Validity of Stages and the Role of Intervention Engagement in the Health Behavior Change Process

Dissertation zur Erlangung des akademischen Grades

Doktor der Philosophie (Dr. phil.)

am Fachbereich Erziehungswissenschaft und Psychologie

der Freien Universität Berlin

vorgelegt von Dipl.-Psych. Jana Richert

Berlin, 2011

Erstgutachter: Prof. Dr. Ralf Schwarzer

Zweitgutachterin: Prof. Dr. Sonia Lippke

Disputation: 23. 05. 2011

Table of Contents

Acknowledgement

First of all, I would like to thank Prof. Dr. Ralf Schwarzer, my dissertation advisor and mentor. By guiding me with your invaluable experience, by challenging me to surpass my own accomplishments and by allowing me to follow my own paths in research, you let me grow from an inexperienced and idealistic intern into a well practiced, independent and more grounded researcher. I will continually draw from the wisdom and advice you passed on over the years.

Secondly, I owe my thanks to Prof. Dr. Sonia Lippke, my supervisor and first source of consult whenever I got stuck. Your pragmatic approach to issues is usually 'spot on' and your feedback on my manuscripts has been tremendously useful in improving my scientific writing. Sonia, without your encouragement to apply as an intern in the department six years ago, I would not be here today. For this, I am very grateful. Also, the fact that you manage to raise two little boys while continuing to further your scientific career sets a great example for all women in academia!

My special recognition goes to Prof. Dr. Leona Aiken, whom I've first met on the Berlin U-Bahn exactly 4 years ago and who has since been a source of advice and support. It is an honor to work with you and learn from how you approach a subject. Leona, your encouragement all those years ago will forever be appreciated.

My gratitude also goes to Prof. Dr. Cynthia Mohr for providing me with some perspective and appreciation when I most needed it.

My heartfelt thanks belongs to Dr. Tabea Reuter. Tabea, your compassion, your positive attitude, and your constructive feedback as well as your open ear, in both academic and personal matters, have helped me get here at last. You are a wonderful role model and I am grateful to call you my friend.

I also want to thank Dr. Benjamin Schüz for working with me, for sharing his experience, for his excellent feedback and encouragement. Benjamin, I greatly appreciate your support.

To Lisa Warner, my accomplice, my companion: We started this journey together and we are finishing it together. And we are *still* friends. But seriously, I am very thankful for being able to share this accomplishment with you. You have been a great source of knowledge, understanding and friendship.

My fellow PhD students, colleagues, co-authors and most importantly, *friends*, Sarah Arpin, Dr. Cécile Bazillier, Sylvia Böhme, Dr. Debi Brannan, Anna Ernsting, Lena Fleig, Paul Gellert, Milena Koring, Daniela Lange, Natalie Schüz, Linda Parschau, Sarah Pomp, and Staci Wendt. I thank you for your open ears, for being my sparring partners to bounce ideas off of, for trusting my ability whenever I lost confidence, and most of all, for making work a *fun* place to go to. You are all such a big part of my life. For everything and more, I thank you.

Recognition also belongs to all the (student) research assistants and interns, who have helped me over the past years in one way or another: Janet Geipel, Katharina Gördeler, Isabel Hübner, Genia Juschkewitsch, Birgitta Kielhorn, Jan Keller, Lucyna Klinicka, Carina Küper, Robi Meyer, Susann Prasse, and Julia Tholeikis. Thank you!

For their help in proofreading parts of my dissertation, I want to thank Monique Kleine, Dr. Francis Quinn, Isabella Vormittag and Sara Wiger.

I am also thankful to Prof. Dr. Falko Sniehotta and Prof. Dr. Irina Todorova, who were somehow important along the way and whose input had an impact on my work.

And at last, *all my family and friends*: Without your support, I would not have finished this dissertation. Seriously. To Kathleen, Veronika and Dr. Waldemar Richert, Naomi Baird, Amanda Rhoads LaGuardia and Skyler Corbett, for believing in me. Always.

Abstract

Stage theories of health behavior change assume that rather than generic 'one size fits all' interventions, individuals need stage appropriate treatments, addressing only those factors that are relevant for a particular stage and mindset. To ensure the receipt of appropriate and consequently effective health promotion interventions, stages need to adequately represent individuals' mindsets towards behavior change. The first part of this dissertation deals with this issue.

In particular, it was investigated whether the standards for behaviors that are embedded in stage measures impact stage allocation and consequently, which standard leads to an allocation that yields the best reflection of individuals' mindsets. A common external standard was compared with an individual standard. Results indicate that the different standards indeed affected stage-allocation. Staging based on an individual standard seemed to reflect individuals' mindsets better than staging based on the external standard. This suggests that misclassification (i.e., an inaccurate reflection of individuals' mindsets) is likely if an arbitrary standard that individuals are measured against is employed. It was concluded that it may be more beneficial to rely on individuals' subjective evaluation when it comes to assessing their stage.

In this first study, mindsets were defined through intentions and behavior only. Consequently, in a second study, mindsets delineated as complete social-cognitive profiles were extracted directly from the data. These 'full' mindsets were then pinned against those stages that were reflective of individuals' behavior and intentions. However, correspondence was low; suggesting that the stage measure based on behavior and intentions might not be effective in assigning individuals to stages that reflect their full mindsets towards behavior change accurately, thus undermining the possibility of stage appropriate interventions.

