores for PA and NA were formed by taking the average of the three items at each assessment. Descriptive statistics are presented in Table 3.2.4.

**Closeness of the replication.** The current study is a replication and extension of the original study by Sherman and colleagues (2015) upon which our Study 1 here was based. However, Study 2 also differs in several aspects from Study 1. The main difference is of course the language and the country the study was conducted in. Whereas the first data collection took place in the US, the second data collection took place in Germany. However, both studies used mainly undergraduate (psychology) students. Further, the participant recruiting strategy, data collection, and remuneration differed between the two studies. In Study 1, participants were invited into the laboratory for the first session (assessment of personality measures), whereas in Study 2 all assessments took place online. Participants in Study 1 received course credit, and participants in

Study 2 received feedback on their measures as well as course credit (approximately 0.13 ECTS). Further differences between the studies concern the number of points used on the Likert-type-rating scales as well as the determination of the sample size. All differences between the two studies are listed in Table 3.2.9. All other elements (e.g., scales used, scoring of scale composites, software used) were kept identical, with the exception that all measures and texts were in German.

To assess the outcome of the replication, it is important that both studies differ as little as possible, or only in aspects irrelevant to the effect examined. However, it is not easy to determine which aspects are relevant to the outcome of the study and which aspects are not. It was therefore our aim to replicate Study 1 as closely as possible. Based on the study differences presented in Table 3.2.9, we argue that Study 2 is a *very close, direct replication* of Study 1 (LeBel, Berger, Campbell, & Loving, 2017). Different outcomes of the two studies should therefore not be attributed to the design, but to the robustness of the effects examined.

Table 3.2.9
*Differences between Study 1 and Study 2*

| **Design Facet** | **Study 1** | **Study 2** | **Explanation/Reason** |
|---|---|---|---|
| Language | English | German | |
| Physical setting | Florida/USA | Berlin/Germany | |
| **Method** | | | |
| Likert Scale Used | | | |
| Subjective Happiness Scale | 1 – 7 | 1 – 8 | avoid middle category, increase variance |
| ESM assessment | 1 – 7 | 1 – 8 | avoid middle category, increase variance |
| Number of Items | | | |
| Trait/Personality assessment | 60 | 80 (some optional) | not all relevant for replication |
| ESM assessment | 18 | 25 (some optional) | added six items for affect, geo-position, removed self-esteem and authenticity |
| **Procedure** | | | |
| Participant recruitment | campus, students | e-mail list to students | |
| Participant briefing | in laboratory | online | |
| Remuneration | money (270USD) | course credit, feedback | no money available |
| Trait assessment | in lab | online | easier to conduct |
| ESM assessment | SMS | E-mail | no money for SMS available |
| Presentation of ESM Items | randomized | same order | |
| Exclusion of measurement occasions | If participants did not respond within one hour | If participants were too quick (i.e., less than 10 sec.) | Time after invitation not tracked, not assumed to be relevant |
| Average Number ESM Assessments | $M$ = 46.44, $SD$ = 9.61 | $M$ = 28.19, $SD$ = 18.23 | |
| **Analysis** | | | |
| exclusion of participants | not responded within one hour | less than 3 measurement occasions | |
| Significance level | .05 | .10 for expected non-significant effects, .05, one-tailed for predicted effects | |
| **Sample characteristics** | | | |
| Sample size | 209 | 274 | power-analyses, at least $N$ = 250 |
| male/female ratio | 136/73 | 41/231 | |
| Age | $M$ = 18.61, $SD$ = 1.78 | $M$ = 24.22, $SD$ = 6.35 | |

**Processing of experience sampling data.** We first excluded participants with less than three measurement occasions. We then excluded measurement occasions in which participants responded in less than 10 seconds (less than 1% of the measurement occasions). Finally, we checked the amount of missing data during experience sampling, and no variable had more than 1.20% missing data. Therefore, no data were imputed. This procedure was pre-registered. However, contrary to our pre-registration, we did not exclude participants that showed very long (> 20 minutes) reaction times as this would have meant to exclude 440 measurement occasions. These reaction times occurred when a participant clicked on a link but finished the survey later.

**Computation of scale scores and centering.** We computed the scale scores for each variable. Specifically, we took the average of all items belonging to one scale to form the scale composite (e.g., all items for extraversion were averaged). The same procedure was applied for state positive and negative affect, as both were assessed with three items each. Finally, level 2 variables (at the person level) were grand-mean centered and level 1 predictors (at the situation level) were person-mean centered (Enders & Tofighi, 2007). The means of level 2 variables (e.g., the mean of all items for Duty) were included in the multi-level models as level 1 predictors to account for between person mean differences. The computation of scale scores and the procedure of multi-level models analyzed exactly mirrors those in Study 1.

**Analyses**

We computed descriptive statistics and correlations of the scales used (see Tables 3, 4, and 10). Subsequently, we ran several multi-level analyses to examine the research questions. First, we replicated the findings for Questions 1-3 in Sherman and colleagues (i.e., , Sources of State-Expression (i.e., behavior), Sources of Situation Experiences (i.e., situation perception), Association between Personality Traits and Situation Experiences).

Second, we examined if specific associations between situation perception and behavior could be detected. To this end, we first predicted behavior using personality traits, mean situation perceptions, as well as in situ situation perceptions. We then examined if the findings from our re-analysis including happiness at person- and situation level could be confirmed. Finally, we included mean positive and mean negative affect at level 2 as well as in situ positive and negative affect at level 1 as additional predictors of behavior. The results for the level 1 and level 2 effects of situation perception are presented in Tables 12 and 13.

Third, we examined if the effects of several situation perception dimensions on behavior could be replicated. Again, after replicating the original findings from Sherman and colleagues (2015), we extended our analyses by including happiness level 2 and level 1, and subsequently positive and negative affect at level 2 and level 1. This last analysis (including happiness and affect

at the trait and state level) was not pre-registered. However, we would expect the effects presented by Sherman and colleagues to replicate, even if happiness as well as affect are included in the model as additional predictors. The results of these last analyses are presented in Table 3.2.8. Due to the large number of pre-registered hypotheses and effects reported, we computed for each research question specificity and sensitivity indices for the tests of our pre-registered hypotheses. Both indices served as a measure for overall success of our predictions.

**Results**

**Descriptive statistics and correlations.** Descriptive statistics and estimates for reliability of the measures used are presented in Table 3.2.3 and 7.2.4. The reliability estimates (Cronbach's alpha) for all trait measures ranged from .70 to .83 and is comparable to the estimates presented in Study 1. All personality measures showed sufficient variability between participants. Regarding experience sampling measures, participants completed a lower number of measurement occasions across all reports. The intra-class correlations (ICC) were similar in magnitude across both studies. Comparing descriptive statistics from Study 1 and Study 2 (even though means should not be directly compared) did not reveal any striking differences that warrant further attention.

Bivariate correlations of the measures used in Study 2 are reported in Table 3.2.10. The convergent correlations of mean behavior and personality traits (i.e., the correlation of a trait score with its corresponding average behavior) were consistently higher than their discriminant correlations, pointing towards construct validity of the bipolar rating scales used for the assessment of behavior. The trait variables also correlated only moderately with the mean affect variables, again with the exception of subjective happiness which showed strong correlations to all three mean affect variables (all $|r|$s $> .46$). Mean situation perception variables correlated moderately to highly among each other as well as with mean affect variables. Finally, mean situation variables correlated moderately with mean behavior variables. Mean affect and mean behavior also correlated moderately to strongly with each other. Indeed, across the whole correlation matrix presented in Table 3.2.10, these correlations were by far the highest, pointing towards the strong nexus between affect and behavior. Overall, the pattern of correlations presented in Table 3.2.10 supports the assumption that participants responded sensibly and differentially to the items instead of just responding in a general pattern during each measurement occasion. We take this as prima facie support of the validity of the scales used.

Table 3.2.10

*Bivariate Pearson-Correlations of Trait Variables and Mean State Variables in Study 2*

| | | | Trait | | | | | | | Mean situational perception | | | | | | | | Mean affect | | | Mean behavior | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | *M* | *SD* | H | E | X | A | C | O | SHS | D | I | A | M | O | N | D | S | PA | NA | SHS | H | E | XS | XD | A | C | O |
| *Trait variables* | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Honesty/Humility | 4.11 | 0.76 | – | | | | | | | | | | | | | | | | | | | | | | | | |
| Emotionality | 3.86 | 0.84 | .10 | – | | | | | | | | | | | | | | | | | | | | | | | |
| eXtraversion | 3.80 | 0.85 | .02 | -.13 | – | | | | | | | | | | | | | | | | | | | | | | |
| Agreeableness | 3.75 | 0.74 | .21 | -.02 | .00 | – | | | | | | | | | | | | | | | | | | | | | |
| Conscientiousness | 4.15 | 0.83 | .10 | .08 | -.03 | -.03 | – | | | | | | | | | | | | | | | | | | | | |
| Openness | 4.42 | 0.77 | .00 | -.14 | .25 | -.03 | .00 | – | | | | | | | | | | | | | | | | | | | |
| Subjective Happiness | 5.15 | 1.37 | .08 | -.16 | .67 | .13 | -.02 | .14 | – | | | | | | | | | | | | | | | | | | |
| *Mean situation variables* | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Duty | 4.49 | 1.40 | -.03 | .01 | .06 | -.02 | .11 | .07 | -.01 | – | | | | | | | | | | | | | | | | | |
| Intellect | 3.80 | 1.25 | -.06 | .07 | -.01 | -.07 | .14 | .06 | -.15 | .68 | – | | | | | | | | | | | | | | | | |
| Adversity | 1.59 | 0.69 | -.14 | .12 | -.13 | -.14 | -.04 | -.03 | -.26 | .06 | .29 | – | | | | | | | | | | | | | | | |
| Mating | 2.55 | 1.43 | .02 | .06 | .13 | -.15 | .08 | .06 | .03 | -.01 | .09 | .25 | – | | | | | | | | | | | | | | |
| pOsitivity | 4.48 | 1.03 | .11 | -.14 | .35 | .10 | -.04 | .17 | .44 | -.22 | -.11 | -.1 | .26 | – | | | | | | | | | | | | | |
| Negativity | 3.20 | 1.30 | -.10 | .22 | -.33 | -.07 | .06 | -.08 | -.47 | .37 | .46 | .48 | .07 | -.39 | – | | | | | | | | | | | | |
| Deception | 1.40 | 0.67 | -.20 | .02 | -.14 | -.12 | -.10 | -.08 | -.26 | .09 | .28 | .75 | .26 | -.12 | .49 | – | | | | | | | | | | | |
| Sociality | 4.21 | 1.28 | -.11 | .11 | .21 | -.10 | .03 | .15 | .16 | .06 | .13 | .24 | .47 | .35 | .08 | .24 | – | | | | | | | | | | |
| *Mean affect variables* | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Positive Affect | 4.16 | 0.98 | .01 | -.07 | .41 | .10 | -.03 | .16 | .46 | -.05 | .08 | .00 | .20 | .73 | -.30 | .00 | .32 | – | | | | | | | | | |
| Negative Affect | 2.39 | 0.94 | -.15 | .18 | -.31 | -.19 | -.05 | -.11 | -.48 | .19 | .25 | .50 | .12 | -.45 | .65 | .45 | .02 | -.31 | – | | | | | | | | |
| Happiness | 5.61 | 1.02 | .06 | -.18 | .51 | .09 | -.04 | .13 | .64 | -.09 | -.13 | -.30 | .02 | .67 | -.64 | -.35 | .12 | .64 | -.65 | – | | | | | | | |
| *Mean behavior variables* | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Honesty/Humility | 6.15 | 0.79 | .24 | .02 | .07 | .07 | .18 | .00 | .12 | .07 | .06 | -.15 | -.06 | .12 | -.13 | -.20 | .06 | .02 | -.24 | .29 | – | | | | | | |
| Emotionality | 4.07 | 1.08 | -.06 | .33 | -.09 | -.12 | .05 | -.01 | -.24 | .12 | .14 | .23 | .18 | -.23 | .51 | .25 | .17 | -.15 | .44 | -.45 | -.17 | – | | | | | |
| Sociability | 5.20 | 0.88 | .10 | -.06 | .53 | .04 | -.04 | .16 | .42 | -.05 | -.01 | -.03 | .13 | .40 | -.32 | -.04 | .30 | .46 | -.29 | .62 | .27 | -.14 | – | | | | |
| Dominance | 5.13 | 0.78 | .00 | -.24 | .43 | -.10 | -.03 | .23 | .38 | .00 | -.08 | -.15 | .05 | .42 | -.31 | -.11 | .18 | .37 | -.35 | .58 | .19 | -.31 | .46 | – | | | |
| Agreeableness | 6.16 | 0.80 | .17 | .12 | .30 | .20 | .03 | .09 | .35 | -.03 | .00 | -.18 | .05 | .42 | -.26 | -.19 | .17 | .33 | -.41 | .57 | .50 | -.16 | .53 | .35 | – | | |
| Conscientiousness | 5.16 | 0.99 | .13 | -.04 | .16 | .00 | .33 | .07 | .19 | .25 | .24 | -.03 | .02 | .07 | -.07 | -.11 | .06 | .12 | -.13 | .27 | .21 | -.01 | .24 | .29 | .13 | – | |
| Openness | 5.34 | 0.91 | .07 | -.18 | .28 | .03 | .14 | .26 | .22 | .11 | .22 | -.05 | .00 | .29 | -.17 | -.06 | .13 | .34 | -.28 | .45 | .24 | -.13 | .42 | .44 | .29 | .60 | – |

**Replication: Original study by Sherman et al., 2015.** We first present the results from the replication of Sherman and colleagues original research. Specifically, we replicate the findings for their Questions 1-3.

***Replication: Sources of behavior (Question 1).*** This research question addresses whether the behaviors assessed are associated primarily with aspects of the person or aspects of the measurement occasion. Replicating Sherman and colleagues' work, unconditional fixed effects models were estimated to obtain the ratio of within- versus between-person variance, which is a form of ICC (Cohen, Cohen, West, & Aiken, 2003). A high ICC (see Table 3.2.4) indicates that the variance of the measure is due to the level 2 cluster, in this case the person. We hypothesized and pre-registered ICCs around .30 for the behavioral measures. Contrary to our expectations, the ICCs for the behavioral items were slightly lower (i.e., ~ .24). Nevertheless, variance in behavior could be attributed to the person and the current situation/state of the person, and further investigation using mixed effects models was therefore possible and warranted.

***Replication: Sources of situation perception (Question 2).*** Similar to the ICCs reported above, ICCs were computed for the situation measures. Contrary to our expectations, these were not lower than the ICCs for the behavioral measures (i.e., ~ .23), which means that this finding by Sherman et al. (2015) could not be replicated in the current study. Furthermore, we expected the ICCs for the affect measures to lie between the ICCs of behavior and situation experiences. However, they were a little bit higher than the other ICCs (i.e., ~ .26). Thus, the first set of hypotheses could not be confirmed.

***Replication: Association between personality traits and situation perception (Question 3).*** Sherman and colleagues (2015) examined if and to what extent situation perceptions were related to personality traits. Although not all of the tested associations were found to be significant in Sherman and colleagues study, we expected all of these effects to be significant in Study 2 due to their theoretical plausibility. The expected effects as well as the results are presented in Table 3.2.11. Seven out of eight tested effects were significant as expected, resulting in a sensitivity of .86 and confirming results and expectations from Sherman et al. (2015).

Table 3.2.11
*Predicting Situation Perception from Personality Traits*

| Situation Characteristic | pred. | *b* | LL | UL | *t* | $R_m$ | $R_c$ |
|---|---|---|---|---|---|---|---|
| Deception | | 1.37 | 1.31 | 1.42 | | .13 | .53 |
| Trait Honesty/Humility | - | -0.17 | -0.25 | -0.09 | **-3.54** | | |
| *SD* in Intercepts | | 0.56 | 0.51 | 0.60 | | | |
| *SD* in Residuals | | 0.93 | 0.91 | 0.94 | | | |
| Negativity | | 3.12 | 3.00 | 3.24 | | .13 | .54 |
| Trait Emotionality | + | 0.32 | 0.18 | 0.47 | **3.71** | | |
| *SD* in Intercepts | | 1.11 | 1.01 | 1.20 | | | |
| *SD* in Residuals | | 1.78 | 1.75 | 1.80 | | | |
| Sociality | | 4.25 | 4.14 | 4.37 | | .10 | .40 |
| Trait eXtraversion | + | 0.33 | 0.18 | 0.47 | **3.95** | | |
| *SD* in Intercepts | | 1.01 | 0.92 | 1.09 | | | |
| *SD* in Residuals | | 2.43 | 2.40 | 2.46 | | | |
| pOsitivity | | 4.55 | 4.45 | 4.65 | | .05 | .45 |
| Trait Agreeableness | + | 0.14 | 0.01 | 0.28 | **1.80** | | |
| *SD* in Intercepts | | 0.86 | 0.78 | 0.93 | | | |
| *SD* in Residuals | | 1.73 | 1.70 | 1.75 | | | |
| Deception | | 1.37 | 1.31 | 1.43 | | .08 | .53 |
| Trait Agreeableness | - | -0.12 | -0.21 | -0.05 | **-2.47** | | |
| *SD* in Intercepts | | 0.56 | 0.52 | 0.61 | | | |
| *SD* in Residuals | | 0.93 | 0.91 | 0.94 | | | |
| Duty | | 4.44 | 4.30 | 4.57 | | .06 | .44 |
| Trait Conscientiousness | + | 0.19 | 0.04 | 0.36 | **1.93** | | |
| *SD* in Intercepts | | 1.17 | 1.06 | 1.29 | | | |
| *SD* in Residuals | | 2.38 | 2.35 | 2.42 | | | |
| Intellect | | 3.82 | 3.69 | 3.94 | | .02 | .48 |
| Trait Openness | + | 0.05 | -0.11 | 0.21 | 0.49 | | |
| *SD* in Intercepts | | 1.14 | 1.03 | 1.23 | | | |
| *SD* in Residuals | | 2.10 | 2.07 | 2.13 | | | |
| pOsitivity | +*a)* | 4.53 | 4.45 | 4.62 | | .22 | .45 |
| Trait Happiness | | 0.32 | 0.25 | 0.39 | **8.33** | | |
| *SD* in Intercepts | | 0.75 | 0.67 | 0.81 | | | |
| *SD* in Residuals | | 1.73 | 1.70 | 1.75 | | | |

*Note.* Effects in bold are significant. An effect is significant if the 90% CI excludes zero as well as if the *t*-value is larger than 1.64 (Wald $\chi$-test).
pred. = predicted and pre-registered effect, + = a positive effect was expected (*a)* the effect was not pre-registered but expected to be positive), - a negative effect was expected; LL and UL are the lower and upper limit of the 90% bootstrapped confidence-interval; *t* = *t*-value of the fixed effect of personality on situational perception; $R_m$ = marginal multiple *R*; $R_c$ = conditional multiple *R*. Effects in bold are significant.

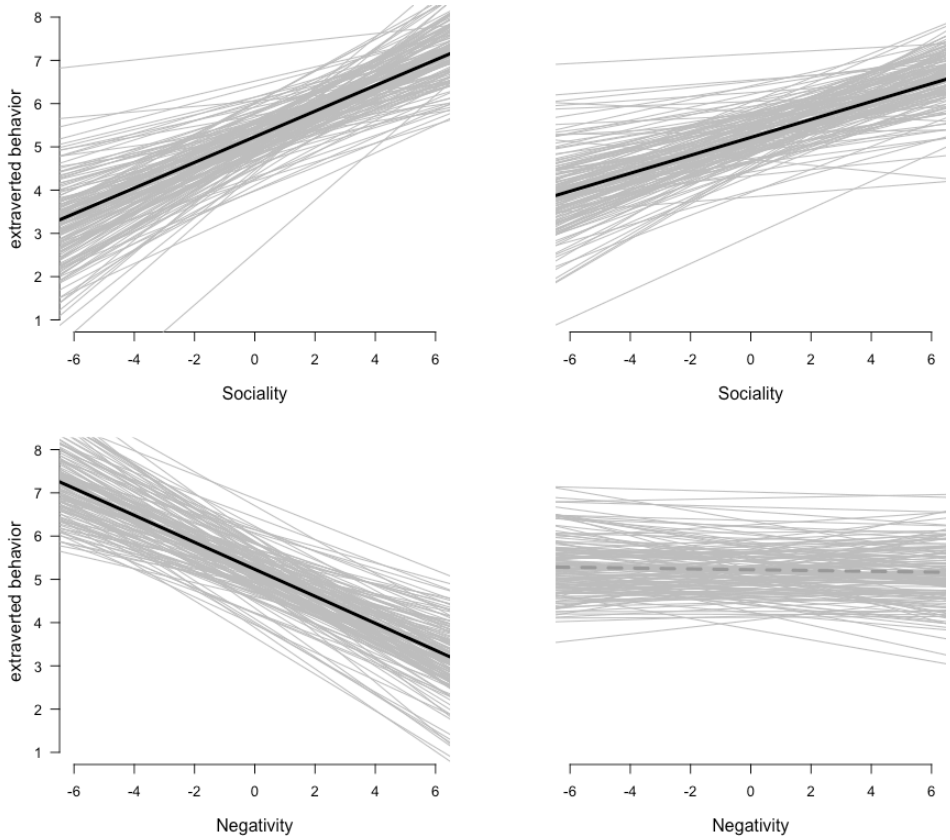
**Replication of Study 1: Specific effects of situation perception on behavior.** To examine the specificity of situation perception effects on behavior, behavior was predicted with the corresponding trait variables (e.g., conscientious behavior with trait conscientiousness), the mean situation variable (e.g., Duty averaged across a situations), and the in situ situation perception (e.g., Duty at a given moment). This model was then extended, and happiness was included at level 1 and level 2, and then positive and negative affect were included at level 1 and level 2. We report the effects of situation perception only, but all other effects (i.e., of personality, happiness, and affect) can be found in the OSM D. The predicted effects are presented in Table 3.2.1, separately for the models with (expected for Model 2 and Model 3, lower third) and without happiness or affect (expected for Model 1, middle third).

***Level 1 effects of situation perception on behavior.*** Table 3.2.12 displays the level 1 effects of in situ situation perception on behavior. For Model 1, we expected nearly all level 1 situation perception variables to exert significant effects on behavior, and our hypotheses of the non-specificity of situation characteristics for the prediction of behavior was largely confirmed (sensitivity: .93, specificity: .20). Comparable to results from Study 1, nearly all situation characteristics were significantly related to all behavioral outcomes.

After the inclusion of happiness at level 1 and level 2 (Model 2, Table 3.2.12), we expected only logically coherent relations between measures of situation perception and behavior to remain significant. This was true for many cases (e.g., the effect of Duty on Emotionality behavior was no longer significant, $t$ = -0.51, whereas it was significant in the model without happiness). For the final model (Model 3, Table 3.2.12) that included mean positive and negative affect and in situ positive and negative affect as additional predictors, we expected the pre-registered effects not to change. This was true for some effects (e.g., the effect of Duty on conscientiousness behavior remained significant in all three models); however, other effects that were predicted not to be significant remained significant in the last model (e.g., the effect of Mating on agreeable behavior remained significant, $t$ = 4.03, although we expected it not to be significant after including affect). This effect is exemplarily presented in Figure 3.2.1. On the left side of Figure 3.2.1, extraverted behavior can be explained with perceived Sociality (upper panel) and perceived Negativity (lower panel, both Model 1). After including PA and NA as well as happiness at level 1 and level 2, only the effect to Sociality remained significant and largely unchanged (Model 3). Consequently, the sensitivity remained unchanged (.93), but the specificity of our predictions increased slightly (.38). This means that many of the effects predicted could be confirmed, but at the same time, several effects that were predicted not to be significant remained significant.

For the last model, sensitivity decreased to .81, but specificity increased to .52. Compared to the previous model, several effects that should have been significant were no longer significant, but fewer effects that should not have been significant also were not significant. Thus, the inclusion of affect and happiness both at situation and person level led to more exclusive and precise effects of situation perception on behavior.

*Figure 3.2.1.* Spaghetti plots for the prediction of extraverted behavior




*Note.* Predictors are Sociality (upper panels) and Negativity (lower panels). Left panels (Model 1): prediction of behavior with situation perception at level 1 and level 2 only. Right panels (Model 3): prediction of behavior with situation perception (level 1 and 2), happiness (level 1 and 2) and positive and negative affect (level 1 and 2). Black solid indicates a significant fixed effect, and the grey dotted line indicates a non-significant fixed effect for situation perception on behavior. Grey lines represent individual regression lines. The predictors (*x*-axis) were centered within person.

***Level 2 effects of situation perception on behavior.*** Similar to the effects of level 1 situation perception, we examined situation perception at level 2. All predictions were derived from Study 1. First, we again examined only situation perception at level 1 and level 2 and personality traits as a predictor (Model 1, Table 3.2.13). Overall, we expected fewer significant effects than at level 1. For the first model, sensitivity was .67, and specificity was .57. The inclusion of happiness at level 1 and level 2 (Model 2, Table 3.2.13) reduced the precision of our predictions further, as effects that were assumed to be significant were no longer significant and effects that were expected not to be significant remained substantial (sensitivity: .61, specificity: .50). Finally, the inclusion of mean positive and negative affect at level 2 (Model 3, Table 3.2.13) changed a lot of the predicted effects such that they were no longer significant, and only five expected effects remained significant. Thus, sensitivity and specificity both further dropped to .36 and .29, respectively. Note that more effects were significant in total, even though (based on Study 1) we did not expect them to be significant. Thus, we conclude that situation perception in general allows explaining variance in behavior, but that our predictions of these effects were rather poor; hence, the effects of mean situation perception on behavior are not yet well understood.

Table 3.2.12

*Effects of Situation Perception at Level 1 in Study 2*

| Sit.-Perc. | Duty | | Intellect | | Adversity | | Mating | | pOsitivity | | Negativity | | Deception | | Sociality | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Behavior | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* |
| *Model 1: Predicting Behavior from Situation Perception at level 1 and level 2* | | | | | | | | | | | | | | | | |
| H | *-0.05 [-0.06; -0.04]* | ***-6.59*** | *-0.01 [-0.02; 0.00]* | *-0.98* | *-0.14 [-0.17; -0.11]* | ***-6.87*** | 0.02 [0.00; 0.04] | **1.99** | *0.12 [0.10; 0.14]* | ***8.75*** | *-0.11 [-0.13; -0.09]* | ***-8.40*** | *-0.21 [-0.26; -0.16]* | ***-7.18*** | *0 [-0.01; 0.01]* | *-0.05* |
| E | *0.01 [-0.01; 0.03]* | *0.87* | *0.05 [0.03; 0.07]* | ***3.90*** | *0.32 [0.29; 0.36]* | ***13.91*** | 0.06 [0.04; 0.09] | **4.79** | *-0.15 [-0.19; -0.12]* | ***-7.37*** | *0.35 [0.33; 0.38]* | ***20.37*** | *0.26 [0.21; 0.32]* | ***8.20*** | *0.07 [0.05; 0.09]* | ***5.92*** |
| X | *-0.07 [-0.08; -0.05]* | ***-5.60*** | *-0.03 [-0.05; -0.01]* | ***-2.53*** | *-0.13 [-0.17; -0.09]* | ***-5.37*** | *0.18 [0.15; 0.21]* | ***11.58*** | *0.45 [0.43; 0.48]* | ***30.09*** | *-0.31 [-0.34; -0.28]* | ***-20.10*** | *-0.13 [-0.18; -0.07]* | ***-4.05*** | *0.30 [0.28; 0.32]* | ***23.5*** |
| A | *-0.07 [-0.09; -0.05]* | ***-7.47*** | *-0.04 [-0.05; -0.02]* | ***-3.50*** | *-0.27 [-0.31; -0.24]* | ***-13.10*** | *0.09 [0.07; 0.11]* | ***7.12*** | *0.34 [0.31; 0.36]* | ***21.14*** | *-0.3 [-0.33; -0.28]* | ***-22.50*** | *-0.28 [-0.34; -0.22]* | ***-8.37*** | *0.11 [0.09; 0.12]* | ***10.52*** |
| C | *0.25 [0.23; 0.27]* | ***17.7*** | *0.22 [0.19; 0.25]* | ***12.69*** | *-0.04 [-0.08; -0.01]* | ***-2.05*** | 0.00 [-0.02; 0.02] | -0.04 | 0.08 [0.05; 0.11] | **4.67** | -0.03 [-0.06; 0.01] | **-1.73** | *-0.06 [-0.10; -0.02]* | ***-2.83*** | *0.07 [0.05; 0.09]* | ***6.15*** |
| O | *0.10 [0.08; 0.12]* | ***8.72*** | *0.17 [0.14; 0.20]* | ***11.30*** | *-0.09 [-0.13; -0.07]* | ***-5.22*** | *0.04 [0.02; 0.06]* | ***3.41*** | *0.22 [0.20; 0.25]* | ***15.43*** | *-0.12 [-0.15; -0.09]* | ***-8.31*** | *-0.07 [-0.11; -0.03]* | ***-3.02*** | *0.10 [0.08; 0.11]* | ***9.64*** |
| D* | 0.02 [0; 0.03] | 1.98 | -0.01 [-0.02; 0.01] | -0.54 | -0.03 [-0.06; 0.00] | -1.62 | 0.02 [0.01; 0.04] | 2.23 | 0.17 [0.14; 0.19] | 12.45 | -0.14 [-0.16; -0.12] | -10.30 | -0.09 [-0.12; -0.05] | -3.93 | 0.06 [0.05; 0.07] | 7.41 |
| *Model 2: Predicting Behavior from Situation Perception at level 1 and level 2, and Happiness at level 1 and level 2* | | | | | | | | | | | | | | | | |
| H | -0.03 [-0.05; -0.02] | **-4.75** | 0.00 [-0.01; 0.02] | 0.31 | *-0.09 [-0.12; -0.06]* | ***-4.74*** | 0.00 [-0.01; 0.02] | 0.08 | 0.05 [0.03; 0.07] | **3.68** | -0.04 [-0.06; -0.01] | **-2.93** | *-0.17 [-0.21; -0.11]* | ***-5.89*** | -0.03 [-0.04; -0.01] | **-3.07** |
| E | -0.01 [-0.02; 0.01] | -0.51 | *0.03 [0.01; 0.05]* | ***2.47*** | *0.22 [0.19; 0.26]* | ***10.58*** | 0.10 [0.07; 0.12] | **7.31** | *0.04 [0.01; 0.07]* | ***†1.82*** | *0.31 [0.28; 0.34]* | ***17.03*** | *0.16 [0.12; 0.21]* | ***5.77*** | 0.12 [0.10; 0.14] | **9.82** |
| X | -0.02 [-0.04; -0.01] | **-2.52** | 0.00 [-0.01; 0.02] | 0.34 | 0.07 [0.05; 0.11] | **3.88** | *0.12 [0.10; 0.13]* | ***10.36*** | *0.19 [0.16; 0.22]* | ***11.56*** | *-0.04 [-0.06; -0.01]* | ***-2.42*** | 0.07 [0.02; 0.11] | **2.64** | *0.22 [0.20; 0.24]* | ***20.57*** |
| A | -0.03 [-0.04; -0.02] | **-4.44** | -0.01 [-0.02; 0.00] | -0.99 | *-0.11 [-0.15; -0.08]* | ***-6.64*** | 0.03 [0.02; 0.05] | **3.93** | *0.09 [0.07; 0.12]* | ***6.76*** | *-0.09 [-0.11; -0.07]* | ***-7.25*** | *-0.12 [-0.16; -0.07]* | ***-4.32*** | *0.03 [0.02; 0.04]* | ***4.29*** |
| C | *0.27 [0.24; 0.29]* | ***19.59*** | *0.24 [0.21; 0.26]* | ***13.85*** | *0.04 [0.01; 0.07]* | ***1.88*** | -0.03 [-0.05; -0.01] | **-2.15** | *-0.11 [-0.15; -0.07]* | ***-5.18*** | *0.13 [0.10; 0.16]* | ***6.70*** | 0.02 [-0.02; 0.05] | 0.69 | *0.04 [0.02; 0.06]* | ***3.43*** |
| O | *0.13 [0.11; 0.15]* | ***12.08*** | *0.19 [0.17; 0.21]* | ***14.49*** | *0.03 [0.00; 0.05]* | *1.62* | 0.00 [-0.02; 0.02] | 0.30 | 0.02 [-0.01; 0.04] | 1.19 | 0.08 [0.06; 0.10] | **5.21** | *0.05 [0.02; 0.08]* | ***2.29*** | *0.05 [0.03; 0.07]* | ***5.15*** |
| D* | 0.03 [0.02; 0.04] | 4.15 | 0.01 [0.00; 0.02] | 1.83 | 0.06 [0.03; 0.08] | 3.39 | -0.01 [-0.02; 0.01] | -0.73 | 0.04 [0.02; 0.07] | 3.19 | -0.03 [-0.06; -0.01] | -2.47 | 0.00 [-0.03; 0.03] | -0.09 | 0.02 [0.01; 0.04] | 3.12 |
| *Model 3: Predicting Behavior from Situational Perception at level 1 and level 2, Happiness at level 1 and level 2, and PA and NA at level 1 and level 2* | | | | | | | | | | | | | | | | |
| H | -0.03 [-0.04; -0.02] | **-4.83** | 0.01 [-0.01; 0.02] | 0.72 | *-0.04 [-0.07; -0.01]* | ***-2.25*** | 0.00 [-0.01; 0.02] | 0.27 | 0.05 [0.02; 0.07] | **3.48** | 0.01 [-0.01; 0.03] | 1.00 | *-0.14 [-0.18; -0.09]* | ***-5.04*** | -0.01 [-0.03; 0.00] | -1.58 |
| E | -0.01 [-0.03; 0.01] | -1.17 | *0.02 [0.01; 0.04]* | ***2.18*** | *0.15 [0.12; 0.18]* | ***7.27*** | 0.08 [0.06; 0.10] | **6.60** | *0.01 [-0.02; 0.05]* | *0.67* | *0.28 [0.25; 0.31]* | ***15.04*** | *0.10 [0.06; 0.15]* | ***3.79*** | 0.09 [0.07; 0.11] | **7.59** |
| X | -0.01 [-0.03; 0.00] | -1.44 | 0.00 [-0.01; 0.02] | 0.45 | 0.08 [0.05; 0.11] | **4.51** | *0.10 [0.08; 0.11]* | ***8.99*** | *0.12 [0.09; 0.14]* | ***7.31*** | *-0.01 [-0.03; 0.02]* | *-0.63* | 0.06 [0.02; 0.10] | **2.61** | *0.21 [0.19; 0.23]* | ***19.62*** |
| A | -0.02 [-0.03; -0.01] | **-2.61** | 0.00 [-0.01; 0.01] | 0.08 | *-0.04 [-0.07; -0.01]* | ***-2.46*** | 0.03 [0.02; 0.05] | **4.03** | *0.06 [0.04; 0.08]* | ***4.90*** | *-0.02 [-0.04; 0.00]* | *-1.54* | *-0.06 [-0.11; -0.02]* | ***-2.5*** | *0.05 [0.04; 0.06]* | ***6.20*** |
| C | *0.27 [0.24; 0.29]* | ***19.97*** | *0.23 [0.20; 0.26]* | ***13.98*** | *0.02 [-0.01; 0.06]* | *1.06* | -0.03 [-0.06; -0.01] | **-2.47** | *-0.15 [-0.19; -0.11]* | ***-5.82*** | *0.14 [0.11; 0.17]* | ***7.42*** | 0.00 [-0.03; 0.04] | 0.15 | *0.03 [0.01; 0.05]* | ***2.77*** |
| O | *0.13 [0.12; 0.15]* | ***12.55*** | *0.19 [0.17; 0.21]* | ***14.50*** | *0.02 [-0.01; 0.05]* | *1.33* | -0.01 [-0.03; 0.01] | -0.57 | -0.02 [-0.05; 0.01] | -1.45 | 0.10 [0.08; 0.13] | **6.64** | *0.04 [0.01; 0.08]* | ***†2.08*** | *0.04 [0.02; 0.05]* | ***3.99*** |
| D* | 0.03 [0.02; 0.04] | 4.02 | 0.01 [-0.01; 0.02] | 0.98 | 0.03 [0.01; 0.05] | 1.69 | -0.01 [-0.03; -0.01] | -1.68 | 0.02 [0.00; 0.05] | 1.64 | -0.06 [-0.08; -0.04] | -4.19 | -0.03 [-0.07; 0.00] | -1.61 | 0.01 [-0.01; 0.02] | 0.84 |

*Note.* Sit.-Perc. = Situational Perception (predictor level 1): *b* = estimate (unstandardized) effect; LL = lower limit of 90% CI, UL = upper limit of 90% CI; *t* = *t*-value of fixed effect.
Predicted behaviors: H = Honesty/Humility; E = Emotionality; X = eXtraversion; A = Agreeableness; C = Conscientiousness; O = Openness;
Bold effects are significant. Effects in *italics* were predicted to be significant. † = Predicted effects were significant in the opposite of the predicted direction.