Researchers often restrict their treatment evaluations to the predictors specified in theories of behavior change and conclude that depending on the outcome of the intervention study, they either addressed the right or the wrong determinants of behavior

change. The second part of this dissertation overcomes this shortcoming by addressing the question of how a behavior change intervention is received.

In particular, it deals with the construct of *engagement* in the context of health behavior change interventions. Specifically, it proposes a comprehensive definition of the construct as well as a new self-report measure, the Task Engagement Scale (TES). In the presented study, the factorial structure of the scale is examined and data on its relationship both with objective indicators of engagement and with theoretically related constructs provided. The items of the TES loaded on four subscales: task-compliance, effort, undivided attention and absorption, which in turn loaded on the higher order factor engagement. The TES captured more information than the objective indicators time-on-task and the completion rate of intervention materials. Associations between theoretically related constructs and the TES score indicate that engagement is a sufficiently distinct construct.

In the final part of this dissertation both aspects were integrated. In particular, the role of intervention engagement was examined in the behavior change process. The mechanism via which the intervention exerted its influence on changes in fruit and vegetable consumption was through changes in planning cognitions. This mediation was moderated by participants' engagement in the way that the treatment led to changes in cognitions only when participants' engagement in the treatment was at a moderate level. This result demonstrates that a theory- and evidence based intervention does not invariably lead to changes in the cognitions it targets. The implication is that when developing health promotion interventions, researchers and practitioners have to give consideration to factors that affect participants' engagement in the intervention.

This dissertation recognized and addressed two issues in health psychology and offered a response. To conclude, this dissertation was a first step in gaining a more comprehensive understanding of what leads to behavior change.

Stadien und Interventions-Engagement im Prozess der Gesunheitsverhaltensänderung

Zusammenfassung

In dieser Dissertation werden *zwei komplementäre Ansätze* verfolgt, die gemeinsam zu einem umfassenderen Wissen über den Prozess der Gesundheitsverhaltensänderung beitragen sollen. Der erste Teil der Arbeit lässt sich in den Bereich der *Theorien der Gesundheitsverhaltensänderung* einordnen, auf deren Basis Interventions*inhalte* entwickelt werden können (1). Der zweite Teil beschäftigt sich mit den *Empfängern* von Interventionen, wobei der Fokus hier vor allem auf dem *Engagement* der Teilnehmer (d.h. dem Ausmaß der Auseinandersetzung mit Interventionsinhalten) liegt (2). Im letzten Teil werden beide Ansätze integriert (3).

(1) Theorien der Gesundheitsverhaltensänderung postulieren Prädiktoren sowie Mechanismen der Verhaltensänderung und können so die Inhalte von Gesundheits-förderungsmaßnahmen bestimmen. Die zentrale Annahme von *Stadientheorien* der Gesundheitsverhaltensänderung ist, dass sich Menschen hinsichtlich ihrer *Mindsets* (d.h. Verhaltens- und Denkweisen) unterscheiden. Diese Unterschiede werden in *maßgeschneiderten* Interventionsmaßnahmen berücksichtigt, indem ausschließlich solche Faktoren angesprochen werden, die für das jeweilige Mindset, in dem sich ein Mensch innerhalb des Prozesses der Gesundheitsverhaltensänderung befindet, relevant sind. Es wird davon ausgegangen, dass solche Maßnahmen effektiv zur Verhaltensänderung führen. Aus pragmatischen Gründen werden aber oft nicht komplette Mindsets erfasst, sondern *Stadien*, die aus dem Verhalten und den Verhaltensabsichten der Menschen abgeleitet werden und Rückschlusse über Mindsets zulassen sollen. Um also zu gewährleisten, dass eine Maßnahme tatsächlich passend und damit wirksam ist, müssen Stadien die Mindsets der Menschen adäquat repräsentieren.

Menschen können Stadien zugeordnet werden, indem ihr Verhalten und ihre

Absichten mit Verhaltens*kriterien* (d.h. Standards und Richtlinien für ausreichend Verhalten, z.B. körperliche Aktivität) verglichen werden. Bislang werden in der Forschung und der Praxis viele verschiedene Kriterien verwendet, scheinbar ohne empirische Grundlage für ihre Notwendigkeit und ohne Evidenz dafür, dass sie zu einer adäquaten Stadienzuordnung führen.