Table 3.2.13

*Effects of Situation Perception at Level 2 in Study 2*

| Sit.-Perc. | Duty | | Intellect | | Adversity | | Mating | | pOsitivity | | Negativity | | Deception | | Sociality | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Behavior | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* | *b* [LL, UL] | *t* |
| *Model 1: Predicting Behavior from Situation Perception at level 1 and level 2* | | | | | | | | | | | | | | | | |
| H | 0.05 [-0.01; 0.12] | 1.49 | 0.06 [-0.01; 0.12] | 1.38 | -0.10 [-0.22; 0.01] | -1.39 | -0.03 [-0.09; 0.02] | -1.01 | *0.10 [0.03; 0.18]* | ***2.12*** | *-0.05 [-0.11; 0.02]* | *-1.33* | *-0.16 [-0.29; -0.02]* | ***-2.02*** | 0.05 [-0.02; 0.12] | 1.26 |
| E | 0.09 [0.02; 0.16] | **2.04** | *0.09 [0.01; 0.17]* | ***1.94*** | *0.26 [0.11; 0.40]* | ***3.04*** | 0.11 [0.04; 0.17] | **2.65** | *-0.12 [-0.22; -0.01]* | ***-1.97*** | *0.36 [0.29; 0.43]* | ***8.12*** | *0.28 [0.13; 0.42]* | ***3.03*** | 0.13 [0.05; 0.21] | **2.69** |
| X | -0.08 [-0.14; -0.02] | **-2.16** | *-0.01 [-0.07; 0.06]* | *-0.25* | *0.02 [-0.1; 0.14]* | *0.25* | 0.03 [-0.03; 0.08] | 0.84 | *0.26 [0.18; 0.33]* | ***5.15*** | *-0.12 [-0.19; -0.06]* | ***-3.10*** | *-0.04 [-0.16; 0.10]* | *-0.49* | *0.17 [0.10; 0.23]* | ***4.42*** |
| A | -0.02 [-0.07; 0.04] | -0.44 | *0.00 [-0.07; 0.06]* | *-0.08* | *-0.25 [-0.36; -0.12]* | ***-3.65*** | 0.02 [-0.03; 0.08] | 0.70 | *0.29 [0.22; 0.36]* | ***6.81*** | *-0.18 [-0.24; -0.13]* | ***-5.18*** | *-0.27 [-0.38; -0.16]* | ***-3.99*** | *0.09 [0.03; 0.15]* | ***2.49*** |
| C | 0.15 [0.09; 0.21] | **3.78** | 0.15 [0.07; 0.23] | **3.23** | *-0.03 [-0.16; 0.10]* | *-0.39* | -0.03 [-0.10; 0.03] | -0.83 | 0.13 [0.04; 0.22] | **2.29** | *-0.09 [-0.17; -0.02]* | ***-2.11*** | *-0.10 [-0.24; 0.05]* | *-1.19* | 0.04 [-0.04; 0.10] | 0.81 |
| O | 0.08 [0.02; 0.15] | **1.91** | 0.15 [0.07; 0.22] | **3.30** | *-0.07 [-0.21; 0.05]* | *-0.86* | -0.05 [-0.11; 0.02] | -1.32 | *0.25 [0.16; 0.34]* | ***4.52*** | *-0.11 [-0.17; -0.03]* | ***-2.44*** | *-0.06 [-0.19; 0.08]* | *-0.74* | 0.07 [-0.01; 0.15] | 1.53 |
| D* | -0.04 [-0.09; 0.02] | -1.15 | -0.05 [-0.11; 0.02] | -1.28 | -0.09 [-0.20; 0.02] | -1.42 | -0.02 [-0.07; 0.03] | -0.67 | 0.20 [0.13; 0.28] | **4.31** | -0.12 [-0.18; -0.05] | -3.24 | -0.06 [-0.18; 0.04] | -0.83 | 0.04 [-0.01; 0.10] | 1.15 |
| *Model 2: Predicting Behavior from Situation Perception at level 1 and level 2, and Happiness at level 1 and level 2* | | | | | | | | | | | | | | | | |
| H | 0.06 [-0.01; 0.12, | 1.57 | 0.06 [0.00; 0.13] | 1.50 | *-0.09 [-0.22; 0.04]* | *-1.23* | -0.03 [-0.09; 0.03] | -0.90 | *0.10 [0.01; 0.20]* | ***1.76*** | *-0.05 [-0.11; 0.03]* | *-1.02* | *-0.16 [-0.31; -0.02]* | ***-2.04*** | 0.05 [-0.02; 0.12] | 1.30 |
| E | 0.09 [0.02; 0.16] | **2.02** | *0.08 [0.00; 0.16]* | ***1.66*** | *0.21 [0.06; 0.35]* | ***2.35*** | 0.12 [0.05; 0.18] | **2.93** | *-0.07 [-0.18; 0.04]* | ***-1.08*** | *0.37 [0.29; 0.46]* | ***7.52*** | *0.23 [0.07; 0.39]* | ***2.47*** | 0.15 [0.08; 0.23] | **3.23** |
| X | -0.07 [-0.13; -0.02] | **-2.18** | *0.00 [-0.06; 0.06]* | *-0.02* | 0.06 [-0.06; 0.17] | 0.80 | 0.03 [-0.02; 0.09] | 1.06 | *0.24 [0.17; 0.33]* | ***4.79*** | *-0.11 [-0.18; -0.05]* | ***-2.72*** | *-0.03 [-0.16; 0.08]* | *-0.43* | *0.18 [0.12; 0.23]* | ***4.89*** |
| A | -0.01 [-0.06; 0.04] | -0.33 | *0.01 [-0.05; 0.07]* | *0.39* | *-0.21 [-0.31; -0.12]* | ***-3.41*** | *0.02 [-0.02; 0.07]* | *0.73* | *0.22 [0.14; 0.29]* | ***4.74*** | *-0.12 [-0.19; -0.06]* | ***-3.45*** | *-0.24 [-0.36; -0.12]* | ***-3.74*** | 0.06 [0.00; 0.12] | **1.69** |
| C | 0.16 [0.09; 0.22] | **4.14** | 0.17 [0.10; 0.24] | **3.86** | *0.06 [-0.07; 0.20]* | *0.72* | -0.04 [-0.09; 0.03] | -0.95 | 0.03 [-0.07; 0.13] | 0.47 | *-0.02 [-0.10; 0.06]* | ***-0.36*** | *-0.02 [-0.17; 0.12]* | *-0.21* | 0.01 [-0.06; 0.07] | 0.14 |
| O | 0.09 [0.03; 0.15] | **2.25** | 0.18 [0.11; 0.25] | **4.18** | *0.01 [-0.13; 0.14]* | *0.08* | *-0.05 [-0.12; 0.02]* | *-1.35* | *0.20[0.1; 0.3]* | ***3.34*** | *-0.04 [-0.12; 0.04]* | ***-0.87*** | *0.00 [-0.14; 0.14]* | *-0.01* | 0.04 [-0.02; 0.12] | 1.01 |
| D* | -0.03 [-0.09; 0.02] | -1.04 | -0.04 [-0.11; 0.02] | -1.17 | -0.07 [-0.19; 0.04] | -1.07 | -0.02 [-0.07; 0.04] | -0.68 | 0.19 [0.11; 0.27] | 3.98 | -0.11 [-0.17; -0.04] | -2.79 | -0.04 [-0.16; 0.07] | -0.53 | 0.03 [-0.02; 0.09] | 0.98 |
| *Model 3: Predicting Behavior from Situation Perception on level 1 and level 2, Happiness at level 1 and level 2, and PA and NA at level 1 and level 2* | | | | | | | | | | | | | | | | |
| H | 0.08 [0.02; 0.14] | **2.34** | 0.10 [0.04; 0.17] | **2.38** | *0.01 [-0.13; 0.14]* | *0.09* | -0.02 [-0.08; 0.04] | -0.57 | *0.09 [-0.05; 0.21]* | *1.10* | *0.04 [-0.05; 0.13]* | *0.84* | *-0.08 [-0.23; 0.08]* | *-0.91* | 0.07 [-0.01; 0.15] | **1.65** |
| E | 0.03 [-0.04; 0.10] | 0.73 | *0.01 [-0.08; 0.09]* | *0.15* | *-0.04 [-0.19; 0.12]* | *-0.44* | 0.10 [0.03; 0.17] | **2.46** | *-0.02 [-0.16; 0.13]* | *-0.20* | *0.28 [0.18; 0.37]* | ***4.81*** | *0.00 [-0.16; 0.16]* | *0.04* | 0.12 [0.05; 0.2] | **2.58** |
| X | -0.05 [-0.10; 0.01] | -1.42 | *0.01 [-0.06; 0.07]* | *0.17* | 0.11 [-0.02; 0.24] | 1.49 | 0.03 [-0.02; 0.08] | 0.96 | *0.12 [0.01; 0.24]* | ***1.67*** | *-0.04 [-0.11; 0.04]* | *-0.88* | *-0.01 [-0.14; 0.11]* | *-0.11* | *0.17 [0.11; 0.23]* | ***4.60*** |
| A | 0.02 [-0.03; 0.07] | 0.79 | *0.04 [-0.01; 0.10]* | *1.24* | *-0.10 [-0.22; 0.02]* | *-1.41* | *0.03 [-0.02; 0.07]* | *0.97* | *0.22 [0.12; 0.33]* | ***3.56*** | *-0.02 [-0.08; 0.06]* | *-0.39* | *-0.15 [-0.27; -0.03]* | ***-2.14*** | 0.06 [0.01; 0.12] | **1.88** |
| C | 0.18 [0.11; 0.24] | **4.54** | 0.19 [0.11; 0.27] | **4.18** | *0.08 [-0.07; 0.25]* | *0.86* | -0.04 [-0.11; 0.03] | -0.98 | -0.05 [-0.19; 0.1] | -0.55 | *0.01 [-0.08; 0.11]* | *0.13* | *-0.03 [-0.21; 0.14]* | *-0.31* | -0.01 [-0.08; 0.07] | -0.11 |
| O | 0.12 [0.06; 0.19] | **3.27** | 0.20 [0.14; 0.28] | **4.83** | *0.09 [-0.05; 0.24]* | *1.01* | *-0.06 [-0.12; 0.00]* | *-1.63* | *0.08 [-0.05; 0.22]* | *0.98* | *0.07 [-0.02; 0.16]* | *1.22* | *0.06 [-0.09; 0.21]* | *0.70* | 0.03 [-0.04; 0.1] | 0.57 |
| D* | 0.00 [-0.06; 0.05] | -0.09 | -0.03 [-0.09; 0.03] | -0.82 | -0.01 [-0.13; 0.09] | -0.18 | -0.03 [-0.08; 0.02] | -0.81 | 0.15 [0.04; 0.27] | **2.22** | -0.07 [-0.15; 0.01] | -1.45 | 0.04 [-0.09; 0.17] | 0.51 | 0.03 [-0.03; 0.09] | 0.70 |

*Note.* Sit.-Perc. = situational perception (predictor level 1): *b* = estimate (unstandardized) effect; LL = lower limit of 90% CI, UL = upper limit of 90% CI; *t* = *t*-value of fixed effect. Predicted behaviors: H = Honesty/Humility; E = Emotionality; X = eXtraversion; A = Agreeableness; C = Conscientiousness; O = Openness;
Bold effects are significant. Effects in *italics* were predicted to be significant. No effect was significant in the opposite of the predicted direction.
*The effects for Dominance were not included in the computation of Sensitivity and Specificity and the hypotheses not pre-registered.

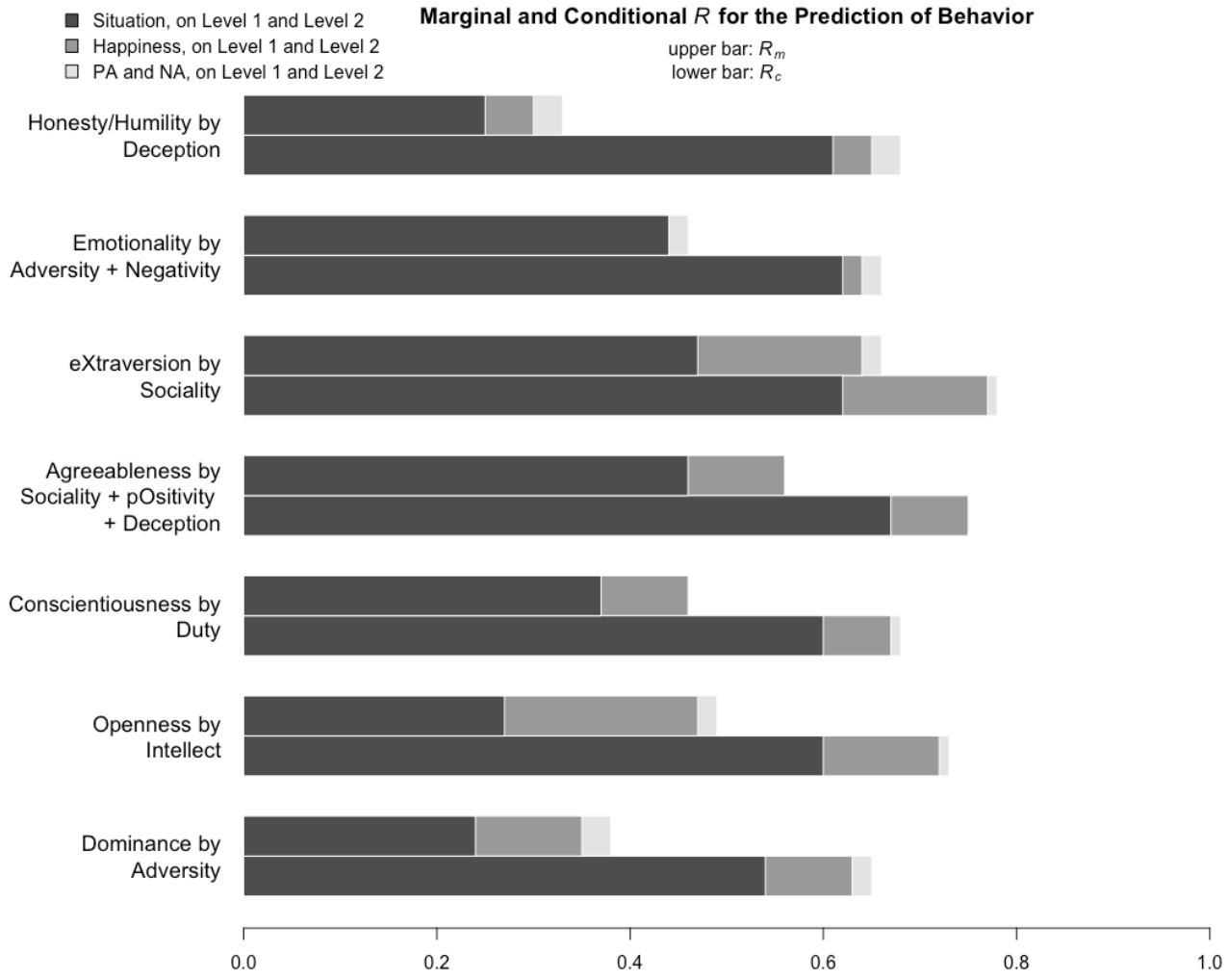
**Replication: Prediction of behavior with several situation perception measures.** Finally, we predicted behavior with several, theoretical meaningful situation perception dimensions at the same time. First, we replicated the original findings from Sherman and colleagues (2015). However, based on the results from Study 1, some of the expectations about effects were updated, and we explicitly expected some effects not to be significant, as can be seen in Table 3.2.8, column *Pred. Effect (Pre-Reg.)*. First, we examined the same models Sherman and colleagues (2015) originally examined. In this direct replication, most predicted effects were significant, and three out of six effects that were assumed not to be significant were also found to not be significant. This resulted in a sensitivity of .87 and a specificity of .50.

Subsequently, happiness was included at level 1 and level 2, reducing both sensitivity to .83 and specificity to .40. At the same time, happiness increased the explained variance in all models. Happiness at level 1 was a significant predictor in all models (evaluated on alpha = .05, two-tailed). Subjective happiness at level 2 significantly predicted sociable behavior, agreeable behavior, conscientious behavior, and open/intellectual behavior.

Finally, positive and negative affect both at level 1 and level 2 were included as additional predictors. Two models (predicting emotional and agreeable behavior) did not converge, and one predictor was removed in each case. The expected effects remained largely unchanged, resulting in an unchanged sensitivity (.83) and slightly decreased specificity (.33; two out of six effects correctly predicted not to be significant). Thus, the effects of situation perception on behavior remained largely unaffected by the additional inclusion of affect and happiness as predictors.

Similar to the previous model, affect and happiness explained substantial amounts of variance: The average $R_m$ was .48 (range .33 to .66), and the average $R_c$ was .70 (range .65 to .78). However, compared to the previous model, this increase in multiple *R* was not large (maximum difference for both indicators was .03), thus suggesting that even if the inclusion of affect at level 2 and level 1 renders the effects of happiness non-significant, no substantial additional variance is explained. This effect is exemplarily presented in Figure 3.2.2.

*Figure 3.2.2.* Marginal and Conditional *R* ($R_m$, $R_c$) for the prediction of behavior with multiple situation perception dimensions at the same time




*Note.* The colors indicate the subsequent expansion of the models. Dark grey: Situation perception (listed on left side) only, medium grey: Happiness included as an additional predictor. Light grey: PA = positive affect, NA = and negative affect included as additional predictors. For exact model parameters, see Table 3.2.8.

**Prediction of behavior with personality traits.** We further pre-registered hypotheses about the effects of personality traits on behavior in the specific models. Specifically, we expected personality traits to predict the corresponding behaviors. As this effect was examined repeatedly in each combination of situation perception and behavior[29], it is most informative to evaluate it in the final models presented in Table 3.2.8. As seen in the last column (Study 2, Extension), personality traits predicted the corresponding behavior in all instances, even if subjective happiness, positive affect, and negative affect were included at level 2. Thus, personality is and remains a viable predictor of behavior in everyday life even when considering additional predictors such as affect and subjective happiness.

**Discussion**

Study 2 consisted of three parts: (1) the replication of the results from Sherman and colleagues (2015); (2) the replication of the re-analysis in Study 1; and (3) the extension of the analyses with better and separate measures of positive and negative affect. First, replicating the examination of the sources of behavior (the person or the situation) did not show a similarly clear pattern as Sherman and colleagues' (2015) original study. We computed ICCs for behavior, and a high ICC means that variance of behavior must be attributed to the person rather than the current occasion. As the ICCs of situation characteristics were similar to those of affect and behavior, it cannot be concluded that behavior or affect influenced more by the person than the perception of situations. However, it is important to note that all in situ measures – behavior, affect, and situation perception – did show substantial between- and within-person variance. As suggested by Baumert and colleagues (2017), we subsequently investigated this variance simultaneously instead of aggregating state measures and considering only between-person variance. As the current study showed, analyzing both within- as well as between-person variability provided a better and more nuanced picture of the relation of a person's characteristics, behaviors, feelings, and perceptions.

Second, we replicated our initial finding that nearly all measures of situation perception predicted nearly all forms of behavior. Further, our predictions were more accurate after the inclusion of happiness and affect. However, it is important to distinguish between level 1 and level 2 effects of situation perception: Whereas the predictions for level 1 situation perception were mainly accurate, the predictions for level 2 situation perception were comparatively poor. However, as the replication from Sherman and colleagues' (2015) original models showed (Study 2, Table 3.2.8, Extension), some mean situation variables indeed predicted behavior in meaningful instances. In six out of seven models, mean situation perception predicted behavior in daily life. Thus, we conclude that mean situation perception can also be considered a stable person-

[29] These effects are not presented in the current article but can be found in the corresponding OSM D.

characteristic (Horstmann, Ziegler, et al., 2018; Ziegler & Horstmann, 2015; Ziegler et al., submitted), but its relation to behavior in the context of experience sampling is not yet well understood.

Finally, we replicated and extended the models examined by Sherman and colleagues (2015), predicting behavior with multiple situation characteristics as well as happiness and affect. These analyses showed that although situation perception and affect/happiness were substantially correlated (Horstmann & Ziegler, 2018), both predicted unique shares of variance in behavior. Furthermore, these analyses highlighted the importance of considering in situ affect as an additional predictor of behavior, especially when examining effects of situation perception. Comparing the effects of affect measures with happiness measures revealed a possible remedy to the problematic overlap of situation perception and affect: The two three-item scales for positive and negative affect, which were assessed at each measurement occasion, did not substantially outperform the single item measure of state happiness and a four-item measure of trait happiness (Lyubomirsky & Lepper, 1999; Swami et al., 2009) in terms of explained variance (see Figure 3.2.2) Thus, a one-item happiness measure could be included as a good proxy for within-person fluctuations of affect, and the four-item trait measure could be used to account for between-person differences in affect.

**Limitations**

Both studies are subject to several limitations, which will be discussed below (see General Discussion, Limitations). However, some limitations are specific to Study 2. First, we used a measure of happiness as well as bipolar rating scales that were translated from English, but not validated for use in experience sampling in a German sample. Nonetheless, as indicated by the correlations between trait-scores and averaged state-scores (mean situation perception and mean happiness), scores from these measures showed good convergent and discriminant validity.

Second, we determined the a priori sample size at level 2 ($N \geq 250$ participants) and level 1 ($N \geq 45$ measurements). Although many participants completed 50 measurements, some participants completed fewer. This leads to a loss of test power for level 1 fixed effects and could explain why some of the effects that were expected to be significant were not.

Finally, we deviated from the pre-registration in some minor aspects. Participants that were slow in their responses during experience sampling were not excluded as this would have meant that more than 440 measurement occasions had to be deleted. Further, the codebook (see pre-registration) falsely stated that some items from the experience sampling phase needed not to be recoded. However, to align the scores on level 1 with the scores on level 2 (e.g., a high value in both cases means high conscientiousness), the items were recoded. The pre-registration further

stated that measures of happiness should be replaced with measures of positive affect in the extended analyses – however, we did not replace the measure of happiness with positive and negative affect, but extended the models and additionally included the measures for positive and negative affect. Also, one equation in the pre-registration stated that the mean affect measures are not considered in the models, which was wrong. In line with general recommendations (Enders & Tofighi, 2007) these variables were included as level 2 predictors to account for between-person variance. Importantly, we deviated from the pre-registration in these aspects not because "they did not work" – we changed them simply because they were wrong. We therefore think that the pre-registration is still valid, and Study 2 can be considered strictly confirmatory.

## General Discussion

Across two independent studies, behavior could be predicted with situation perception, affect, happiness, and personality traits. Even though affect and situation perception were correlated and shared variance with behavior in a situation, both contributed uniquely to explaining intra-individual differences in behavior. Accounting for shared variance between affect or happiness and situation perception revealed that the latter was related to behavior in logically coherent and predictable ways. These findings have important theoretical consequences for understanding and examining situation perception and its effects.

### Overlap of Affect and Situation Perception

Affect and situation perception were related in the current study, and this finding is in line with previous literature (Edwards & Templeton, 2005; Gerpott et al., 2017; Horstmann & Ziegler, 2018; Parrigon et al., 2017). This corroborates the simple process model of situation perception, suggested by Rauthmann, Sherman, and Funder (2015b). Rauthmann and colleagues (2015) suggested that situation perception is a product of person aspects (e.g., personality traits, social roles, affect) and situation cues. The current study showed that affect, situation perception, and personality traits are indeed related in a meaningful way among each other, but also predict unique shares of variance in behavior. More importantly, we showed that situation perception is *more* than just affect or happiness; situation perceptions shared systematic variance with behavior after controlling for happiness and affect. As Rauthmann and colleagues (2015) suggested, this effect should be attributable to situation cues. A participants' rating of a situation and the rating of his or her current affect were differentiable. Thus, participants' ratings distinguished between situational information (e.g., "the situation is pleasant") and their personal affect (e.g., "I am in a bad mood"). Under the assumption that all relevant person parameters could be accounted for, effects of situation perception should thus reflect situation cues and their individual processing. Controlling for individual person parameters – especially idiosyncratic person states – should, for

example, increase the consensus different perceivers reach on the characteristics of a certain situation. Similarly, but only true under the assumption that affect influences situation perception (as suggested by Rauthmann et al., 2015), the perception of a situation should be more similar between two perceivers the more similar their affective states are.

The second finding of our study was that nearly all dimensions of situation perception were tied to behavior, but that after controlling for affect, specific predicted combinations of situation characteristics and behavior emerged. This finding supports the *discriminant criterion validity* of situation perception measures. For example, after controlling for affect, perceiving a situation as intellectual was related to open behavior, but not to agreeable behavior.

Together, our findings suggest that situation perception is thus not only a powerful, but also logically coherent predictor of behavior. Wilson, Thompson, and Vazire (2017) asked if "fluctuations in personality states [are] more than just fluctuations in affect?" (p. 110). The authors came to the conclusion that situations might account for additional variance after controlling for positive and negative affect for some behaviors. The current study showed that situation characteristics accounted for variance in behavior, even after controlling for positive and negative affect on level 1 and level 2. Of course, it may be possible that by accounting for further person-states (e.g., fatigue, hunger, excitement, or motivation; Rauthmann, 2016), or by considering the arousal component of affective states, effects of situation characteristics will further be reduced. Such a reduction would not be problematic as affect and situation perception are expected to be related (Horstmann & Ziegler, 2018; Kuppens, 2009). However, if effects of situation perception on behavior were indeed *fully* accounted for by other affective states, then this would mean that (a) situation cues were not reflected in measures of situation perception, (b) situation cues did not affect behavior directly, and/or (c) the effects of cues on behavior were fully mediated by affective processes. However, our findings suggest that situation perceptions cannot be reduced to affect only.

**Happiness and Affect as Predictors of Behavior**

Even though the main purpose of the study was the investigation of situation perception, we also gained important insights into the role of happiness and affect as predictors of behavior. As mentioned earlier, a plethora of studies has already shown that emotions and affect explain behavior (e.g., Lench et al., 2011; Wilson et al., 2016; Wilt & Revelle, 2017), and the results presented here further corroborate this. Affect at the person level explained variance in several behaviors across Studies 1 and 2. Nevertheless, the effects of the corresponding personality traits remained substantial (in Study 2), and affect explained additional variance previously unaccounted for. For example, in Study 2, happiness predicted conscientious, agreeable, sociable, and open

behavior. It is possible that, as most participants were enrolled in a university during participation (90.50% enrolled during participation, 5.84% were at some time enrolled, and only 3.65% were never enrolled in a university), behaving conscientiously, agreeably, openly, and sociably leads to success (Ziegler et al., 2010) and subsequently to more positive affect. Unfortunately, it is not possible to investigate this or any other matters related to causal forces with the current data. Thus, which behaviors affect predicts most across more heterogeneous groups of people remains a question to be addressed by future studies specifically interested in the predictive power of affect on behavior.

**Situation Perception, Affect, and Personality as Predictors of Behavior**

We examined how behavior could be predicted with a combination of affect, situation perception, and personality traits. First, it has to be noted that prior assessed personality traits *predicted* future behavior during later experience sampling. This supports the idea of traits as density distributions of states, where the mean of the state distribution corresponds roughly to the level of the (self-reported) trait (Fleeson, 2001; Fleeson & Gallagher, 2009; Fleeson & Jayawickreme, 2015; Rauthmann et al., 2018). Variance unaccounted for by the personality trait must thus be explained by other traits (or the same trait assessed via different measures if there is method variance), situational influences, or current affective states of the person.

Sherman and colleagues (2015) showed that "[b]oth personality traits and situation characteristics *independently* predicted real-time state expressions" ( p. 884). After extending and replicating their study, we conclude that personality, situation characteristics, *and* happiness or affect independently predict behavior. Note that affect and happiness do not independently predict behavior as the inclusion of the latter mostly reduced the effect of the former. Similar to the study by Sherman and colleagues (2015), we found that both level 1 as well as level 2 effects of situation perception predicted behavior. However, as outlined above, we were not successful in predicting which effects of mean situation perception would be significant once controlling for affect at level 2. Further, the inclusion of mean affect oftentimes reduced effects of mean situation perception. Thus, the effect of mean situation perception needs to be examined in more detail in future research. One possible avenue would be to ask participants about their "general situation experience" (Rauthmann et al., 2018) or to use a set of standardized situational vignettes to assess people's trait-like situational perception tendencies (Horstmann & Ziegler, 2018; Ziegler et al., submitted). It will be an important next step to distinguish between stable characteristics of participants' environments and stable tendencies of a person to perceive their situations, and their independent role in jointly explaining and predicting behavior (Rauthmann et al., 2018).

**Implications for Personality Theories**

Sherman and colleagues (2015) discussed implications for several trait theories of personality. The current study does not alter the conclusions of Sherman and colleagues' original study but corroborates them and adds more nuance. Situations and their perceptions have a place in all personality theories (e.g., Fleeson & Jayawickreme, 2015; McCrae & Costa, 2008; Mischel & Shoda, 1995), the exception being trait activation theory (Tett & Guterman, 2000) which calls explicitly for person-situation interactions (Sherman et al., 2015). However, even main effects of a situation that remain after controlling for personality traits can be considered an interaction with the situation as they technically explain changes in rank-ordering among the participants that cannot be accounted for by personality traits alone (Horstmann, Ziegler, et al., 2018).

Furthermore, we additionally referred to Latent State-Trait theory as another framework for the conceptualization of personality traits and states (Steyer et al., 1992, 2015, 1999). The current work shows that variance unaccounted for by personality traits can be explained by occasion-specific person-aspects (states), and not only by situation characteristics. Similarly, Read and colleagues (2017) recently presented a model for personality that reconciles stable trait and variable state aspects of personality by introducing temporal "motive affordances." These motive affordances lead to certain behaviors in a situation due to current needs of the person. For example, if a person has just spent a lot of time talking due to situational demands of sociality, they might feel the need to talk less and therefore seek a situation with less sociality or change the current situation so that it contains less sociality (Rauthmann & Sherman, 2016b). One could argue that affect is a suitable proxy for motive affordances (de Vries et al., 2016), guiding the behavior and giving a person an indication as towards the current satisfaction of needs (Baumeister, Vohs, DeWall, & Zhang, 2007; Giner-Sorolla, Kupfer, & Sabo, 2018). This means that affect could be considered a situationally influenced link between stable person-aspects (personality traits, including needs), and variable behaviors (personality states). To clarify this, future studies could employ experimental approaches to situation manipulation and examine the temporal sequence of situation perception, affect, and behavior.

Finally, much research on situational information and influences was initially triggered by the person-situation debate (Donnellan et al., 2009; Funder, 2006). Initially, the observation that personality traits explained only a limited amount of intra-individual variance in behavior (Mischel, 1968) led to the search for other viable predictors for behavior. Most researchers would now agree that both *stable* characteristics of the person as well as *fleeting* situation factors shape behavior (Funder, 2001; Mischel et al., 2002; Sherman et al., 2010). Consistency can therefore be established by examining persons in certain situations: *if* Person X is in Situation Y, *then* he/she behaves

extraverted (Mischel & Shoda, 1995; Shoda et al., 1994). However, given that momentary affect in a situation predicts behavior *independently* from situation characteristics and traits, momentary affect should be recognized as well in this approach: *if* Person X with his/her affective state A is in Situation Y, *then* he/she behaves extraverted. The current work highlights the importance of conceptualizing the person-side in a much broader way as it was initially intended (Cattell, 1963; Horstmann, Ziegler, et al., 2018; Lewin, 1936; Rauthmann, Sherman, & Funder, 2015b). If the search for behavioral consistency is to be continued, affective states of a person in a situation should also be considered explicitly in conceptualizations and tests of personality theories (Baumert et al., 2017).

**The Role of Situation Perception for Appraisal Theories of Emotion**

The overlap of situation perception and affect has previously been discussed as a direct consequence of appraisal theories of emotion (Kuppens, 2009, 2015; Kuppens & Van Mechelen, 2007; Sander et al., 2005; Scherer, 2001). The results of the current study support this idea and call for a more thorough investigation of appraisal mechanisms in situation perception. As suggested by Rauthmann and colleagues (Rauthmann, 2012; Rauthmann, Sherman, & Funder, 2015b), situation perception is a result of appraisal mechanisms of cues and current affect, whereas appraisal theories of emotions state that affect results from the appraisal of situation cues (Giner-Sorolla et al., 2018; Kuppens, 2009; Sander et al., 2005).

Related to this issue is the question of causality of the effects and their chronological order. Affect or emotions may shape behavior in two different ways. First, a certain emotion leads directly to a certain behavior. Second, emotions serve as feedback for certain behaviors and thereby incentivize or inhibit behavior indirectly (Baumeister, Vohs, DeWall, et al., 2007). A person may select and thus perceive certain situations because they have previously led to positive feelings, or they may search for certain situations because they already experience a positive affect and would like to behave accordingly. Measuring situation perception as well as situation selection will allow investigating these processes on a more fine-grained level.

**The Examination of Personality States and Traits**

Baumert and colleagues (Baumert et al., 2017) called for the integration of personality structure, processes, and development, with the specific examination of "measures of cognitive, affective, motivational, and behavioral states under specified situational conditions" (p. 517). To examine psychological *processes*, these measures will need to be assessed repeatedly. The pre-registration of the research questions and the methods and analyses of the current study present a starting point for such an endeavor. Furthermore, as both data-sets from Study 1 and Study 2 are openly available online, researches are invited to explore and test further hypotheses.

When examining repeated person states, researchers face the challenge of selecting a number of constructs and items while at the same time keeping participant burden at a minimum. Based on the results of the current study, we suggest that at least one item (e.g., a bipolar rating scale) of happiness should be included in future experience sampling studies. As the correlational structure of situation characteristics has been examined extensively, future studies may also consider examining only one or two dimensions of situation perception in more detail by replacing other dimensions with measures of situation management (Rauthmann & Sherman, 2016b) or situational affordances (N. A. Brown et al., 2015). Alternatively, of course, more refined measures of behavior, affect, and emotions could also be used.

**Limitations**

Similar to the limitations in Sherman and colleagues' (2015) original study, the current study used a convenience sample of undergraduate students, thus limiting generalizability. At the same time, though, this enhances the closeness of the replication and allows for a better comparison between Study 1 and Study 2 (LeBel et al., 2017). Nevertheless, future studies should investigate if the results can be generalized across different person groups. The sample of participants also affects the sample of situations. Most of the situations that contained Duty (e.g., studying) also contained Intellect (e.g., thinking is required). Although these two dimensions are in principle independent (i.e., it is possible to imagine a situation that contains high Intellect, but little Duty, and vice versa), they were highly correlated in Study 1. Thus, for Study 2, we expected Intellect to be also related to conscientiousness behavior. This reflects the expectations we had about the sample that we were going to collect. However – and this needs to be addressed in future research – the specifics of the sample that led us to update these hypotheses are not explicitly reflected in any other quality than "being university students." Investigating such moderator variables may be considered in future research examining person-situation-behavior contingencies. Additionally, university students (and participants in general) may only experience or report certain situations (Horstmann, Ziegler, et al., 2018), which also threatens the generalizability of the findings.

Finally, both studies relied on self-reports. Despite all evidence available pointing towards validity of these assessments, no external criteria were available for external validation (e.g., peer-reports, grades, time spent studying). This is, for example, problematic with respect to the behavioral measures as they might contain wishful thinking ("how would I like to behave?") or socially desirable responding. Future studies assessing behavior, emotions, and situations via other methods (e.g., smartphone or laptop use, physical arousal, geo-spatial location) will be very beneficial and advance the field in many ways.

**Recommendations**

The results of the study bear several implications for the examination of situation perception, situation taxonomies, the examination of affect in daily life, and personality and social psychology more generally. Below, we compile some areas that can benefit from our findings.

**Considering Null-Effects and Replication.** Based on the exploratory re-analyses of Sherman and colleagues' data, we discovered that the proposed links between situation perception and behavior were not exclusive (i.e., there were many statistically significant unpredicted correlations). To this end, we examined effects that were not considered in the first place (see Table 3.2.1) but were implicitly assumed to be not present. Similar to the idea of Multi-Trait Multi-Method analyses (Campbell & Fiske, 1959) where discriminant correlations should be low and not significant, considering and examining effects that were assumed not to be significant has been insightful. We would thus encourage future investigations of person-situation relations to explicitly test those effects that "do not make sense."

Even though Study 1 and Study 2 were similar in design, several effects were significant in Study 1 but not in Study 2, and vice versa. It is impossible to state which one of both cases is actually the false-positive (or false-negative), or if any unknown variables exist that explain these differences in study outcome. Theses specific results should therefore be interpreted cautiously, and future (pre-registered) replications would be helpful.

**Assessing Situational Information**. As Horstmann and Ziegler (2018) pointed out, the overlap of situation perception and affect is problematic when assessing situation characteristics and constructing taxonomies for the description of situational information. The shared affective variance in situation characteristics will be reflected in factor structures during the construction of situation perception taxonomies, similar to the effect of common method variance in the construction of personality questionnaires (Bäckström, 2007; Bäckström et al., 2009; Podsakoff et al., 2003; Ziegler & Buehner, 2009).

Further, during the assessment of situation characteristics, the variance of affect or happiness – an aspect of the person – will be included in the situation perception measure, which

should be primarily an aspect of the situation, thereby further conflating person- and situation variance (Rauthmann & Sherman, 2018b; Rauthmann, Sherman, & Funder, 2015b). This effect can be taken care of by triangulating the situation from different perspectives, specifically raters in situ, raters juxta situm, and raters ex situ: Variance due to affect will be cancelled out by using multiple raters to assess the same situation (Horstmann, Ziegler, et al., 2018), yielding better estimates of the true nature of the situation.

**Summary and Conclusion**

What happens when a person reports their behavior in a current situation, and which other psychological variables can help explain their behavior? Based on the two studies presented here, several processes seem to take place at the same time – some at the enduring level of persons (i.e., trait-like aspects) and some at the momentary level (i.e., state-like aspects).

First, persons act in accordance with their personality traits. Second, and independently from personality traits, a person's general happiness explains variability in certain behaviors. The happier a person, the more sociable, conscientious, and open the reported behavior within a certain time-span may be. Similar to the effects of happiness, mean positive and negative affect (from repeated measurements) also predict behavior, yet this effect is not substantially different from the effects of trait happiness. Finally, mean situation perception explains behavior. Persons may possess general tendencies to perceive and interpret situations, which are reflected their behavior. As Horstmann, Ziegler, and Ziegler (2018; 2015; submitted) point out, assessing these tendencies – which are independent from "classic" (Big Five/Six) personality traits – may open up new avenues for personality assessments and the prediction of relevant outcomes.

At the level of time-variant person aspects, there are also processes to consider. First, happiness and affect exhibited comparatively strong effects on behavior. The happier a person in a situation, the more of "the good" behaviors were displayed; being happy in a situation corresponded to behaving more honest/humble, less neurotic, more social/extraverted, more agreeable, more conscientious, more open, and also more dominant, and similar effects were found for affect. Yet, and despite these substantial effects of happiness and affect, the perception of situation characteristics explained behavior independently from it. A person thus responded not only based on internal, current states, but also according to situation characteristics. It is striking that these interpretable effects occurred despite the short and rapid assessments of behavior. This indicates that these processes are at play even if a person does not consciously elaborate and examine a situation in all details. Looking only at the initial results from Study 1 and seeing that all situation characteristics were related to nearly all behaviors, one could have concluded that the situation measures do not pick up on specific situational information. However, and quite to the

contrary, the analyses and results throughout Studies 1 and 2 paint a different, clearer picture: Situation perception, affect, happiness, and personality traits can be seen as sufficiently distinct predictors of behavior, each contributing uniquely to the explanation of it. Only the holistic view presented here allowed examining all aspects of within- and between-person variability (Baumert et al., 2017) and unveiling the entire picture. It remains an open question, though, *how* a person is capable of making these quick, distinct, and theoretically meaningful judgments in any given situation, and which processes guide these evaluations. For now, we can conclude *that* all of these aspects presented here play an important role in the explanation of human behavior in daily life.

## 3.3 Conclusion Part 3

Affect and dimensions of situation perception are correlated, yet both predict unique shares of variance in experience sampling studies (Horstmann et al., submitted; Horstmann & Ziegler, 2018). Apparently, the evaluation of situations and current affect are distinct, but both are systematically tied to behavior. The principles of situation research (see Part 2, section 2.1) provide initial explanations for this finding.

First, a person's affect should not be used as a situation variable. Of course, a person may experience a certain affect before, during, and after a situation. However, perceiving and interpreting situation cues, and experiencing affect are different processes, and both can guide behavior: Imagine a person that experiences severe negative affect but finds him or herself in a surprisingly nice situation. Negative affect may reduce how agreeable this person behaves (Horstmann et al., submitted; Wilson et al., 2017; Wilt et al., 2017; Wilt & Revelle, 2017). However, the perception of the situation as positive may lead to a little bit more agreeable behavior compared a person who would experience comparable levels of negative affect in a neutral situation or even negative situation.

Second, situation ratings contain shared or consensual variance (see agreement corollary) and idiosyncratic variance. The consensual reality reflects what most perceivers would agree upon when rating or perceiving a situation. Consider a person experiencing very negative affect, taking the subway back home from work. If this person was asked how negative the current situation is, he or she might think "Well, I am in a pretty negative mood, but I know that this situation right here is actually not too bad." Because a person knows that taking a subway home is generally not perceived as a very negative situation (consensual reality), this could be reflected in its rating. This theory, however, demands that affect overlaps particularly with the idiosyncratic variance in situation ratings. This can be tested in future studies with multiple raters of the same situation. I will review this distinction in more detail below. For now, it is important to conclude that situation perception and affect both account for unique shares of variance in behavior in daily life.

# 4 Part 4: Implications and Conclusion

The work compiled in this dissertation bears several implications for the description and the assessment of situations, the application of situational taxonomies, the conceptualization of personality theories, the prediction of behavior, and for better understanding consistency. Although these topics are of course related, implications will be reviewed separately. I conclude each section with specific recommendations for future research.

## 4.1 Status Quo of Situation Research

Situation research has received a lot of attention in recent years. This research includes general recommendations for the description of situations (Rauthmann & Sherman, 2018b; Rauthmann, Sherman, & Funder, 2015b) as well as the assessment of situational information (Horstmann, Rauthmann, et al., 2018; Horstmann, Ziegler, et al., 2018; Rauthmann, 2015). Although several situational taxonomies had been developed previously, they were only recently applied in personality psychology, most likely due to the availability of specific assessment tools for situation perception. This widespread application of taxonomies will help shaping the nomological net of situation perception dimensions. At the same time, the current surge in situational taxonomies calls for an integration: As we have suggested, it is likely that situation dimensions possess an overarching structure, which will make future research on situations more comparable. Yet, this is not an easy task and requires probably multiple studies and refinements over a longer period of time (Horstmann, Rauthmann, et al., 2018; Rauthmann & Sherman, 2018a).

Although considerable progress has been made on the description of situation characteristics, this is not the case for the description of situation cues or situation classes. As outlined above, research on situation cues and their psychological relevance will most likely benefit from technical innovations and the increased ability to collect and analyze larger data sets. Examples include pictures taken in daily life (N. A. Brown et al., 2017) or the electronically activated recorder (EAR), which samples audio-snippets in daily life (Holtzman, Vazire, & Mehl, 2010; Matthias R. Mehl, Pennebaker, Crow, Dabbs, & Price, 2001; Matthias R. Mehl, Robbins, & Deters, 2012). Once these data sets have been coded and a meaning has been assigned manually to situation cues, it might be possible to extract psychological meaning from ratings of situation cues that apply similarly to most persons. Situation cues could thus be made available for psychological research.

Concerning situation classes, several taxonomies have already been suggested but no conclusive taxonomy exists so far (Horstmann, Rauthmann, et al., 2018). Given that situation

classes were successfully used to explain variance in behavior (e.g., Geukes et al., 2017; Oud et al., 2017), they seem to contain psychologically relevant information. Further research could examine if situation classes can be extracted that hold similar or even more information than situation characteristics. Situation classes are appealing as they reflect the way situations are usually conceptualized by lay people (e.g., party, home, work) and may therefore be easier to assess. Furthermore, if situations could be assigned to meaningful classes, this would make an application of situation research easier. As an example, it could be possible to group certain job situations in similar classes (e.g., manual work, work at desks, meetings, work outside, etc.) that have similar demands and therefore suit the needs and personality profiles of specific persons.

Overall, it can be concluded that situation research is in a promising, healthy state. However, as considerable progress is achieved in some areas of situation research (e.g., development of taxonomies, application of situation characteristics), other important areas should not be overlooked (e.g., measurement issues, definition of situation classes) and should receive further attention in future research.

## 4.2 Assessing Situations via Situation Perceptions

Given that the perception of a situation correlates with state affect, state affect should be considered when examining situations via situation perceptions. As the componentiality corollary (Rauthmann, Sherman, & Funder, 2015b) states, perceptions of situations consist of person, situation, and person x situation variance, and each share in variance may be accounted for by different predictors.

The person-variance in situation perception may best be accounted for by stable characteristics of the person, for example personality traits, as it is a function of the person across several situations. The situation and the person x situation variance may best be accounted for by variable states of the person, such as affect and situation perception. By definition, situation perception should best account for situation variance. Affect, which is a characteristic of the person *in* the situation, should best account for the person x situation variance. Note that "situation x person" must not be understood as a purely statistical interaction term of personality traits x situation characteristics, but rather as the person *in* the situation, similar to the conceptualization of the latent state residual in latent state-trait theory. Hence, person x situation variance may best be accounted for by state affect or other characteristics of the person *in* the situation. This idea is elaborated in more detail in section 4.4.2.