In der ersten Studie dieser Dissertation wurde daher die Stadienzuordnung basierend auf einem häufig gebrauchten, externen Verhaltenskriterium für körperliche Aktivität (die Richtlinie der Weltgesundheitsorganisation an mindestens vier Tagen in der Woche für mindestens 30 Minuten körperlich aktiv zu sein) mit der Stadienzuordnung basierend auf einem individuellen Kriterium (subjektive Einschätzung der Teilnehmer bezüglich regelmäßiger körperlicher Aktivität) verglichen. Die Ergebnisse weisen darauf hin, dass die verschiedenen Kriterien einen Einfluss darauf haben, welchem Stadium ein Teilnehmer zugeordnet wird, d.h. dass es nicht unbedeutend ist, an welchem Kriterium das Verhalten und die Absichten der Menschen gemessen werden. Eine Stadieneinteilung, die sich auf das individuelle Kriterium stützt, scheint darüber hinaus die Mindsets der Teilnehmer besser zu repräsentieren als die Zuordnung, die sich auf den externen Standard stützt. Dies deutet darauf hin, dass eine Klassifikation ins *falsche* Stadium (d.h. ein Stadium, welches das Mindset inadäquat repräsentiert) möglich und wahrscheinlich ist, sofern ein ungünstiger Standard gewählt wird. Es scheint daher von Nutzen zu sein, sich an der subjektiven Einschätzung der Teilnehmer zu orientieren, wenn es darum geht, ihre Stadien zu erfassen.

Eine Einschränkung dieser ersten Studie ist, dass Mindsets lediglich durch Verhalten und Intention charakterisiert wurden (Begründung, siehe oben). Deshalb sind in einer zweiten Studie Mindsets als sozial-kognitive Profile erfasst worden. Diese wurden direkt aus den Daten extrahiert und anschließend mit den Stadien verglichen, welche das Verhalten und die Intention von Menschen widerspiegeln. Die Übereinstimmung war gering, was darauf hindeutet, dass der Stadienalgorithmus basierend auf Verhalten und Intention nicht zu einer Stadienzuordnung führt, welche Mindsets hinsichtlich eines Zielverhaltens vollständig und angemessen zu reflektieren scheint. Dies stellt in Frage, ob

als stadienpassend deklarierte Interventionsmaßnahmen dies wirklich sind, und erklärt möglicherweise, weshalb sich solche Maßnahmen gegenüber generischen Interventionen nicht immer als wirksam erweisen.

(2) Die Evaluation von Interventionen wird oft auf die in Theorien spezifizierten Prädiktoren und Mechanismen von Verhaltensänderung beschränkt. Je nach Ergebnis wird dann der Schluss gezogen, dass in der Intervention die jeweils adäquaten bzw. unpassenden Determinanten angesprochen wurden. Für eine erschöpfende Evaluation reicht dies jedoch meistens nicht aus. Der zweite Teil dieser Dissertation versucht deswegen dieses Defizit zu überwinden, indem sich der Frage gewidmet wird, inwieweit Interventionen von den Personen überhaupt empfangen werden.

Dazu wurde in einer Studie eine umfassende Definition des Konstrukts *Engagement* im Kontext von Gesundheitsförderungsmaßnahmen vorgeschlagen. In Einklang mit dieser Definition wurde darauf aufbauend eine neue Selbstberichtskala zur Erfassung von Engagement (die *Task Engagement Scale*, TES) entwickelt. Es wurde die faktorielle Struktur der Skala sowie die Beziehungen zu objektiven Indikatoren von Engagement und theoretisch verwandten Konstrukten untersucht. Es konnten vier Unterskalen identifiziert werden, die individuell zur Erfassung der Teilkomponenten von Engagement: Komplianz, Anstrengung, ungeteilte Aufmerksamkeit und Vertiefung eingesetzt werden können. Die TES hatte gegenüber den objektiven Indikatoren Zeit und Vervollständigungsrate mehr Informationsgehalt. Beziehungen zwischen theoretisch verwandten Konstrukten und dem TES Wert deuteten darauf hin, dass Engagement ausreichend distinkt ist. Damit wurde die Grundlage geschaffen, um Engagement zu erfassen und im Verhaltensveränderungsprozess zu berücksichtigen. Dies wurde in der anschließenden Studie gemacht.

(3) Im dritten Teil der Dissertation wurden beide Herangehensweisen kombiniert. Insbesondere wurde die Rolle des Interventions-Engagements im Prozess der Gesundheitsverhaltensänderung untersucht. Hierfür wurde eine experimentelle Interventionsstudie zum Thema Obst- und Gemüseverzehr durchgeführt. Der

Mechanismus, durch den die Intervention Einfluss auf Veränderungen im Obst- und Gemüseverzehr hatte, wirkte über Veränderungen in Planungskognitionen. Diese Mediation wurde moderiert vom Engagement der Teilnehmer. Dies bedeutet, dass nur wenn das Engagement der Teilnehmer bezüglich des Interventionsmaterials relativ hoch war, bewirkte die Intervention Veränderungen in den Planungskognitionen. Dieses Ergebnis demonstriert, dass theorie- und evidenzbasierte Interventionen allein möglicherweise nicht zu Veränderungen in den Variablen führen, die durch die Intervention verändert werden sollen. Die Implikation ist, dass Forscher und Anwender bei der Entwicklung von Gesundheitsförderungsmaßnahmen nicht nur Prädiktoren der Verhaltensänderung berücksichtigen sollten, sondern auch Faktoren, welche einen Einfluss auf das Interventions-Engagement haben könnten.