This reasoning further supports the approximation corollary: If multiple raters rate the same situation, it can be assumed that the person variance is unique and specific for each rater.

Thus, this unique variance would be cancelled out if multiple ratings of the same situation were aggregated. This effect may even be stronger if ex situ raters (from outside of the situation) rated accounts of the same situation, for example based on video-tapes. Under the assumption that situation perception does not immediately influence affect similarly in all raters (otherwise, different ex situ raters would be in similar affective states right after the examination of the situational material; but see below for a more detailed explanation of the overlap of affect and situation perception), this would yield more accurate ratings of the situation. Of course, this is only true if the raters' affect can be assumed to be independent from one another. Thus, as Rauthmann, Sherman, and Funder (2015) already argued, assessing a situation from multiple perspectives with multiple raters may give the best picture of the situation.

## 4.3 Application of Situational Taxonomies

Dimensions of situation perception are meaningfully related to behavior and can be used to explain variability in behavior in daily life (Horstmann et al., submitted; Sherman et al., 2015). However, there are several challenges that need to be overcome in future research. Currently, situation perception is mainly used to explain variability in behavior: An in situ measure of situation perception is used to account for behavior in the very same situation. However, the assessment of in situ situation perception does not allow real predictions of behavior, that is, in the future and under unknown circumstances. Yet, there may be several possible ways to address this problem.

First, the trait component of situation perception can be used to predict future behavior (Horstmann, Rauthmann, et al., 2018; Ziegler & Horstmann, 2015; Ziegler et al., submitted). Even if future situations were unknown, it is possible to assess a person's general tendency to perceive situations and thus to make predictions about a persons' most likely behavior in future situations.

Second, understanding how a person acts under certain circumstances can inform the selection of situations for certain persons. If a person was examined during their daily life or at work, a specific pattern could be observed for this particular person, which reflects a situation-contingent pattern of behavior. For example, one could find out how a person acts when perceiving high or low Duty. Additionally, environments (for example occupations) could be described using situational taxonomies. Occupations could then be selected to optimally fit a person's unique situation-contingent pattern of behavior. Note, however, that this builds on the assumption that a specific perception of a situation leads (in a causal way) to behavior. However, causal effects of situation perception on behavior have not yet been examined. Connections between situation perception and behavior that have been reported in previous studies may also have occurred due to situation selection. It is plausible that persons select situations due to their

desire to express certain behaviors or their previous experience with similar situations. Situation perception assessed in experience sampling studies (Horstmann et al., submitted; Sherman et al., 2015) would then merely be a correlate of behavior, and not the antecedent. Although this correlation is the first pre-requisite to establish causality, results from these correlational studies alone should not inform interventions.

Furthermore, it is currently unknown if real – that is, directly observable and not self-reported – behavior can be predicted with situation perception. As Baumeister, Vohs, and Funder (2007) noted, real behavior is rarely assessed in psychology, and research on situation perceptions is no exception. Although studies did show that situations can have effects on real behavior (e.g., Sherman et al., 2010), similar studies using measures of situation perception do not yet exist. Thus, the prediction of real (and not self-reported) behavior is one of the next steps that studies on situation perception should address.

So far, one of the strongest applications for situation perception remains the examination of basic principles of human behavior. Knowing how behavior could be explained opens new research avenues. Although some have called for "prediction over explanation" (Yarkoni & Westfall, 2017), meaning that psychology should be more concerned with predicting behavior rather than explaining it, I would argue that in the case of situation perception, understanding the processes that lead to certain behaviors and establishing sound measures for situations will allow much better predictions of behavior in the future.

## 4.4 Process Model of Situation Perception

One question addressed especially in our last study was how variance in (self-reported) behavior can be described and explained with personality, affect, and situation perception (Horstmann et al., submitted). Rauthmann, Sherman, and Funder (2015b) suggested a model that highlights possible predictors of behavior. In their model, the effect of personality, affect, and situation cues on behavior is mediated via situation perception (see Figure 1, Rauthmann, Sherman, & Funder, 2015b, p. 367). In this model, situation perception is the result of "bottom up and top-down information processing" (Rauthmann, Sherman, & Funder, 2015, p. 367), which is essentially a black box. Situation cues are interpreted and evaluated, but the exact process of this evaluation is not depicted. The results presented here suggest several additions and changes to this model. These changes reflect, in a large part, (1) effects that were examined in the studies presented here, (2) effects that were examined in other studies, and also (3) theoretically assumed effects. The effects are displayed in Figure 8.4.1, and each effect can be understood as follows:

**(1) Effects examined in this dissertation.** Several effects were examined independently or simultaneously during this dissertation. Some of these effects can be seen as pre-requisites for later studies (e.g., the correlation of affect and situation perception, or the effect of trait situation perception on behavior) and others as the litmus test for situation perception (i.e., the combined effect of situation perception and affect on behavior).

**(a)** Effects of stable personality characteristics on behavior. Effects of personality traits and affect have been examined in the literature (e.g., Ozer & Benet-Martínez, 2006; Paunonen & Ashton, 2001). Additionally, effects of personality traits, affect, and situation perception were examined in two studies presented in part 2 and 3 (Horstmann et al., submitted; Ziegler et al., submitted).

**(b)** Effects of state affect on behavior were examined in part 3, Horstmann et al., (Horstmann et al., submitted). Additionally, previous studies also examined and reported effect of affect or emotions on behavior (e.g., Lench et al., 2011; Wilson et al., 2017; Wilt & Revelle, 2017).

**(c)** The correlation or overlap of affect and situation perception was examined in part 3, Horstmann and Ziegler (2018), but also in other studies (e.g., Edwards & Templeton, 2005; Parrigon et al., 2017).

**(d)** Effects of situation perception on behavior were examined in part 3, Horstmann et al., (submitted), and also by Sherman and colleagues (2015) and Parrigon et al. (2017).

**(2) Effects examined in other studies.** Some effects have not been examined in the current dissertation but were extensively examined in previously published studies.

**(e)** Effects of personality traits on situation perception were in part examined in part 3, by Horstmann et al., (submitted), but were first reported by Sherman et al. (2015) and Rauthmann, Sherman, Nave, et al. (2015).

**(f)** Effects of personality traits and trait happiness on affect were, among others, examined by Sherman and colleagues (2015).

**(g)** Effects of behavior (in this case, state extraversion and state neuroticism) on affect were, even in an experimental design, examined and reported by McNiel and Fleeson (2006). Acting extraverted or neurotic led to more positive or negative affect, respectively.

**(h)** How behavior shapes situations was, indirectly, examined by Gosling et al. (Gosling et al., 2002). Gosling and colleagues showed that personality traits can be inferred from perceiving situation cues associated with a person (e.g., a dorm-room of the person). Without effects of a person's behavior on situation cues (h) as well as a link between personality and behavior (a), a person's personality could not have been estimated by simply observing situation cues.

**(3) Assumed effects.** Furthermore, some effects must be assumed for the model to be complete. However, not all possible effects are displayed in the model. For example, it is plausible that personality traits influence the perception and interpretation of situation cues. Although it has been examined how personality traits relate to situation perception, the relation of traits with individual steps of the evaluative process has not been examined.

**(i)** Situation cues are perceived and evaluated. This evaluation is described in appraisal theories of emotion (Sander et al., 2005; Scherer, 2001).

**(j)** The evaluation of situation cues can lead to emotions and affect (Gendolla, 2000; Sander et al., 2005; Scherer, 2001).

**(k)** As I propose below in section 4.4.1, situation perception may reflect the information collected during the evaluative process of situation cues.

**(l)** Other person states are likely to influence behavior as well, such as social roles, motivation, but also hunger, tiredness, etc. (Rauthmann, 2016; Rauthmann, Sherman, & Funder, 2015b).

Some changes were made to the model proposed by Rauthmann, Sherman, and Funder (2015). First, given that personality and affect exhibit effects on behavior after controlling for situation perception, a full mediation of personality and affect via situation perception as suggested by Rauthmann, Sherman, and Funder (2015b) seems unlikely. Instead, situation perception, affect, and personality all exhibit independent effects on behavior and predict unique variance in behavior. This is reflected in Figure 8.4.1. In Figure 8.4.1, all effects that were examined in the current thesis are displayed as solid black arrows, all hypothetical effects are displayed as dashed black arrows. Effects not examined in this dissertation are presented as grey dotted arrows. All graphical elements are explained in Table 8.4.1.

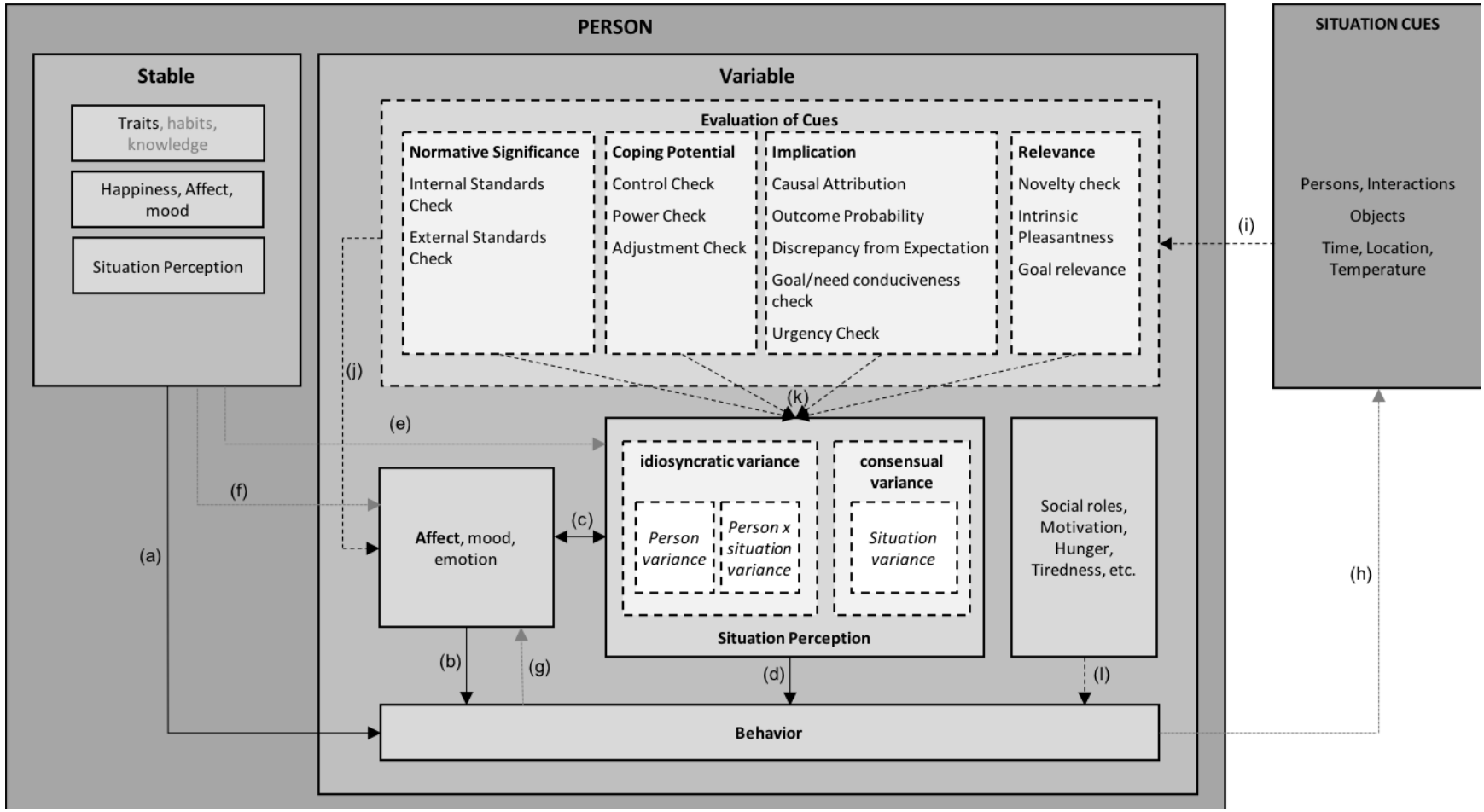
Table 8.4.1

*Explanations for graphical elements presented in Figure 8.4.1*

| **Graphical elements** | **Meaning assigned to elements** |
|---|---|
| Boxes | |
| Color of the Boxes | No particular meaning, graphical purposes only |
| Solid black frame | Manifest Variables assessed in one of the studies |
| Dashed black frame | Not directly assessed elements, but suggested by theory |
| Arrows | |
| Solid black ⟶ | Effects examined in one of the presented studies |
| Dashed black ⇢ | Assumed processes (directed path) |
| Dotted grey ⇢ | Effects examined in other studies (specified in text) |
| Double headed ⟷ | Correlations |

Starting in the top-right, situation cues are perceived and then evaluated (i). This evaluation process may be similar to those suggested by appraisal theories of emotion (Scherer, 2001; see below for a more detailed description of the process). In appraisal theories, this evaluation leads to emotions or affect (j). At the same time, a perception of the situation is formed (k), and situation perception and affect, which are correlated (c), then allow accounting for variance in behavior (d and b, respectively). Behavior then influences situation cues (h). Additionally, stable person characteristics allow explaining variance in behavior (a), but also predict affect (f) and situation perception (e). Furthermore, it was shown that behavior can influence affect (g). Finally, some assumed effects were not examined in this dissertation, but may be examined in future studies (l). However, if affect and situation perception are a product of the same evaluative process, why are situation perception and affect differentiable? As they predicted unique shares in variance in behavior, they must represent different information (Horstmann et al., submitted). There are several possible explanations for this finding, and these will be discussed in the next section.

*Figure 8.4.1.* Process model of situation perception and affect




*Note.* This model is based on a model presented by Rauthmann, Sherman, and Funder (2015). Solid arrows depict effects that were examined in the dissertation, dotted arrows depict effects examined elsewhere, dashed arrows depict assumed processes or effects. Note that not all assumed or possible effects are depicted. See text for a detailed explanation.

### *4.4.1 Overlap of Affect and Situation Perception*

The evaluation process in emotion appraisal involves four steps (Scherer, 2001):

1) Relevance of the situation
2) Implication of the situation
3) Coping Potential of the person in that situation
4) Normative Significance of the situation

Each step consists of several components or questions that are answered (see Figure 8.4.1). For the relevance of the situation, these are the novelty check (is the event or cue new?), the intrinsic pleasantness check (will a stimulus result in pleasure or pain?), and goal relevance check (how relevant is the stimulus for current needs or goals?). For the implication assessment, these are the causal attribution check (who is responsible for the event?), outcome probability check (how likely are certain outcomes?), discrepancy from expectation check (how consistent or discrepant is the current situation from a person's expectations?), goal/need conduciveness check (how helpful is a situation for current needs/goals?), and the urgency check (how urgent is the situation for an accomplishment of goals?). The coping potential determination assessed the coping abilities of the person; the control check (can the event be controlled at all?), power check (if control over the situation is possible, does the person have the power to do it?), and the adjustment check (can the person adjust to the consequences after control has been exerted?). For the evaluation of the normative significance, a person checks internal standards (does an action in/reaction to the situation satisfy or fall short of internal, personal standards?) and external standards (which social consequences does an action in/reaction to the situation have?).

The difference of situation perception and affect now lies in the appraisal of each component and the response the person provides to each question. Consider someone doing a job such as working on a conveyor belt. This person may perceive some items (packages, tools, etc.). To evaluate the task, the person now performs the four steps and answers each of the questions at each step. For example, the task is relevant (it is not new, but also not unpleasant, and relevant to secure income), it has implications (the person needs to stand there, and not doing the work results directly in loss of income), but the person has a good coping potential and thus can do the job reasonably well. Furthermore, both internal standards (how the person would like to perform) and external standards are met (the supervisor is satisfied with the work). Thus, no change in affect or emotions occurs. Affect- or emotion-wise, the situation is not particularly relevant. Yet, during this process, a lot of information was collected, which now shapes situation perception, namely that the situation contains some Duty. One could say that for affect or emotions to result from

this process, the four steps are connected in a multiplicative way: If one is not relevant/does not apply, no change in affect or emotions will occur. For situation perception, these septs are additive: Each individual step provides additional information on the situation and forms situation perception.

To conclude, the information collected during the process shapes situation perception, and the personal evaluation of the information leads to affect. Furthermore, if certain evaluations take place that lead to affect, these can at the same time influence situation perception, thus leading to the correlation of situation perception and affect. If a person, for example, does not possess the means to cope with a task, this could lead to negative affect, but also to the perception of enhanced Duty.

Some situation perception dimensions (e.g., Negativity and pOsitivity from the DIAMONDS taxonomy) are highly correlated with affect (Horstmann & Ziegler, 2018), and yet they account for variance in behavior after controlling for positive and negative affect (Horstmann et al., submitted). In the example of the conveyor belt job, it could be assumed that the person concludes that the situation is negative (e.g., because it contains heavy work, some colleagues suffer under the job-demand), but the person him- or herself does not suffer. Note that in the context of occupational psychology, this process resembles that of the Job-Demand-Resources Model (Karasek, 1979), which describes how job demands and available resources form job-satisfaction and well-being. Here, the process of situation perception can be employed to describe the stage in the model at which the job demands and resources are perceived.

As mentioned above (section 3), affect and emotions are related, but not indifferentiable. Whereas emotions are related to objects, this is not necessarily the case for affect. More specifically, affect is also influenced by diurnal rhythms, the weather, or the temperature (see Gendolla, 2000). Affect also influences the processing of visual information (Gasper & Clore, 2002), and affect is also considered to provide information about a stimulus (Schwarz & Clore, 1983), although it is questionable if these effects are robust (Yap et al., 2017). One could therefore argue that affect also influences the evaluation of situations directly, and emotions are a consequence of situation perception. However, this needs to be examined empirically and should, for now, be understood as a hypothesis.

### *4.4.2 Explaining idiosyncratic and consensual variance in situation perception.*

As mentioned above, situation perception contains two forms of variance: Idiosyncratic and shared or consensual variance. Furthermore, variance in situation perception can be attributed to the person, the situation, and their interaction.

Based on the process described here, I suggest that person variance and person x situation variance should mainly be part of the idiosyncratic variance and thus be aligned with person effects. More specifically, I suggest that person variance should best be explained with stable person characteristics (i.e., traits), and person x situation variance should best be accounted for by characteristics of the person in the situation (e.g., affect, emotion). Finally, consensual variance, and thus situation variance, should be mostly be accounted for by effects of situation cues.

For example, all students taking exam will perceive some a similar level of Duty (consensual variation) as they all have to take the exam to successfully complete their degree. This is an effect that should be attributed to the situation, the exam. Yet, not all students will perceive the same level of Duty, which is idiosyncratic variance. Some of this is person-variance, which can be explained with person effects (e.g., conscientiousness), and some variance can be attributed to the person x situation interaction (e.g., having little knowledge about some specific items, and thus little coping ability), which leads to affect and a perception of more Duty. This variance in situation perception should therefore correspond mostly with affect.

It should be possible to test this assumption, for example, by assessing situation perception and affect from several perceivers in multiple situations. Raters would then be nested in situations. At situation level, the average score across raters in situations would represent the shared variance (how each situation is seen on average), on perceiver level, the average score within rater, across situations would be the idiosyncratic situation variance per person, and the variance that remains unexplained is the idiosyncratic, person and situation specific variance, which should best be explained by affect or person states of the person in the situation. Such a model would be a multilevel confirmatory factor analysis with non-interchangeable raters (Eid et al., 2008; Nussbeck, Eid, Geiser, Courvoisier, & Lischetzke, 2009).

### *4.4.3 The Role of Time*

The investigation of the model depicted in Figure 8.4.1 clearly requires a more thorough recognition of time as a central component. Behavior feedbacks to situation cues, for example, through changes made to the situation (e.g., Gosling et al., 2002) or by leaving the situation and selecting a new one (e.g., Rauthmann, Sherman, Nave, et al., 2015). Additionally, it might be possible that affect from one situation influences the perception and interpretation of another, later situation. Affect piggybacks on the person and thus "travels" from one situation to another, whereas situation perception – by definition – should not. The classic pink-glasses metaphor describes this effect: A person in a certain affective state perceives situations differently than others

in a different affective state. Investigating such a hypothesis requires the manipulation of affect and a *subsequent* assessment of situation perception.

As we have mainly ignored time throughout all analyses and models, this constitutes a limitation to the interpretation of the current work. Establishing a sequence of events (e.g., does affect come before or after situation perception?) and the duration of a process (e.g., how long does it take to evaluate a situation?) are key requirements to establishing causality and thus for gaining a deeper understanding of the processes involved. The current work may be seen as a first stepping stone towards such an endeavor. We can now confidently say that affect, personality, and situation perception all have a place in the prediction of behavior. Yet, the dynamics of the processes involved are unknown as of now. Nevertheless, the model suggested here (Figure 8.4.1) provides an additional starting point for such an investigation.

## 4.5 Relevance for Personality Theories

Personality models, frameworks, and theories that were developed in the wake of the person-situation debate aimed at incorporating behavioral variability and the stability of personality traits. While Trait theories and LST theory focused more on the description of stable and variable aspects of personality, respectively, TAT and CAPS focused more on explaining the processes that are involved in creating behavioral variability. WTT has finally combined both approaches by splitting personality traits into a descriptive part and an explanatory part. These personality theories highlight variables that should be investigated if the contradiction of stable personality traits and variable personality states should be bridged, namely the situation and person states. The current work shows that these variables indeed play an important role in accounting for variance in human behavior, and recent personality theories therefore serve well as guidelines for future studies.

At the same time, these theories are vague in their formulation of testable hypotheses, especially with respect to the explanatory processes. For example, LST theory states that personality states can be "decomposed into a latent trait variable and an occasion-specific residual which represents the effect of the situation in which the person is when assessment is performed and the person-situation interaction." (Steyer et al., 1999, p. 392). Although the theory makes personality states measurable (e.g., latent state residuals could be predicted using other variables assessed on the state-level), no specific explanation is offered what "the effect of the situation" actually means. CAPS, and thus also WTT which explicitly builds on CAPS, offer some more insight into these processes. Specifically, Fleeson and Jayawickreme (2015) propose five different processes that are involved in creating trait-expressions and thus explaining intra-individual variability (these five processes are the interpretative process, motivational process, stability-

inducing process, temporal process, and random error process). The interpretative process, which "represents the cognitive aspects of the mind – the manner in which information is processed and which results in implications for behavior" (Fleeson & Jayawickreme, 2015, p. 392), could resemble the evaluation-process depicted in Figure 8.4.1. It is important to precisely articulate these processes as this is what makes them testable. Further, this articulation facilitates integration of theories, as similarities in the assumed and underlying processes become apparent.

The current work shows how testing little pieces of larger theories separately can lead to a better understanding of personality theories and subsequently the rearrangement of certain aspects of these theories. Fleeson and Jayawickreme (2015) acknowledge that WTT is not yet final, and that the more knowledge is generated the more will it be possible to re-arrange these theories to accommodate new findings. At the same time, untested elements of personality models, frameworks, or theories can guide new studies and help with the specification of precise, testable hypotheses.

## 4.6 Consequences for Consistency

The person-situation debate was closely aligned with the search for consistency in behavior and finding a "new locus of consistency" (Mischel et al., 2002). This work has highlighted several important points that need to be considered in the future when consistency should be examined and variability in behavior should be accounted for:

1. It must be made clear which type of consistency is studied (see Article: Distinguishing Simple and Residual Consistency, section 1.2.2).
2. Depending on the type of consistency examined, the influence of situations has to be considered.
3. Regardless of the type of consistency, affect (and potentially other states) should be considered as additional explanatory variables for behavior (see section 3.2).

Without heeding these three steps, the search for consistency will likely be a never-ending story. If affect, for example, was not considered in the prediction of behavior, a person might behave differently in two functionally equivalent situations (Horstmann et al., in preparation). This difference in behavior could then not be explained with variability in situation cues – there is none – nor with changes in personality. However, short-term fluctuations in affect, emotions, or other person states can account for such fluctuations above and beyond the effects of situational influences (Horstmann et al., submitted). Figure 8.4.1 as well as the model proposed by Rauthmann and colleagues (2015) suggest further factors that might be able to account for additional variability

in behavior. These include social roles, motivation, physiological states, or experience with or knowledge of the situation.

Baumert and colleagues (2017) recently called for an integration of personality structure, processes, and development. As the current work has demonstrated, such an integration is worthwhile. It helps understanding why persons vary in their behaviors, even across similar situations. Although some argued that the person-situation debate was pointless (Hogan, 2009) or old wine in new bottles (J. A. Johnson, 1999), it has moved personality psychology considerably forward. Personality traits exist and are meaningful, and yet, much more about the processes involved in shaping behavior awaits to be discovered.

# 5 Part 5: Outlook

This dissertation bears several implications for future research. These implications concern personality theories, the understanding of personality traits, a recognition of processes involved, and, on a more abstract level, the value and insights that can be gained from replication studies.

As several personality theories have been suggested, it is now time to examine and test them. Advancements in situation research and methods for data collection and analysis allow testing and integrating personality theories. This will require making underlying processes explicit, thereby setting them up to criticism and debate. A deeper understanding of the processes involved requires also an integration of different areas of psychology. The current work included research from personality psychology, psychological assessment, occupational psychology, and psychological methods as well as research on emotion and affect, and research on situations and effects of situation cues (an area primarily associated with social psychology). Combining research from these different areas has allowed finding explanations for unexpected effects and the formulation of testable hypotheses.

For personality psychology more specifically, the research presented here reflects a call made recently by Baumert and colleagues (2017). The authors argued for an integration of personality processes, structure, and development. The research presented here as only addressed personality processes and personality structure; however, it has already been insightful and allowed a better understanding of the relation of personality and behavior. Adding a developmental perspective to this research will undoubtedly increase both the effort required to undertake it but also the knowledge gained from it. Furthermore, it may be possible to apply similar methods to other theories of personality psychology. For example, research on self-other-agreement may benefit from considering psychological situations (or contexts) in which peers form their peer-reports (Finnigan & Vazire, 2017). Similarly, current research on situations and personality usually operationalizes the latter as the Big Five personality traits, but other traits' expression in daily life could also be considered, for example narcissism (e.g., Maaß & Ziegler, 2017) or morality (e.g., Hofmann, Wisneski, Brandt, & Skitka, 2014).

A large chunk of the current work involved the re-analysis of existing data and the examination of a published research finding, namely the study by Sherman and colleagues (2015). This re-analysis has triggered a subsequent pre-registration and replication of previous work and therefore an estimation of the robustness of previously published effects. Furthermore, the re-analysis allowed making specific predictions of expected effects, it simplified study design (e.g., conducting a power analysis), and thus limited the overall number of researcher degrees of

freedom, which leads to more informative results in general (Wicherts et al., 2016). Replications and extensions of prior work should therefore be encouraged in future research of personality psychology, and the current work may serve as an example of the informativeness of such endeavors. At the same time, all critique and cautiousness that applies to previous research must, of course, also apply to current research, and thus also this dissertation. The results presented here are not set in stone, and future research may cast a new light on some of the findings, their generalizability, and their interpretation.

Referring to the person-situation debate, Fleeson wrote in 2004 "this is an exciting time to be a personality psychologist, unshackled by doubts about the value of one's field and encouraged by the promise of future productive integration of opposing viewpoints" (Fleeson, 2004, p. 86). Research from previous years as well as the current work may be seen as part of the integration of opposing viewpoints, having added further value to the field of personality psychology. At the same time, none of the excitement about possible future discoveries is lost – quite the contrary, new avenues for research on the dynamic person were opened. It undoubtedly is still an exciting time to be a personality psychologist.

# 6 References


Allport, G. W. (1936). *Personality: a psychological interpretation.* Oxford, England: Holt.

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, *47*, i-171.

Andersen, S. M., & Thorpe, J. S. (2009). An If–Then theory of personality: Significant others and the relational self. *Journal of Research in Personality*, *43*, 163–170.

Arendasy, M. (2011). *Big-Five Struktur Inventar (BFSI)*. Mödling: Schuhfried GmbH.

Arslan, R. C., & Tata, C. (2015). formr.org survey software (Version v0.8.2).

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, *70*, 1–70.

Asendorpf, J. B. (1990). The measurement of individual consistency. *Methodika*, *4*, 1–22.

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J. A., Fiedler, K., … Wicherts, J. M. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality*, *27*, 108–119.

Asendorpf, J. B., Penke, L., & Back, M. D. (2011). From dating to mating and relating: Predictors of initial and long-term outcomes of speed-dating in a community sample. *European Journal of Personality*, *25*, 16–30.

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, *11*, 150–166.

Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, *91*, 340–345.

Augustine, A. A., & Larsen, R. J. (2012). Is a trait really the mean of states? *Journal of Individual Differences*, *33*, 131–137.

Back, M. D., Schmukle, S. C., & Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology*, *97*, 533–548.

Bäckström, M. (2007). Higher-order factors in a Five-Factor personality inventory and its relation to social desirability. *European Journal of Psychological Assessment*, *23*, 63–70.

Bäckström, M., Björklund, F., & Larsson, M. R. (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality*, *43*, 335–344.

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*, 512–527.

Bargh, J. A. (2007). Social psychological approaches to consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge Handbook of Consciousness* (pp. 555–570). Cambridge: Cambridge University Press.

Barker, R. (1968). *Ecological psychology: Concepts and methods for studying the environment of human behavior*. Stanford, California: Stanford University Press.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, *9*, 9–30.

Bartoń, K. (2016). MuMIn: Multi-Model Inference.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2016). lme4: Linear mixed-effects models using "Eigen" and S4.

Battistich, V. A., & Thompson, E. G. (1980). Students' perceptions of the college milieu: A multidimensional scaling analysis. *Personality and Social Psychology Bulletin*, *6*, 74–82.

Baumeister, R. F., Vohs, K. D., DeWall, N. C., & Zhang, L. (2007). How emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation. *Personality and Social Psychology Review*, *11*, 167–203.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, *2*, 396–403.

Baumert, A., Schmitt, M., Perugini, M., Johnson, W., Blum, G., Borkenau, P., … Wrzus, C. (2017). Integrating personality structure, personality process, and personality development. *European Journal of Personality*, *31*, 503–528.

Behrendt, S. (2014). lm.beta: Add standardized regression coefficients to lm-objects.

Bem, D., & Allen, A. (1972). Constructing cross-situational consistencies in behavior: Some thoughts on Alker's critique of Mischel. *Journal of Personality*, *40*, 17–26.

Bem, D., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review*, *81*, 506–520.

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., … Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*, 6–10.

Bettencourt, B. A., Talley, A., Benjamin, A. J., & Valentine, J. (2006). Personality and aggressive behavior under provoking and neutral conditions: A meta-analytic review. *Psychological Bulletin*, *132*, 751–777.

Biderman, M. D., Nguyen, N. T., Cunningham, C. J. L., & Ghorbani, N. (2011). The ubiquity of common method variance: The case of the Big Five. *Journal of Research in Personality*, *45*, 417–

429.

Bless, H., Bohner, G., Schwarz, N., & Strack, F. (1990). Mood and persuasion. *Personality and Social Psychology Bulletin*, *16*, 331–345.

Block, J. (1978). *The Q-sort method in personality assessment and psychiatric research.* Palo Alto, CA: Consulting Psychologists Press. (Original work published 1961).

Bond, M. H. (2013). Refining Lewin's formula: A general model for explaining situational influence on individual social behavior. *Asian Journal of Social Psychology*, *16*, 1–15.

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.

Brennan, R. L. (1992). *Elements of generalizability theory (rev. ed.).* Iowa City: ACT.

Brogden, H. E., & Taylor, E. K. (1950). The theory and classification of criterion bias. *Educational and Psychological Measurement*, *10*, 159–183.

Brown, N. A., Blake, A. B., & Sherman, R. A. (2017). A snapshot of the life as lived. *Social Psychological and Personality Science*, *8*, 592–600.

Brown, N. A., Neel, R., & Sherman, R. A. (2015). Measuring the evolutionarily important goals of situations: Situational affordances for adaptive problems. *Evolutionary Psychology*, *13*, 1–15.

Brown, N. A., & Rauthmann, J. F. (2016). Situation characteristics are age graded: Mean-level patterns of the situational eight DIAMONDS across the life span. *Social Psychological and Personality Science*, *7*, 667–679.

Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology, 1904-1920*, *3*, 296–322.

Burns, M. N., Begale, M., Duffecy, J., Gergle, D., Karr, C. J., Giangrande, E., & Mohr, D. C. (2011). Harnessing context sensing to develop a mobile intervention for depression. *Journal of Medical Internet Research*, *13*, e55.

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

Campus, N. (1974). Transituational consistency as a dimension of personality. *Journal of Personality and Social Psychology*, *29*, 593–600.

Cattell, R. B. (1963). Personality, role, mood, and situation-perception: A unifying theory of modulators. *Psychological Review*, *70*, 1–18.

Chaplin, W. F., & Goldberg, L. R. (1984). A failure to replicate the Bem and Allen study of individual differences in cross-situational consistency. *Journal of Personality and Social Psychology*, *47*, 1074–1090.

Chen, J., Furukawa, T. A., Nakano, Y., Ietsugu, T., Ogawa, S., Funayama, T., … Rapee, R. M. (2010). Video feedback with peer ratings in naturalistic anxiety-provoking situations for social anxiety disorder: Preliminary report. *Journal of Behavior Therapy and Experimental Psychiatry*, *41*, 6–10.

Ching, C. M., Church, A. T., Katigbak, M. S., Reyes, J. A. S., Tanaka-Matsumi, J., Takaoka, S., … Ortiz, F. A. (2014). The manifestation of traits in everyday behavior and affect: A five-culture study. *Journal of Research in Personality*, *48*, 1–16.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Cooper, M. L. (2016). Editorial. *Journal of Personality and Social Psychology*, *110*, 431–434.

Costa, P. T., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology*, *38*, 668–678.

Costa, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the revised NEO Personality Inventory. *Journal of Personality Assessment*, *64*, 21–50.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.

Cuperman, R., & Ickes, W. (2009). Big Five predictors of behavior and perceptions in initial dyadic interactions: personality similarity helps extraverts and introverts, but hurts "disagreeables." *Journal of Personality and Social Psychology*, *97*, 667–684.

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, *8*, 377–383.

Deary, I. (2009). The trait approach to personality. In G. Matthews & P. J. Corr (Eds.), *The Cambridge handbook of personality psychology*. (pp. 89–109). Cambridge: Cambridge University Press.

de Vries, R. E., Tybur, J. M., Pollet, T. V., & van Vugt, M. (2016). Evolution, situational affordances, and the HEXACO model of personality. *Evolution and Human Behavior*, *37*, 407–421.

Deinzer, R., Steyer, R., Eid, M., Notz, P., Schwenkmezger, P., Ostendorf, F., & Neubauer, A. (1995). Situational effects in trait assessment: The FPI, NEOFFI, and EPI questionnaires. *European Journal of Personality*, *9*, 1–23.

Dejonckheere, E., Mestdagh, M., Houben, M., Erbas, Y., Pe, M., Koval, P., … Kuppens, P. (2018). The bipolarity of affect and depressive symptoms. *Journal of Personality and Social Psychology*, *114*, 323–341.

Denissen, J. J. A., Geenen, R., Selfhout, M., & van Aken, M. A. G. (2008). Single-item big five ratings in a social network design. *European Journal of Personality*, *22*, 37–54.

DeYoung, C. G. (2010). Toward a theory of the Big Five. *Psychological Inquiry*, *21*, 26–33.

DeYoung, C. G. (2015). Cybernetic Big Five theory. *Journal of Research in Personality*, *56*, 33–58.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*, 880–896.

Diener, E., Oishi, S., & Lucas, R. E. (2003). Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life. *Annual Review of Psychology*, *54*, 403–425.

Donnellan, M. B., Lucas, R. E., & Fleeson, W. (2009). Introduction to personality and assessment at age 40: Reflections on the legacy of the person–situation debate and the future of person–situation integration. *Journal of Research in Personality*, *43*, 117–119.

Dragulescu, A. A. (2014). xlsx: Read, write, format Excel 2007 and Excel 97/2000/XP/2003 files.

Dudenredaktion. (2006). *Die deutsche Rechtschreibung. Bd. 1–24. völlig neu bearb. u. erweiterte Auflage*. Mannheim: Dudenverlag.

Eckes, T. (1995). Features of situations: A two-mode clustering study of situation prototypes. *Personality and Social Psychology Bulletin*, *21*, 366–374.

Edwards, J., & Templeton, A. (2005). The structure of perceived qualities of situations. *European Journal of Social Psycholgy*, *723*, 705–723.

Eid, M., Nussbeck, F. W., Geiser, C., Cole, D. A., Gollwitzer, M., & Lischetzke, T. (2008). Structural equation modeling of multitrait-multimethod data: Different models for different types of methods. *Psychological Methods*, *13*, 230–253.

Emmons, R. A., & Diener, E. (1986). Situation selection as a moderator of response consistency and stability. *Journal of Personality and Social Psychology*, *51*, 1013–1019.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*, 121–138.

Endler, N. S., Hunt, J. M., & Rosenstein, A. J. (1962). An S-R inventory of anxiousness. *Psychological Monographs: General and Applied*, *76*, 1–33.

Epstein, S. (1979). The stability of behavior: On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, *37*, 1097–1126.

Epstein, S. (1980). The stability of behavior: II. Implications for psychological research. *American*

*Psychologist*, *35*, 790–806.

Epstein, S. (1983a). Aggregation and beyond: Some basic issues on the prediction of behavior. *Journal of Personality*, *51*, 360–392.

Epstein, S. (1983b). The stability of confusion: A reply to Mischel and Peake. *Psychological Review*, *90*, 179–184.

Finnigan, K. M., & Vazire, S. (2017). The incremental validity of average state self-reports over global self-reports of personality. *Journal of Personality and Social Psychology*. Advance online publication.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, *51*, 327–358.

Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, *80*, 1011–1027.

Fleeson, W. (2004). Moving personality beyond the person-situation debate. The challenge and the opportunity of within-person variability. *Current Directions in Psychological Science*, *13*, 83–87.

Fleeson, W. (2007). Situation-based contingencies underlying trait-content manifestation in behavior. *Journal of Personality*, *75*, 825–862.

Fleeson, W., & Gallagher, P. (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis. *Journal of Personality and Social Psychology*, *97*, 1097–1114.

Fleeson, W., & Jayawickreme, E. (2015). Whole Trait Theory. *Journal of Research in Personality*, *56*, 82–92.

Fleeson, W., & Law, M. K. (2015). Trait enactments as density distributions: The role of actors, situations, and observers in explaining stability and variability. *Journal of Personality and Social Psychology*, *109*, 1090–1104.

Fleeson, W., Malanos, A. B., & Achille, N. M. (2002). An intraindividual process approach to the relationship between extraversion and positive affect: Is acting extraverted as "good" as being extraverted? *Journal of Personality and Social Psychology*, *83*, 1409–1422.

Fleeson, W., & Noftle, E. E. (2008). Where does personality have its influence? A supermatrix of consistency concepts. *Journal of Personality*, *76*, 1355–86.

Fleeson, W., & Noftle, E. E. (2009). In favor of the synthetic resolution to the person–situation debate. *Journal of Research in Personality*, *43*, 150–154.

Forgas, J. P. (1976). The perception of social episodes: Categorical and dimensional representations in two different social milieus. *Journal of Personality and Social Psychology*, *34*, 199–209.

Forgas, J. P. (2000). Affect and information processing strategies: An interactive relationship. In J. P. Forgas (Ed.), *Studies in emotion and social interaction, second series. Feeling and thinking: The role of affect in social cognition* (pp. 253–280). New York: Cambridge University Press.

Fournier, M. A., Moskowitz, D. S., & Zuroff, D. C. (2008). Integrating dispositions, signatures, and the interpersonal domain. *Journal of Personality and Social Psychology*, *94*, 531–545.

Fournier, M. A., Moskowitz, D. S., & Zuroff, D. C. (2009). The interpersonal signature. *Journal of Research in Personality*, *43*, 155–162.

Fox, J., & Weisberg, S. (2011). *An {R} Companion to Applied Regression* (Second). Thousand Oaks, CA: Sage.

Funder, D. C. (1997). *The personality puzzle*. (7th ed.). W W Norton & Co.

Funder, D. C. (2001). Personality. *Annual Review of Psychology*, *52*, 197–221.

Funder, D. C. (2006). Towards a resolution of the personality triad: Persons, situations, and behaviors. *Journal of Research in Personality*, *40*, 21–34.

Funder, D. C. (2009). Persons, behaviors and situations: An agenda for personality psychology in the postwar era. *Journal of Research in Personality*, *43*, 120–126.

Funder, D. C. (2016). Taking situations seriously. *Current Directions in Psychological Science*, *25*, 203–208.

Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology*, *60*, 773–94.

Funder, D. C., & Ozer, D. J. (1983). Behavior as a function of the situation. *Journal of Personality and Social Psychology*, *44*, 107–112.

Furr, R. M. (2009a). Personality psychology as a truly behavioural science. *European Journal of Personality*, *23*, 369–401.