Im abschließenden Teil dieser Dissertation werden alle Befunde integrativ diskutiert und Implikationen für weiterführende Forschung abgeleitet. Insgesamt ist festzuhalten, dass diese Arbeit eine Hauptfragestellung der Gesundheitspsychologie, nämlich was den Erfolg von Verhaltensänderungsmaßnahmen bedingt, von zwei Seiten betrachtet und durch ein integratives Vorgehen zum besseren Verständnis und zum Fortschritt dieses Forschungsgebiets beiträgt. Konkret wurde gezeigt, dass Mindsets hinsichtlich eines Zielverhaltens nicht adäquat mit Hilfe von Stadien erfasst werden konnten, was die Angemessenheit und Passung der Inhalte von Gesundheitsförderung in Frage stellt. Des Weiteren wurde demonstriert, dass die Auseinandersetzung des Interventionsteilnehmers mit dem Interventionsmaterial nicht unbedeutend für den Erfolg von Gesundheitsförderung ist.

1

Introduction

Validity of Stages and the Role of Intervention Engagement in the Behavior Change Process

In this introductory chapter, the field of health behavior change is briefly introduced, thereby providing the theoretical framework of this dissertation and setting the rationale for the research questions investigated in the following empirical chapters.

Health Behaviors

All behaviors that are connected to *health* can be referred to as health behaviors. However, while *health promoting behaviors* are carried out "for the purpose of preventing or detecting disease or for improving health and well being" (Connor, 2001, p. 1), *risk behaviors* are behaviors that could be detrimental to one's health. Examples of health promoting behaviors are physical activity, a healthy diet, flossing, use of condoms, putting on seat belts, medical checkups and screenings, as well as the giving up of risk behaviors such as smoking and alcohol- or drug abuse (Schwarzer, 2008). Two health risk behaviors that operate as the leading cause of chronic morbidity and early mortality in the developed world and that are associated with reduced quality of life are *physical inactivity* and an *unhealthy diet* (Khaw et al., 2008).

Theories of Health Behaviors and Health Behavior Change

As individuals engage in lifestyles that are comprised of both health enhancing and health compromising behaviors, psychologists try to uncover the mechanisms that will help individuals engage in lifestyles that are attractive yet won't put their health at risk. In order to explain what drives behavior change, various health behavior theories can be drawn from. All available theories of health behavior change can be classified into the category *continuum theories* or the category *stage theories* (Biddle, Hagger, Chatzisaranitis, & Lippke, 2007; Schwarzer, 2008). Both types of theories will be described briefly below in order to provide a theoretical framework as well as offer a rationale for the aims of this dissertation.

Continuum Theories

Continuum Models propose a number of social-cognitive predictors of (health) behaviors. Depending on each predictor's value, individuals are assumed to have different likelihoods of engaging in the targeted behavior. In other words, individuals are located at different points of a continuum of probabilities. Interventions concentrate on at least one of these predictors with the objective of moving individuals further along the continuum (Schwarzer, 2008). In accordance with such theories, all individuals should benefit from the same intervention and this idea is expressed in the idiom "one size fits all" (Nigg, 2003; Kreuter, Stretcher, & Glasman, 1999).

The most well-known examples of continuum models are the Health Belief Model (HBM; Becker 1974; Rosenstock, 1966), the Protection Motivation Theory (PMT; Rogers, 1983), the Theory of Planned Behavior (TPB; Ajzen, 1985), the Social-Cognitive Theory (SCT; Bandura, 1997), and the continuum version of the Health Action Process Approach (HAPA, Schwarzer, 1992, 2008). A detailed description of these models can be obtained elsewhere (for meta-analytical review regarding HBM see Harrison, Mullen, & Green, 1992, regarding PMT see Milne, Sheeran, & Orbell, 2000; and for further information on the TPB see Conner, & Sparks, 2005; for a review see Richert, & Lippke, 2008).

Stage Theories

The central assumption of *stage theories of health behavior change* is that individuals differ in their *mindsets* (manifested in social cognitions) towards behavior change. Those with similar cognitions are said to be in the same *stage*. An important notion that distinguishes stage theories from continuum theories is that determinants of transitions towards behavior change are assumed to be differential for each stage. It follows that interventions need to be *stage appropriate* in order to be effective, addressing only those factors that are relevant for a particular stage and mindset (Sutton, 2000; Weinstein, Rothman, & Sutton, 1998). The most popular stage theories are the Transtheoretical Model (TTM, Prochaska & DiClemente, 1983), the Precaution Adoption Process Model (PAPM, Weinstein & Sandman, 1992) and the stage version of the Health

Action Process Approach (HAPA, Schwarzer, 1992, 2008).

Validity of Stages and Stage Theories

Tests of the validity of stage theories are diverse, all having different strengths and weaknesses (for an overview on different ways to test the validity of stages, cf., Sutton, 2000; Weinstein, Rothman, & Sutton, 1998). Evidence from these tests is inconclusive. Sometimes findings are interpreted in favor of stages (e.g., Armitage, Sheeran, Conner, & Arden, 2004, Armitage, Povey, & Arden, 2003; Schüz, Sniehotta, Mallach, Wiedemann, & Schwarzer, 2009; Lippke, Sniehotta, and Luszczynska, 2005; Lippke, Ziegelmann and Schwarzer, 2005; Sniehotta, Luszczynska, Scholz, & Lippke, 2005; Wiedemann, Lippke, Reuter, Schüz, Ziegelmann, & Schwarzer, 2009). Other times, findings are construed as challenging the notion of stage theories (e.g., Adams & White, 2005; Bridle et al., 2005; Conn, Hafdahl, Brown & Brown, 2008; Dijkstra, Conijn, & De Vries, 2006; Lippke, Schwarzer, Ziegelmann, Scholz, & Schüz, 2010; Noar, Benac & Harris, 2007; Weinstein, & Sandman, 1992).