Furr, R. M. (2009b). Profile analysis in person–situation integration. *Journal of Research in Personality*, *43*, 196–207.

Furr, R. M., & Funder, D. C. (2004). Situational similarity and behavioral consistency: Subjective, objective, variable-centered, and person-centered approaches. *Journal of Research in Personality*, *38*, 421–447.

Gasper, K., & Clore, G. L. (2002). Attending to the Big Picture: Mood and Global Versus Local Processing of Visual Information. *Psychological Science*, *13*, 34–40.

Geiser, C., Litson, K., Bishop, J., Keller, B. T., Burns, G. L., Servera, M., & Shiffman, S. (2015). Analyzing person, situation and person × situation interaction effects: Latent state-trait models for the combination of random and fixed situations. *Psychological Methods*, *20*, 165–192.

Gelman, A., & Loken, E. (2014). The statistical crisis in science. *American Scientist*, *102*, 460–465.

Gendolla, G. H. E. (2000). On the impact of mood on behavior: An integrative theory and a review. *Review of General Psychology*, *4*, 378–408.

Gerpott, F. H., Balliet, D., Columbus, S., Molho, C., & de Vries, R. E. (2017). How do people think about interdependence? A multidimensional model of subjective outcome interdependence. *Journal of Personality and Social Psychology*. Advance online publication.

Geukes, K., Nestler, S., Hutteman, R., Küfner, A. C. P., & Back, M. D. (2017). Trait personality and state variability: Predicting individual differences in within- and cross-context fluctuations in affect, self-evaluations, and behavior in everyday life. *Journal of Research in Personality*, *69*, 124–138.

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78.

Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on "Estimating the reproducibility of psychological science." *Science*, *351*, 1037–1037.

Giner-Sorolla, R., Kupfer, T., & Sabo, J. (2018). What makes moral disgust special? An integrative functional review. *Advances in Experimental Social Psychology*, *57*, 223–289.

Goldberg, L. R. (1990). An alternative "description of personality": the Big Five factor structure. *Journal of Personality and Social Psychology*, *59*, 1216–1229.

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology*, *82*, 379–398.

Greenwald, A. G. (1976). An Editorial. *Journal of Personality and Social Psychology*, *33*, 1–7.

Greiff, S., & Heene, M. (2017). Why psychological assessment needs to start worrying about model fit. *European Journal of Psychological Assessment*, *33*, 313–317.

Griner, L. A., & Smith, C. A. (2000). Contributions of motivational orientation to appraisal and emotion. *Personality and Social Psychology Bulletin*, *26*, 727–740.

Guillaume, E., Baranski, E., Todd, E., Bastian, B., Bronin, I., Ivanova, C., … Funder, D. C. (2016). The World at 7:00: Comparing the experience of situations across 20 countries. *Journal of Personality*, *84*, 493–509.

Gurtman, M. B. (2009). Exploring personality with the interpersonal circumplex. *Social and Personality Psychology Compass*, *3*, 601–619.

Hall, P. A., Fong, G. T., & Epp, L. J. (2013). Cognitive and personality factors in the prediction of health behaviors: an examination of total, direct and indirect effects. *Journal of Behavioral Medicine*, *37*, 1057–1068.

Hampel, R. (1977). Adjektiv-Skalen zur Einschätzung der Stimmung (SES). *Diagnostica*, *23*, 43–60.

Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016). Using smartphones to collect behavioral data in psychological science. *Perspectives on Psychological Science*, *11*, 838–854.

Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Bühner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: a cautionary note on the usefulness of cutoff values of fit indices. *Psychological Methods*, *16*, 319–336.

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58*, 78–79.

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science*, *345*, 1340–1343.

Hofstee, W. K. B. (1994). Who should own the definition of personality? *European Journal of Personality*, *8*, 149–162.

Hogan, R. (2009). Much ado about nothing: The person–situation debate. *Journal of Research in Personality*, *43*, 249.

Holtzman, N. S., Vazire, S., & Mehl, M. R. (2010). Sounds like a narcissist: Behavioral manifestations of narcissism in everyday life. *Journal of Research in Personality*, *44*, 478–484.

Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review*, *14*, 332–346.

Horstmann, K. T. (2015). *Putting Lewin's Equation to the Test: Assessing the Person-Situation Interaction with the B5PS*. https://doi.org/10.17605/OSF.IO/Z32RU

Horstmann, K. T., & Rauthmann, J. F. (in preparation). How many states make a trait? A comprehensive meta-analysis of experience sampling studies.

Horstmann, K. T., Rauthmann, J. F., & Sherman, R. A. (2018). Measurement of situational influences. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *The SAGE Handbook of Personality and Individual Differences* (pp. 465–484). SAGE Publications.

Horstmann, K. T., Rauthmann, J. F., Sherman, R. A., & Ziegler, M. (submitted). Unveiling an exclusive link: Predicting behavior with personality, situation perception, and affect in a pre-registered experience sampling study.

Horstmann, K. T., Rauthmann, J. F., & Ziegler, M. (in preparation). Distinguishing simple and residual consistency in functionally equivalent situations: Evidence from variable- and person-centered analyses in longitudinal data.

Horstmann, K. T., Ziegler, J., & Ziegler, M. (2018). Assessment of situational perceptions:

Measurement issues and a joint taxonomization of persons and situations. In D. C. Funder, J. F. Rauthmann, & R. A. Sherman (Eds.), *The Oxford Handbook of Psychological Situations*. Oxford University Press.

Horstmann, K. T., & Ziegler, M. (2016). Situational perception: Its theoretical foundation, assessment, and links to personality. In U. Kumar (Ed.), *The Wiley Handbook of Personality Assessment* (1st ed., pp. 31–43). Oxford: Wiley Blackwell.

Horstmann, K. T., & Ziegler, M. (2018). Situational perception and affect: Barking up the wrong tree? *Personality and Individual Differences*. Advance online publication

Hox, J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY, US: Routledge.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55.

Ickes, W., Snyder, M., & Garcia, S. (1997). Personality Influences on the Choice of Situations. In *Handbook of Personality Psychology* (pp. 165–195). Elsevier.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*, e124.

IPIP. (2015). International Personality Item Pool: A scientific collaboratory for the development of advanced measures of personality traits and other individual differences. Retrieved from http://ipip.ori.org/

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). New York, NY, US.

John, O. P., & Srivastava, 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., pp. 102–38). New York, NY: Guilford Press.

Johnson, J. A. (1999). Persons in situations: Distinguishing new wine from old wine in new bottles. *European Journal of Personality*, *13*, 443–453.

Johnson, J. V, & Hall, E. M. (1988). Job strain, work place social support, and cardiovascular disease: a cross-sectional study of a random sample of the Swedish working population. *American Journal of Public Health*, *78*, 1336–1342.

Jones, A. B., Brown, N. A., Serfass, D. G., & Sherman, R. A. (2017). Personality and density distributions of behavior, emotions, and situations. *Journal of Research in Personality*, *69*, 225–236.

Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology*, *87*, 530–541.

Judge, T. a, Simon, L. S., Hurst, C., & Kelley, K. (2014). What I experienced yesterday is who I am today: Relationship of work motivations and behaviors to within-individual variation in the five-factor model of personality. *Journal of Applied Psychology*, *99*, 199–221.

Kanfer, F. H., & Saslow, G. (1965). Behavioral Analysis: An alternative to diagnostic classification. *Archives of General Psychiatry*, *12*, 529–538.

Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly*, *24*, 285–308.

Keller, M. C., & Nesse, R. M. (2006). The evolutionary significance of depressive symptoms: Different adverse situations lead to different depressive symptom patterns. *Journal of Personality and Social Psychology*, *91*, 316–330.

Kelley, H. H., Holmes, J. G., Kerr, N. L., Reis, H. T., Rusbult, C. E., & van Lange, P. A. M. (2003). *An atlas of interpersonal situations*. New York, NY, US: Cambridge University Press.

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations: a theory of interdependence. Interpersonal relations: A theory of interdependence*. New York, NY, US: Wiley.

Kelley, T. L. (1927). *Interpretation of Educational Measurement*. Yonkers-on-Hudson, NY: World Book Co.

Kerr, N. L. (1998). HARKing: Hypothesizing After the Results are Known. *Personality and Social Psychology Review*, *2*, 196–217.

Kim, H. J., Shin, K. H., & Swanger, N. (2009). Burnout and engagement: A comparative analysis using the Big Five personality dimensions. *International Journal of Hospitality Management*, *28*, 96–104.

King, G. a., & Sorrentino, R. M. (1983). Psychological dimensions of goal-oriented interpersonal situations. *Journal of Personality and Social Psychology*, *44*, 140–162.

Klages, L. (1926). *Die Grundlagen der Charakterkunde*. Bonn: Bouvier Verlag.

Klehe, U.-C., Kleinmann, M., Hartstein, T., Melchers, K. G., König, C. J., Heslin, P. A., & Lievens, F. (2012). Responding to personality tests in a selection context: The role of the ability to identify criteria and the ideal-employee factor. *Human Performance*, *25*, 273–302.

Krause, M. S. (1970). Use of social situations for research purposes. *American Psychologist*, *25*, 748–753.

Krohne, W. H., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der Positive and Negative Affect Schedule (PANAS). *Diagnostica*, *42*, 139–156.

Krueger, J. I. (2009). A componential model of situation effects, person effects, and situation-by-person interaction effects on social behavior. *Journal of Research in Personality*, *43*, 127–136.

Kulas, J. T., & Stachowski, A. A. (2009). Middle category endorsement in odd-numbered Likert response scales: Associated item characteristics, cognitive demands, and preferred meanings. *Journal of Research in Personality*, *43*, 489–493.

Kuppens, P. (2009). The legacy of the person–situation debate for understanding variability in emotional experience. *Journal of Research in Personality*, *43*, 255–256.

Kuppens, P. (2015). It's about time: A special section on affect dynamics. *Emotion Review*, *7*, 297–300.

Kuppens, P., & Van Mechelen, I. (2007). Interactional appraisal models for the anger appraisals of threatened self-esteem, other-blame, and frustration. *Cognition & Emotion*, *21*, 56–77.

Lakens, D., Adolfi, F. G., Albers, C. J., Anvari, F., Apps, M. A. J., Argamon, S. E., … Zwaan, R. A. (2018). Justify your alpha. *Nature Human Behaviour*, *2*, 168–171.

Lämmle, L., Oedl, C., & Ziegler, M. (2014). Don't threaten me and my dark side or even self-harm won't stop me from hurting you. *Personality and Individual Differences*, *67*, 87–91.

Lang, F. R., Lüdtke, O., & Asendorpf, J. B. (2001). Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen. *Diagnostica*, *47*, 111–121.

Larsen, R. J. (2000). Toward a science of mood regulation. *Psychological Inquiry*, *11*, 129–141.

Latané, B., & Nida, S. (1981). Ten years of research on group size and helping. *Psychological Bulletin*, *89*, 308–324.

LeBel, E. P., Berger, D., Campbell, L., & Loving, T. J. (2017). Falsifiability is not optional. *Journal of Personality and Social Psychology*, *113*, 254–261.

Leckelt, M., Küfner, A. C. P., Nestler, S., & Back, M. D. (2015). Behavioral processes underlying the decline of narcissists' popularity over time. *Journal of Personality and Social Psychology*, *109*, 856–871.

Lee, K., & Ashton, M. C. (2008). The HEXACO personality factors in the indigenous personality lexicons of english and 11 other languages. *Journal of Personality*, *76*, 1001–1054.

Lefcheck, J. S. (2016). piecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, *7*, 573–579.

Leikas, S., Lönnqvist, J.-E., & Verkasalo, M. (2012). Persons, situations, and behaviors:

Consistency and variability of different behaviors in four interpersonal situations. *Journal of Personality and Social Psychology*, *103*, 1007–1022.

Lench, H. C., Flores, S. a, & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin*, *137*, 834–855.

Lewin, K. (1936). *Principles of topological psychology. New York-London*. New York: McGraw Hill.

Lievens, F. (2017). Assessing personality-situation interplay in personnel selection: toward more integration into personality research. *European Journal of Personality*, *31*, 424–440.

Little, T. D., Cunningham, W. a., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: exploring the question, weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*, 151–173.

Lucas, R. E., Diener, E., Grob, A., Suh, E. M., & Shao, L. (2000). Cross-cultural evidence for the fundamental features of extraversion. *Journal of Personality and Social Psychology*, *79*, 452–468.

Lucas, R. E., & Fujita, F. (2000). Factors influencing the relation between extraversion and pleasant affect. *Journal of Personality and Social Psychology*, *79*, 1039–1056.

Lynott, D., Corker, K. S., Wortman, J., Connell, L., Donnellan, M. B., Lucas, R. E., & O'Brien, K. (2014). Replication of "experiencing physical warmth promotes interpersonal warmth" by Williams and Bargh (2008). *Social Psychology*, *45*, 216–222.

Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research*, *46*, 137–155.

Maaß, U., Lämmle, L., Bensch, D., & Ziegler, M. (2016). Narcissists of a feather flock together: narcissism and the similarity of friends. *Personality and Social Psychology Bulletin*, *42*, 366–384.

Maaß, U., & Ziegler, M. (2017). Narcissistic self-promotion is not moderated by the strength of situational cues. *Personality and Individual Differences*, *104*, 482–488.

Maassen, G. H. (2004). The standard error in the Jacobson and Truax Reliable Change Index: The classical approach to the assessment of reliable change. *Journal of the International Neuropsychological Society*, *10*, 888–893.

MacCann, C., Danay, E., Ziegler, M., & Roberts, R. (2011). A comprehensive facet structure for agreeableness/extraversion. *Paper Presented at the ISSID, London.*

Magnusson, D. (1971). An aanalysis of situational dimensions. *Perceptual and Motor Skills*, *32*, 851–867.

Mathieu, J. E., Aguinis, H., Culpepper, S. a., & Chen, G. (2012). "Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling": Correction to Mathieu, Aguinis, Culpepper, and Chen (2012). *Journal of Applied Psychology*, *97*, 981–981.

Mauno, S., Kinnunen, U., & Ruokolainen, M. (2007). Job demands and resources as antecedents of work engagement: A longitudinal study. *Journal of Vocational Behavior*, *70*, 149–171.

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

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, *60*, 577–605.

McCrae, R. R., & Costa, P. T. (2008). Empirical and theoretical status of the five-factor model of personality traits. In *The SAGE Handbook of Personality Theory and Assessment: Volume 1 — Personality Theories and Models* (pp. 273–294). London: SAGE Publications Ltd.

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, *15*, 28–50.

McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, *86*, 730–740.

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, *22*, 114–140.

McNiel, J. M., & Fleeson, W. (2006). The causal effects of extraversion on positive affect and neuroticism on negative affect: Manipulating state extraversion and state neuroticism in an experimental approach. *Journal of Research in Personality*, *40*, 529–550.

Mehl, M. R., Pennebaker, J. W., Crow, D. M., Dabbs, J., & Price, J. H. (2001). The Electronically Activated Recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, & Computers*, *33*, 517–523.

Mehl, M. R., Robbins, M. L., & Deters, F. groe. (2012). Naturalistic observation of health-relevant social processes. *Psychosomatic Medicine*, *74*, 410–417.

Mehl, M. R., Vazire, S., Ramirez-Esparza, N., Slatcher, R. B., & Pennebaker, J. W. (2007). Are women really more talkative than men? *Science*, *317*, 82–82.

Miguel, E., Camerer, C., Casey, K., Cohen, J., Esterling, K. M., Gerber, A., … Van der Laan, M. (2014). Promoting transparency in social science research. *Science*, *343*, 30–31.

Miller, J. D., Gaughan, E. T., Maples, J., & Price, J. (2011). A comparison of agreeableness scores from the Big Five Inventory and the NEO PI-R: Consequences for the study of narcissism and psychopathy. *Assessment*, *18*, 335–339.

Mischel, W. (1968). *Personality and assessment.* New York: Wiley.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality.

*Psychological Review*, *80*, 252–283.

Mischel, W. (1977). The interaction of person and situation. In D. Magnusson & N. S. Endler (Eds.), *Personality at the cross-roads: Current issues in interactional psychology* (pp. 333–352). Hilsdale, NY: Lawrence Erlbaum.

Mischel, W. (2009). From Personality and Assessment (1968) to Personality Science, 2009. *Journal of Research in Personality*, *43*, 282–290.

Mischel, W., & Peake, P. K. (1982). Beyond déjà vu in the search for cross-situational consistency. *Psychological Review*, *89*, 730–755.

Mischel, W., & Shoda, Y. (1994). Personality Psychology Has Tow Goals: Must It Be Two Fields? *Psychological Inquiry*, *5*, 24–27.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*, 246–68.

Mischel, W., Shoda, Y., & Mendoza-Denton, R. (2002). Situation-Behavior profiles as a locus of consistency in personality. *Current Directions in Psychological Science*, *11*, 50–54.

Moors, A. (2014). Flavors of Appraisal Theories of Emotion. *Emotion Review*, *6*, 303–307.

Moos, R. H. (1973). Conceptualizations of human environments. *American Psychologist*, *28*, 652–665.

Morse, P. J., Sauerberger, K. S., Todd, E., & Funder, D. (2015). Relationships among personality, situational construal and social outcomes. *European Journal of Personality*, *29*, 97–106.

Moshagen, M., Hilbig, B. E., & Zettler, I. (2014). Faktorenstruktur, psychometrische Eigenschaften und Messinvarianz der deutschsprachigen Version des 60-Item HEXACO Persönlichkeitsinventars. *Diagnostica*, *60*, 86–97.

Mount, M. K., Barrick, M. R., Scullen, S. M., & Rounds, J. (2005). Higher-Order dimensions of the Big Five personality traits and the Big Six Vocational interest types. *Personnel Psychology*, *58*, 447–478.

Murray, H. A. (1938). *Explorations in personality*. Oxford, England: Oxford Univ. Press.

Muthén, L. K., & Muthen, B. O. (2015). Mplus.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R 2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142.

Nascimento-Schulze, C. M. (1981). Towards situational classification. *European Journal of Social Psychology*, *11*, 149–159.

Nezlek, J. B. (2017). A practical guide to understanding reliability in studies of within-person variability. *Journal of Research in Personality*, *69*, 149–155.

Nussbeck, F. W., Eid, M., Geiser, C., Courvoisier, D. S., & Lischetzke, T. (2009). A CTC( M −1) Model for different types of raters. *Methodology*, *5*, 88–98.

Ones, D. S., Mount, M. K., Barrick, M. R., & Hunter, J. E. (1994). Personality and job performance: A critique of the Tett, Jackson, and Rothstein (1991) meta-analysis. *Personnel Psychology*, *47*, 147–156.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*, aac4716-aac4716.

Oreg, S., Edwards, J., & Rauthmann, J. F. (in preparation). The Situation Six: Uncovering six basic dimensions of psychological situations from the Hebrew lexicon.

Orom, H., & Cervone, D. (2009). Personality dynamics, meaning, and idiosyncrasy: Identifying cross-situational coherence by assessing personality architecture. *Journal of Research in Personality*, *43*, 228–240.

Oshio, A., Abe, S., Cutrone, P., & Gosling, S. D. (2014). Further Validity of the Japanese Version of the Ten Item Personality Inventory (TIPI-J). *Journal of Individual Differences*, *35*, 236–244.

Oud, J. H. L., Voelkle, M. C., & Driver, C. C. (2017). SEM based CARMA time series modeling for arbitrary N. *Multivariate Behavioral Research*, 1–21.

Ozer, D. J. (1985). Correlation and the coefficient of determination. *Psychological Bulletin*, *97*, 307–315.

Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, *57*, 401–421.

Pace, V. L., & Brannick, M. T. (2010). How similar are personality scales of the "same" construct? A meta-analytic investigation. *Personality and Individual Differences*, *49*, 669–676.

Parrigon, S. (2017). The lexical approach to situations: History, theory, and practice. In *The Oxford Handbook of Psychological Situations*. Oxford University Press. Advance online publication.

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.

Paunonen, S. V, & Ashton, M. C. (2001). Big five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology*, *81*, 524–539.

Pervin, L. A. (1976). A free-response description approach to the analysis of person-situation interaction. *Journal of Personality and Social Psychology*, *34*, 465–474.

Pervin, L. A., & John, O. P. (1999). *Handbook of personality: Theory and research*. New York, NY:

Guilford Press.

Petty, R. E., Fabrigar, L. R., & Wegener, D. T. (2003). Emotional factors in attitudes and persuasion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 72–772). Oxford: Oxford University Press.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879–903.

Price, R. H. (1974). The taxonomic classification of behaviors and situations and the problem of behavior-environment congruence. *Human Relations*, *27*, 567–585.