Stage allocation

A lack of supportive results is often interpreted in favor of continuum models. However, such verdicts have wide implications for theory and practice, for example for the development of health promotion interventions. It may therefore be useful to take another look at these results and explore an alternative explanation that does not challenge the core concept of stage theories (Richert, Lippke, & Schwarzer, 2010).

To ensure that treatments are suitable for an individual and consequently effective, stages need to represent individuals' mindsets towards behavior change in an adequate manner. In other words, stage *allocation* needs to be valid and reliable (Lippke, Ziegelmann, Schwarzer, & Velicer, 2009). Thus far, this idea that misclassification (i.e., an inadequate representation of mindsets through stages) may be the cause of stage-specific interventions' failure to outperform generic interventions has been no concern for health psychologists and has not been addressed in previous research. One aim of this dissertation is therefore to bridge this gap in research by addressing this question of the validity of

stage allocation.

Criteria in Stage Measures

Measures used to allocate individuals into stages often contain some standard for behavior and intention that individuals can be compared against. Prominent standards that are used frequently are the guidelines on health behaviors advocated by the World Health Organization (for an example measure cf., Lippke, Ziegelmann, Schwarzer, & Velicer, 2009; Richert, Lippke, & Schwarzer, 2010). Although it seems intuitive that the criterion inherent in stage measures may affect subsequent stage allocation, a large variety of criteria is currently employed (cf., e.g., Lippke, Ziegelmann, Schwarzer, & Velicer, 2009). Which of these criteria will lead to an allocation that adequately reflects individuals' mindsets? This question draws awareness to the risk of misclassification should a wrong criterion be used. As misclassification undermines the possibility of stage appropriate health interventions, it is an issue worth addressing. The first study of this dissertation (*Chapter 2*) will therefore take a first step and compare two criteria in terms of their ability to lead to stage allocation that reflects individuals' mindsets correctly. As most stage measures rely on individuals' self-reports on behavior and intentions (the idea behind it being that these are correlated with social cognitions), mindsets will be represented by behavior and intentions to change.

Correspondence between mindsets and stages based on behavior and intentions

Though all stage theories of health behavior change share the same core ideas (see paragraph on stage theories), they differ in their assumptions on how many mindsets or stages exist (Warner, & Lippke, 2008). While the Transtheoretical Model, for example, proposes five stages of change (Prochaska & DiClemente, 1983), the Precaution Adoption Process Model assumes no less than six different stages (Weinstein, 1988). Rather than selecting one of these theories and relying on the accuracy of the 'a priori' specified stages as well as the validity of the employed stage measures, in a second study (*Chapter 3*), social-cognitive profiles (i.e., *mindsets*) are extracted directly from the data. This

empiricism-based approach to identifying stages avoids both theoretical and methodological issues of confirmatory approaches and consequently allows for a more accurate consideration of mindsets. Again, stages need to be reflective of individuals' mindsets in order to allow for adequate treatments, so in a second step, the stage measure used in study 1 is tested for correspondence. That is, the notion that stages based on reports about behavior and intentions are representative of mindsets is put to the test. Thereby this second study extends the investigation of the issue of the validity of stages.

Going beyond intervention content

Health psychologists often restrict their investigation of the effectiveness of health interventions to the determinants and mechanisms of behavior change. These variables and mechanisms can and should inform the content of an intervention (Michie, Johnston, Francis, Hardeman, & Eccles, 2008). However, researchers often incorrectly conclude that depending on the results of the intervention study, they either addressed the right or the wrong determinants of behavior change when it could just as well be that the participants did not read the materials (Richert & Lippke, 2010). Investigating to which degree an intervention is *received* by the target population would therefore allow for a more comprehensive and accurate evaluation of health promotion (Richert, Lippke, & Ziegelmann, in press). Another aim of this dissertation was to address this issue and thereby overcome the shortcoming of previous research.

Intervention Engagement

Engagement may briefly be explained as the time that is actively used to work on achieving a goal, such as participating in an intervention. It is more than being physically present. Engagement refers to the use of the provided time for its designated purpose and to the avoidance of distractions as well as engagement in goal unrelated actions (adapted from Singh, Granville, & Dika, 2002). Despite this definition, there is no consistent terminology or established definition of the construct in the literature (Guthrie et al., 2004). Likewise, several ways to measure engagement exist. These include direct observation,

(e.g., Amato-Zech, Hoff, & Doepke, 2006; Chafouleas, Riley-Tillman, Sassu, LaFrance, & Patwa, 2007), behavioral indicators such as time spent on task (e.g., Richert, Lippke, & Ziegelmann, 2010), self-report scales such as the Motivation and Engagement Scale (MES, Martin, 2007) and the Utrecht work engagement scale (UWES, Schaufeli, Martinez, Pinto, Salanova, & Bakker, 2002). However, there is no comprehensive and feasible measure of the construct in the context of health behavior change interventions.

To fill this gap, one endeavor was to provide an exhaustive definition of the construct as well as a parsimonious and feasible self-report measure that may help making more comprehensive intervention research possible. The factorial structure of the scale is examined and data on its relationship both with objective indicators of engagement and with theoretically related constructs provided in *Chapter 4*.

Generating an integrated understanding on what leads to health behavior change

The last aim of this dissertation (*Chapter 5*) is to integrate the construct of engagement into the behavior change process. Thus, it aims at contributing to a more accurate understanding on what is necessary for effective health behavior change. In particular, whether a theory- and evidence based health promotion intervention leads to changes in the cognitions it targets and subsequently to changes in behavior is examined. Additionally, whether intervention engagement affects this mechanism is investigated. The implications of such an influence may be that researchers and practitioners need to give consideration to factors that affect participants' behavior as well as their engagement in the intervention.

References

Adams, J., and White, M. (2005). Why don't stage-based activity promotion interventions work? *Health Education Research, 20*(2), 237-243. doi:10.1093/her/cyg105

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl and J. Beckmann (Eds.), *Action control: From cognition to behavior.* (pp. 11-39). Berlin: Springer.

Amato-Zech, N. A., Hoff, K. E., & Doepke, K. J. (2006). Increasing on-task behavior in the classroom: Extension of self-monitoring strategies. *Psychology in the Schools, 43*(2)*,* 211-221. doi:10.1002/pits.20137

Armitage, C. J., Povey, R., & Arden, M. A. (2003). Evidence for discontinuity patterns across the stages of change: A role for attitudinal ambivalence. *Psychology and Health, 18*(3), 373-386. doi:10.1080/0887044031000066553

Armitage, C. J., Sheeran, P., Conner, M., & Arden, M. A. (2004). Stages of Change or Changes of Stage? Predicting Transitions in Transtheoretical Model Stages in Relation to Healthy Food Choice. *Journal of Consulting and Clinical Psychology, 72*(3), 491-499. doi:10.1037/0022-006X.72.3.491

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Becker, M. H. (Ed.) (1974). The health belief model and personal health behavior. Thorofare, NJ: Slack.

Biddle, S. J. H., Hagger, M. S., Chatzisaranitis, N. L. D., and Lippke, S. (2007). Theoretical frameworks in exercise psychology. In G. Tenenbaum and R. C. Eklund (Eds.), *Handbook of Sport Psychology* (3rd edition) (pp. 537-559). New York: Wiley.

Bridle, C., Riemsma, R. P., Pattenden, J., Sowden, A. J., Mather, L., Watt, I. S., et al. (2005). Systematic review of the effectiveness of health behavior interventions based on the Transtheoretical model. *Psychology and Health, 20*(3), 283-301. doi:10.1080/08870440512331333997

Chafouleas, S. M., Riley-Tillman, T. C., Sassu, K. A., LaFrance, M. J., & Patwa, S. S. (2007). Daily Behavior Report Cards: An Investigation of the Consistency of On-Task Data Across Raters and Methods. *Journal of Positive Behavior Interventions, 9*(1), 30-37. doi:10.1177/10983007070090010401

Conn, V., Hafdahl, A., Brown, S., & Brown, L. (2008). Meta-analysis of patient education interventions to increase physical activity among chronically ill adults. *Patient Education and Counseling*, *70*, 157-172. doi:10.1016/j.pec.2007.10.004

Connor, M. (2001). Health behaviors. In P. B. Baltes & N. J. Smelser (Eds.), The International Encyclopedia of the Social & Behavioral Sciences (Vol. 10, pp. 6506-6512). Oxford, England: Elsevier Science.

Conner, M., and Sparks, P. (2005). Theory of planned behaviour and health behaviour. In M. Conner and P. Norman (Eds.). *Predicting health behaviours.* (pp. 170-222). London: Open University Press.

Dijkstra, A., Conijn, B, & De Vries, H. (2006). A match-mismatch test of a stage model in tobacco smoking. *Addiction, 101,* 1035-1043. doi:10.1111/j.1360-0443.2006.01419.x

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., et al. (2004). Increasing Reading Comprehension and Engagement Through Concept-Oriented Reading Instruction. *Journal of Educational Psychology, 96*(3), 403-423. doi:10.1037/0022-0663.96.3.403

Harrison, J. A., Mullen, P. D., and Green, W. (1992). A meta-analysis of studies of the Health Belief Model with adults. *Health Education Research, 7*(1), 107-116. doi:10.1093/her/7.1.107

Khaw, K. T., Wareham, N., Bingham, S., Welch, A., Luben, R., & Day, N. (2008). Combined impact of health behaviours and mortality in men and women: the EPIC-Norfolk prospective population study. *PLoS Med, 5*(1), e12. doi:10.1371/journal.pmed.0050012

Kreuter, M. W., Stretcher, V. J., and Glasman, B. (1999). One size does not fit all. *Annals of Behavioral Medicine, 21*, 276-283. doi:10.1007/BF02895958

Lippke, S., Schwarzer, R., Ziegelmann, J. P., Scholz, U., & Schüz, B. (2010). Testing stage-specific effects of a stage-matched intervention: a randomized controlled trial targeting physical exercise and its predictors. *Health Education & Behavior, 37*(4), 533-546. doi: 10.1177/1090198109359386

Lippke, S., Sniehotta, F. F., and Luszczynska, A. (2005). Social cognitions across the stages of behavior change. A comparison of two stage models. *Polish Psychological Bulletin, 36*(1), 43-50.