Price, R. H., & Blashfield, R. K. (1975). Explorations in the taxonomy of behavior settings. *American Journal of Community Psychology*, *3*, 335–351.

Pustejovsky, J. (2016). clubSandwich: Cluster-robust (sandwich) variance estimators with small-sample corrections.

R Core Team. (2016). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria*. http://www.R-project.org/.

Rauthmann, J. F. (2012). You say the party is dull, i say it is lively: A componential approach to how situations are perceived to disentangle perceiver, situation, and perceiver x situation variance. *Social Psychological and Personality Science*, *3*, 519–528.

Rauthmann, J. F. (2015). Structuring situational information. *European Psychologist*, *20*, 176–189.

Rauthmann, J. F. (2016). Motivational factors in the perception of psychological situation characteristics. *Social and Personality Psychology Compass*, *10*, 92–108.

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.

Rauthmann, J. F., & Horstmann, K. T. (2017). Overview of Situation Characteristic Taxonomies. http://doi.org/10.17605/OSF.IO/M3R6M

Rauthmann, J. F., Horstmann, K. T., & Sherman, R. A. (2018). Do self-reported traits and aggregated states capture the same thing? A nomological perspective on trait-state homomorphy. *Social Psychological and Personality Science*. Advance online publication.

Rauthmann, J. F., Jones, A. B., & Sherman, R. A. (2016). Directionality of person-situation transactions: are there spillovers among and between situation experiences and personality states? *Personality and Social Psychology Bulletin*, *42*, 893–909.

Rauthmann, J. F., & Sherman, R. A. (2016a). Measuring the situational eight DIAMONDS

characteristics of situations. *European Journal of Psychological Assessment*, *32*, 155–164.

Rauthmann, J. F., & Sherman, R. A. (2016b). Situation change: Stability and change of situation variables between and within Persons. *Frontiers in Psychology*, *6*.

Rauthmann, J. F., & Sherman, R. A. (2016c). Ultra-brief measures for the situational eight DIAMONDS domains. *European Journal of Psychological Assessment*, *32*, 165–174.

Rauthmann, J. F., & Sherman, R. A. (2018a). The description of situations: Towards replicable domains of psychological situation characteristics. *Journal of Personality and Social Psychology*, *114*, 482–488.

Rauthmann, J. F., & Sherman, R. A. (2018b). Toward a research agenda for the study of situation perceptions: a variance componential framework. *Personality and Social Psychology Review*, *114*, 482–488.

Rauthmann, J. F., Sherman, R. A., & Funder, D. C. (2015a). New horizons in research on psychological situations and environments. *European Journal of Personality*, *29*, 419–432.

Rauthmann, J. F., Sherman, R. A., & Funder, D. C. (2015b). Principles of situation research: Towards a better understanding of psychological situations. *European Journal of Personality*, *29*, 363–381.

Rauthmann, J. F., Sherman, R. A., Nave, C. S., & Funder, D. C. (2015). Personality-driven situation experience, contact, and construal: How people's personality traits predict characteristics of their situations in daily life. *Journal of Research in Personality*, *55*, 98–111.

Read, S. J., Smith, B. J., Droutman, V., & Miller, L. C. (2017). Virtual personalities: Using computational modeling to understand within-person variability. *Journal of Research in Personality*, *69*, 237–249.

Revelle, W. (2014). psych: Procedures for Psychological, Psychometric, and Personality Research. http://CRAN.R-project.org/package=psych Version = 1.4.5.

Revelle, W. (2016). psych: Procedures for Psychological, Psychometric, and Personality Research.

Revelle, W., & Condon, D. M. (2015). A model for personality at three levels. *Journal of Research in Personality*, *56*, 70–81.

Richard, E., & Diener, E. (2009). Personality and Subjective Well-Being. In E. Diener (Ed.), *The Science of Well-Being: The Collected Works of Ed Diener* (pp. 75–102). Dordrecht: Springer Netherlands.

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*, 331–363.

Riemann, R., Angleitner, A., Borkenau, P., & Eid, M. (1998). Genetic and environmental sources of consistency and variability in positive and negative mood. *European Journal of Personality*,

*12*, 345–364.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*, 313–345.

Robinson, D. (2016). broom: Convert statistical analysis objects into tidy data frames.

Rockstuhl, T., Ang, S., Ng, K.-Y., Lievens, F., & Van Dyne, L. (2015). Putting judging situations into situational judgment tests: Evidence from intercultural multimedia SJTs. *Journal of Applied Psychology*, *100*, 464–480.

Rosseel, Y. (2015). lavaan: Latent Variable Analysis.

RStudio Team. (2015). RStudio: Integrated development environment for R. Boston, MA.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*, 581–592.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*, 145–172.

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*, 805–819.

Sander, D., Grandjean, D., & Scherer, K. R. (2005). A systems approach to appraisal mechanisms in emotion. *Neural Networks*, *18*, 317–352.

Saucier, G., Bel-Bahar, T., & Fernandez, C. (2007). What modifies the expression of personality tendencies? Defining basic domains of situation variables. *Journal of Personality*, *75*, 479–503.

Schaufeli, W. B., & Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: a multi-sample study. *Journal of Organizational Behavior*, *25*, 293–315.

Schaufeli, W. B., Salanova, M., Gon Alez-ro, V. A., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness Studies*, *3*, 71–92.

Scherer, K. R. (1999). On the sequential nature of appraisal processes: indirect evidence from a recognition task. *Cognition & Emotion*, *13*, 763–793.

Scherer, K. R. (2001). Appraisal considered as a process of multi-level sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 92–120). New York: Oxford University Press.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings.

*Psychological Bulletin*, *124*, 262–274.

Schmit, M. J., & Ryan, A. M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology*, *78*, 966–974.

Schmitt, M. (1990a). Further evidence on the invalidity of self-reported consistency. In *European perspectives in psychology* (Vol. I, pp. 57–68).

Schmitt, M. (1990b). *Konsistenz als Persönlichkeitseigenschaft? Moderatorvariablen in der Persönlichkeits- und Einstellungsforschung.* Berlin: Springer.

Schmitt, M. (1992). Interindividuelle Konsistenzunterschiede als Herausforderung für die Differentielle Psychologie. *Psychologische Rundschau*, *43*, 30–45.

Schmitt, M. (2009). Person×situation-interactions as moderators. *Journal of Research in Personality*, *43*, 267.

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*, 609–612.

Schwartz, G. E., & Weinberger, D. A. (1980). Patterns of emotional responses to affective situations: Relations among happiness, sadness, anger, fear, depression, and anxiety. *Motivation and Emotion*, *4*, 175–191.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.

Seppälä, P., Mauno, S., Feldt, T., Hakanen, J., Kinnunen, U., Tolvanen, A., & Schaufeli, W. (2009). The construct validity of the Utrecht Work Engagement Scale: Multisample and longitudinal evidence. *Journal of Happiness Studies*, *10*, 459–481.

Serfass, D. G., & Sherman, R. A. (2013). Personality and perceptions of situations from the Thematic Apperception Test. *Journal of Research in Personality*, *47*, 708–718.

Shaffer, J. A., & Postlethwaite, B. E. (2012). A matter of context: A meta-analytic investigation of the relative validity of contextualized and noncontextualized personality measures. *Personnel Psychology*, *65*, 445–494.

Sherman, R. A. (2015). multicon: Multivariate Constructs.

Sherman, R. A., Nave, C. S., & Funder, D. C. (2010). Situational similarity and personality predict behavioral consistency. *Journal of Personality and Social Psychology*, *99*, 330–343.

Sherman, R. A., Nave, C. S., & Funder, D. C. (2012). Properties of persons and situations related to overall and distinctive personality-behavior congruence. *Journal of Research in Personality*, *46*, 87–101.

Sherman, R. A., Nave, C. S., & Funder, D. C. (2013). Situational construal is related to

personality and gender. *Journal of Research in Personality*, *47*, 1–14.

Sherman, R. A., Rauthmann, J. F., Brown, N. A., Serfass, D. G., & Jones, A. B. (2015). The independent effects of personality and situations on real-time expressions of behavior and emotion. *Journal of Personality and Social Psychology*, *109*, 872–888.

Shoda, Y., & Mischel, W. (2000). Reconciling contextualism with the core assumptions of personality psychology. *European Journal of Personality*, *14*, 407–428.

Shoda, Y., Mischel, W., & Wright, J. C. (1994). Intraindividual stability in the organization and patterning of behavior: Incorporating psychological situations into the idiographic analysis of personality. *Journal of Personality and Social Psychology*, *67*, 674–87.

Siemer, M. (2009). Mood experience: Implications of a dispositional theory of moods. *Emotion Review*, *1*, 256–263.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*, 107–120.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359–1366.

Smith, R. E., Shoda, Y., Cumming, S. P., & Smoll, F. L. (2009). Behavioral signatures at the ballpark: Intraindividual consistency of adults' situation–behavior patterns and their interpersonal consequences. *Journal of Research in Personality*, *43*, 187–195.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.

Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology, 1904-1920*, *3*, 271–295.

Stanley, D. (2015). apaTables: Create American Psychological Association (APA) style tables.

Steyer, R., Ferring, D., & Schmitt, M. (1992). States and traits in psychological assessment. *European Journal of Psychological Assessment*, *8*, 79–98.

Steyer, R., Mayer, A., Geiser, C., & Cole, D. A. (2015). A theory of States and Traits—revised. *Annual Review of Clinical Psychology*, *11*, 71–98.

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state–trait theory and research in personality and individual differences. *European Journal of Personality*, *13*, 389–408.

Stroebe, W., & Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science*, *9*, 59–71.

Swami, V., Stieger, S., Voracek, M., Dressler, S. G., Eisma, L., & Furnham, A. (2009). Psychometric evaluation of the tagalog and German Subjective Happiness Scales and a

cross-cultural comparison. *Social Indicators Research*, *93*, 393–406.

Tackett, J. L., Lilienfeld, S. O., Patrick, C. J., Johnson, S. L., Krueger, R. F., Miller, J. D., … Shrout, P. E. (2017). It's time to broaden the replicability conversation: Thoughts for and from clinical psychological science. *Perspectives on Psychological Science*, *12*, 742–756.

Ten Berge, M. A., & De Raad, B. (2001). The construction of a joint taxonomy of traits and situations. *European Journal of Personality*, *15*, 253–276.

Ten Berge, M. A., & De Raad, B. (2002). The structure of situations from a personality perspective. *European Journal of Personality*, *102*, 81–102.

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, *88*, 500–517.

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*, *34*, 397–423.

Thalmayer, A. G., & Saucier, G. (2014). The questionnaire Big Six in 26 Nations: Developing cross-culturally applicable Big Six, Big Five and Big Two inventories. *European Journal of Personality*, *28*, 482–496.

Thomas, K. M., Yalch, M. M., Krueger, R. F., Wright, A. G. C., Markon, K. E., & Hopwood, C. J. (2013). The convergent structure of DSM-5 personality trait facets and Five-Factor model trait domains. *Assessment*, *20*, 308–311.

Van Heck, G. L. (1984). The construction of a general taxonomy of situations. In H. Bonarius, G. L. Van Heck, & N. Smid (Eds.), *Personlity psychology in Europe: Theoretical and empirical developments* (pp. 149–164). Lisse, The Netherlands: Swets and Zeitlinger.

Van Heck, G. L. (1989). Situation concepts: Definitions and classification. *Personality and Environment: Assessment of Human Adaptation*, *53–69*.

Van Mechelen, I. (2009). A royal road to understanding the mechanisms underlying person-in-context behavior. *Journal of Research in Personality*, *43*, 179–186.

van Mechelen, I., & De Raad, B. (1999). Editorial: personality and situations. *European Journal of Personality*, *13*, 333–336.

Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*, 281–300.

Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: the accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, *95*, 1202–1216.

Vickers, A. J., & Altman, D. G. (2001). Statistics notes: Analysing controlled trials with baseline

and follow up measurements. *Bmj*, *323*, 1123–1124.

Wagenmakers, E., Beek, T., Dijkhoff, L., Gronau, Q. F., Acosta, A., Adams, R. B., … Zwaan, R. A. (2016). Registered Replication Report. *Perspectives on Psychological Science*, *11*, 917–928.

Wagerman, S., & Funder, D. (2009). Personality psychology of situations. In P. J. Corr & G. Matthews (Eds.), *The Cambridge handbook of personality psychology* (pp. 27–42). New York, NY, US: Cambridge University Press.

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, *54*, 426–483.

Ware, J. J., & Munafò, M. R. (2015). Significance chasing in research practice: causes, consequences and possible solutions. *Addiction*, *110*, 4–8.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of Positive and Negative Affect: The PANAS Scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.

Watson, D., & Gray, E. K. (2007). Assessing positive and negative affect via self-report. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 171–183). New York, NY, US: Oxford University Press.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*, 219–235.

Weidman, A. C., Steckler, C. M., & Tracy, J. L. (2017). The jingle and jangle of emotion assessment: Imprecise measurement, casual scale usage, and conceptual fuzziness in emotion research. *Emotion*, *17*, 267–295.

Weisbuch, M., Slepian, M. L., Clarke, A., Ambady, N., & Veenstra-VanderWeele, J. (2010). Behavioral stability across time and situations: Nonverbal versus verbal consistency. *Journal of Nonverbal Behavior*, *34*, 43–56.

Westhoff, K., & Kluck, M. L. (2008). *Psychologische Gutachten: schreiben und beurteilen. Entspricht deutschen und europäischen Richtlinien zur Erstellung psychologischer Gutachten.* Berlin: Springer.

Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M. (2016). Degrees of Freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking. *Frontiers in Psychology*, *7*.

Wickham, H. (2016a). purrr: Functional programming tools.

Wickham, H. (2016b). stringr: Simple, consistent wrappers for common string operations.

Wickham, H., & Francois, R. (2016). dplyr: A Grammar of data manipulation.

Williams, L. E., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science. 322*, 606–607.

Wilson, R. E., Thompson, R. J., & Vazire, S. (2017). Are fluctuations in personality states more than fluctuations in affect? *Journal of Research in Personality*, *69*, 110–123.

Wilt, J. A., Bleidorn, W., & Revelle, W. (2017). Velocity explains the links between personality states and affect. *Journal of Research in Personality*, *69*, 86–95.

Wilt, J. A., & Revelle, W. (2017). The Big Five, everyday contexts and activities, and affective experience. *Personality and Individual Differences*. Advance online publication.

Winkielman, P., Knutson, B., Paulus, M., & Trujillo, J. L. (2007). Affective influence on judgments and decisions: Moving towards core mechanisms. *Review of General Psychology*, *11*, 179–192.

Wittmann, W. W., & Süß, H. M. (1999). Investigating the paths between working memory, intelligence, knowledge, and complex problem-solving performances via Brunswik symmetry. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences. Process, trait content determinants.* (pp. 77–104). Washington DC: American Psychological Association.

Wrzus, C., Hänel, M., Wagner, J., & Neyer, F. J. (2013). Social network changes and life events across the life span: A meta-analysis. *Psychological Bulletin*, *139*, 53–80.

Wrzus, C., & Mehl, M. R. (2015). Lab and/or field? Measuring personality processes and their social consequences. *European Journal of Personality*, *29*, 250–271.

Wrzus, C., & Roberts, B. W. (2017). Processes of Personality Development in Adulthood: The TESSERA Framework. *Personality and Social Psychology Review*, *21*, 253–277.

Wrzus, C., Wagner, G. G., & Riediger, M. (2016). Personality-situation transactions from adolescence to old age. *Journal of Personality and Social Psychology*, *110*, 782–799.

Wyer, R. S., Clore, G. L., & Isbell, L. M. (1999). Affect and information processing. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (pp. 1–77). Academic Press.

Xie, Y. (2016). knitr: A General-Purpose Package for Dynamic Report Generation in R.

Yang, Y., Read, S. J., & Miller, L. C. (2006). A taxonomy of situations from Chinese idioms. *Journal of Research in Personality*, *40*, 750–778.

Yap, S. C. Y., Wortman, J., Anusic, I., Baker, S. G., Scherer, L. D., Donnellan, M. B., & Lucas, R. E. (2017). The effect of mood on judgments of subjective well-being: Nine tests of the judgment model. *Journal of Personality and Social Psychology*, *113*, 939–961.

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, *12*, 1100–1122.

Yik, M. S. M., Russell, J. A., & Barrett, L. F. (1999). Structure of self-reported current affect: Integration and beyond. *Journal of Personality and Social Psychology*, *77*, 600–619.

Youyou, W., Stillwell, D., Schwartz, H. A., & Kosinski, M. (2017). Birds of a feather do flock together. *Psychological Science*, *28*, 276–284.

Ziegler, M. (2014a). *B5PS. Big Five Inventory of Personality in Occupational Situations*. Mödling, Austria: Schuhfried GmbH.

Ziegler, M. (2014b). Comments on item selection procedures. *European Journal of Psychological Assessment*, *30*, 1–2.

Ziegler, M. (2014c). Stop and state your intentions! *European Journal of Psychological Assessment*, *30*, 239–242.

Ziegler, M., & Bäckström, M. (2016). 50 Facets of a Trait – 50 Ways to Mess Up? *European Journal of Psychological Assessment*, *32*, 105–110.

Ziegler, M., Bensch, D., Maaß, U., Schult, V., Vogel, M., & Bühner, M. (2014). Big Five facets as predictor of job training performance: The role of specific job demands. *Learning and Individual Differences*, *29*, 1–7.

Ziegler, M., Booth, T., & Bensch, D. (2013). Getting entangled in the nomological net. *European Journal of Psychological Assessment*, *29*, 157–161.

Ziegler, M., & Brunner, M. (2016). Test standards and psychometric modeling . In A. A. Lipnevich, F. Preckel, & R. Roberts (Eds.), *Psychosocial Skills and School Systems in the 21st Century* (pp. 29–55). Göttingen: Springer.

Ziegler, M., & Buehner, M. (2009). Modeling socially desirable responding and its effects. *Educational and Psychological Measurement*, *69*, 548–565.

Ziegler, M., Cengia, A., & Roberts, R. (in prep.). A facetted Big Five measure: Construction, evaluation and measurement invariance across German and American samples.

Ziegler, M., Danay, E., Schölmerich, F., & Bühner, M. (2010). Predicting academic success with the big 5 rated from different points of view: Self-rated, other rated and faked. *European Journal of Personality*, *24*, 341–355.

Ziegler, M., Ehrlenspiel, F., & Brand, R. (2009). Latent state-trait theory: An application in sport psychology. *Psychology of Sport and Exercise*, *10*, 344–349.

Ziegler, M., & Hagemann, D. (2015). Testing the unidimensionality of items. *European Journal of Psychological Assessment*, *31*, 231–237.

Ziegler, M., & Horstmann, K. (2015). Discovering the second side of the coin. *European Journal of Psychological Assessment*, *31*, 69–74.

Ziegler, M., Horstmann, K. T., & Ziegler, J. (submitted). Personality in situations: Going beyond the OCEAN and introducing the Situation Five.

Ziegler, M., Poropat, A., & Mell, J. (2014). Does the length of a questionnaire matter? *Journal of*

*Individual Differences*, *35*, 250–261.

Ziegler, M., & Ziegler, J. (2015). Better understanding of psychological situations: opportunities and challenges for psychological assessment. *European Journal of Personality*, *29*, 418–419.

# 7 Statement on Independence

Hiermit erkläre ich, die Dissertation selbstständig und nur unter Verwendung der angegebenen Hilfen und Hilfsmittel angefertigt zu haben. Ich habe mich anderwärts nicht um einen Doktorgrad beworben und besitze keinen entsprechenden Doktorgrad. Ich erkläre, dass ich die Dissertation oder Teile davon nicht bereits bei einer anderen wissenschaftlichen Einrichtung eingereicht habe und dass sie dort weder angenommen noch abgelehnt wurde. Ich erkläre die Kenntnisnahme der dem Verfahren zugrunde liegenden Promotionsordnung der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin vom 5. März 2015. Weiterhin erkläre ich, dass keine Zusammenarbeit mit gewerblichen Promotionsbearbeiterinnen/ Promotionsberatern stattgefunden hat und dass die Grundsätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis eingehalten wurden.

*I hereby declare that I completed the doctoral thesis independently based on the stated resources and aids. I have not applied for a doctoral degree elsewhere and do not have a corresponding doctoral degree. I have not submitted the doctoral thesis, or parts of it, to another academic institution and the thesis has not been accepted or rejected. I declare that I have acknowledged the Doctoral Degree Regulations which underlie the procedure of the Faculty of Life Sciences of Humboldt-Universität zu Berlin, as amended on 5th March 2015. Furthermore, I declare that no collaboration with commercial doctoral degree supervisors took place, and that the principles of Humboldt-Universität zu Berlin for ensuring good academic practice were abided by.*

_____

Kai T. Horstmann