Lippke, S., Ziegelmann, J. P., and Schwarzer, R. (2005). Stage-specific adoption and maintenance of physical activity: Testing a three-stage model. *Psychology of Sport and Exercise, 6*, 585-603. doi:10.1016/j.psychsport.2004.11.002

Lippke, S., Ziegelmann, J. P., Schwarzer, R., & Velicer, W. F. (2009). Validity of stage assessment in the adoption and maintenance of physical activity and fruit and vegetable consumption. *Health Psychology, 28*(2), 183-193. doi: 10.1037/a0012983

Martin, A. J. (2007). Examining a multidimensional model of student motivation and engagement using a construct validation approach. *British Journal of Educational Psychology, 77*(2), 413-440. doi:10.1348/000709906X118036

Michie, S., Johnston, M., Francis, J., Hardeman, W., & Eccles, M. (2008). From theory to intervention: Mapping theoretically derived behavioural determinants to behavior change techniques. Applied Psychology: An International Review, 57(4), 660-680. doi:10.1111/j.1464-0597.2008.00341.x

Milne, S., Sheeran, P., and Orbell, S. (2000). Prediction and intervention in health-related behavior: A meta-analytic review of protection motivation theory. *Journal of Applied Social Psychology, 30*(1), 106-143. doi:10.1111/j.1559-1816.2000.tb02308.x

Nigg, C. R. (2003). Technology's influence on physical activity and exercise science: the present and the future. *Psychology of Sport and Exercise, 4*, 57-65. doi:10.1016/S1469-0292(02)00017-1

Noar, S. M., Benac, C., & Harris, M. (2007). Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychological Bulletin*, *133*(4), 673-693. doi:10.1037/0033-2909.133.4.673

Prochaska, J. O., and DiClemente, C. C. (1983). Stages and processes of self-change of smoking : Toward an integrative model of change. *Journal of Consulting and Clinical Psychology*, *51*, 390-395. doi:10.1037/0022-006X.51.3.390

Richert, J., & Lippke, S. (2008). Theoretical approaches (Health knowledge, attitudes, and practices - theoretical approaches for evidence-based practice. In R. W. Pierce & R. I. Schwartz (Eds.), *New Perspectives on Knowledge, Attitudes and Practices in Health* (pp. 133-156). Hauppauge, NY: Nova Science Publishers.

Richert, J., & Lippke, S. (2010). Assessment and Validation of the Task-Engagement Scale (TES). Manuscript submitted for publication.

Richert, J., Lippke, S., & Schwarzer, R. (2010). Comparison of individual criteria and externally imposed criteria for stage allocation: Findings from an internet study addressing physical activity, *Measurement in Physical Education and Exercise Science, 14,* 1–16. doi: 10.1080/1091367X.2010.520242

Richert, J., Lippke, S., & Ziegelmann, J. P. (in press). Intervention-engagement and its role in the effectiveness of stage-matched interventions promoting physical exercise. *Research in Sports Medicine: An International Journal.*

Rogers, R. W. (1983). Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. In J. R. Cacioppo and R. E. Petty (Eds.), *Social psychology: A sourcebook.* (pp. 153-176). New York: Guilford.

Rosenstock, I. M. (1966). Why people use health services. *Milbank Memorial Fund Quarterly*, *44*, 94-127. doi:10.2307/3348967

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and Science Achievement: Effects of Motivation, Interest, and Academic Engagement. *Journal of Educational Research, 95*(6), 323. doi:10.1080/00220670209596607

Schaufeli, W. B., Martinez, I. M., Pinto, A. M., Salanova, M., & Bakker, A. B. (2002). Burnout and Engagement in University Students: A Cross-National Study. *Journal of Cross-Cultural Psychology, 33*(5), 464-481. doi:10.1177/0022022102033005003

Schüz, B., Sniehotta, F. F., Mallach, N., Wiedemann, A. U., & Schwarzer, R. (2009). Predicting transitions from preintentional, intentional and actional stages of change. *Health Educ. Res., 24*(1), 64-75. doi: 10.1093/her/cym092

Schwarzer, R. (1992). Self-efficacy in the adoption and maintenance of health behaviours: Theoretical approaches and a new model. In R. Schwarzer (Ed.), *Self-efficacy: Thought control of action.* (pp. 217-243). Bristol, PA: Taylor and Francis.

Schwarzer, R. (2008). Modeling health behavior change: How to predict and modify the adoption and maintenance of health behaviors. *Applied Psychology (*57), 1-29.

Sniehotta, F. F., Luszczynska, A., Scholz, U., & Lippke, S. (2005). Discontinuity patterns in stages of the precaution adoption process model: Meat consumption during a livestock epidemic. *British Journal Of Health Psychology, 10*(2), 221-235. doi:10.1348/135910705X26137

Sutton, S. (2000). Interpreting cross-sectional data on stages of change. *Psychology and Health*, 15, 163-171. doi:10.1080/08870440008400298

Warner, L. M., and Lippke, S. (2007). Stage models of physical exercise. In F. Columbus (Ed.), *Sport and Exercise Psychology Research Advances*. Hauppauge, NY: Nova Science Publishers.

Weinstein, N. D. (1988). The precaution adoption process. *Health Psychology, 7*(4), 355-386. doi:10.1037/0278-6133.7.4.355

Weinstein, N. D., Rothman, A. J., and Sutton, S. R. (1998). Stage theories of health behavior: Conceptual and methodological issues. *Health Psychology, 17*(3), 290-299. doi:10.1037/0278-6133.17.3.290

Weinstein, N. D., and Sandman, P. M. (1992). A model of the precaution adoption process: Evidence from home radon testing. *Health Psychology*, *11*, 170-180. doi:10.1037/0278-6133.11.3.170

Wiedemann, A. U., Lippke, S., Reuter, T., Schüz, B., Ziegelmann, J. P., & Schwarzer, R. (2009). Prediction of Stage Transitions in Fruit and Vegetable Intake. *Health Education Research, 24*, 596-607. doi:10.1093/her/cyn061

2

Comparison of individual criteria and externally imposed criteria for stage-allocation: Findings from an internet study addressing physical activity.

Jana Richert, Sonia Lippke, & Ralf Schwarzer

Freie Universität Berlin, Germany

Richert, J., Lippke, S., & Schwarzer, R. (2010). Comparison of individual criteria and externally imposed criteria for stage allocation: Findings from an internet study addressing physical activity, *Measurement in Physical Education and Exercise Science, 14,* 1–16. doi: 10.1080/1091367X.2010.520242

Abstract

Stage-matched interventions can only be more effective than one-size-fits-all interventions if they target participants' specific needs. Therefore, individuals have to be allocated to a stage that truly reflects their mindsets. Various criteria for stage-allocation exist. This study's objective was to demonstrate the impact of different classification criteria on stage-allocation, and which criterion yields the best reflection of individuals' mindsets. The sample consists of 569 internet users. Physical activity, intention to change, and four stages of change (Non-Intender, Intender, Maintaining Actor, and Changing Actor) were assessed. Staging was based on two criteria: individual criteria and externally imposed criteria (30 minutes of physical activity on at least 4 days of the week). Hypotheses were tested by multivariate analyses and validity tests. As predicted, the different criteria affected stage-allocation, and staging based on the individual criterion seemed to reflect individuals' intention to change and behavior better than staging based on the external standard.

Keywords: Stage theories, diagnostic, stage-algorithm, matched interventions

Comparison of individual criteria and externally imposed criteria for stage-allocation: Findings from an internet study addressing physical activity

Stage theories such as the Transtheoretical Model (TTM, Prochaska & DiClemente, 1983), the Precaution Adoption Process Model (PAPM, Weinstein & Sandman, 1992) and the Health Action Process Approach (HAPA, Schwarzer, 1992, 2008) are of high intuitive appeal and have gained popularity over the past years (Brug et al., 2005). The central assumption of all existing stage models (in the field of exercise science and health psychology) is that individuals of a comparable mindset have psychological resources that are alike, encounter similar obstacles, and subsequently need the same kind of support when it comes to the adoption or maintenance of behaviors. Based on this notion, stage theories identify a number of stages that each represent a distinct mindset. Individuals are classified accordingly (for an overview, cf. Warner & Lippke, 2008; Weinstein, Rothman & Sutton, 1998). Treatments addressing stage-appropriate variables should be more effective and also more parsimonious than standard "one-size-fits-all" interventions because they specifically target the particular needs of the individuals. This assumption only holds true if individuals are allocated to the correct stage. If participants are misclassified and their allotted stage is not reflective of their mindset, the supposedly matched treatment will not be appropriate (Armitage, 2009).

Often, study results are not clearly supporting the notion of stage-matched interventions (Adams & White, 2005; Bridle et al., 2005; Conn, Hafdahl, Brown & Brown, 2008; Noar, Benac & Harris, 2007). Conclusions drawn from such findings are of vast theoretical and practical implications and are interpreted in favor of continuum models and against stage theories. When taking the possibility of misclassifications into account, these findings might be re-evaluated in a more accurate way.

Currently, a large assortment of criteria is employed to allocate people into stages and the risk of misclassification is given (e.g., Lippke, Ziegelmann, Schwarzer, & Velicer, 2009; Richert, 2009). Therefore, this paper will focus on the question of which criterion to

use for stage-allocation. We will address which criterion may adequately reflect individuals' mindsets, or, which criterion leads to the least misclassification.

In the present study, a four-stage model derived from the Health Action Process Approach (HAPA, Schwarzer, 1992, 2008) has been chosen. This was done because the model is deemed simple, sufficiently distinctive as well as practical (Schwarzer, 2008). The four-stage HAPA distinguishes between Non-Intenders, Intenders, Maintaining Actors, and Changing Actors. The latter two groups have only been studied as combined action stage in previous studies (Lippke, Ziegelmann & Schwarzer, 2005; Schwarzer, 2008). *Non-Intenders* are individuals who are not active with regard to a target behavior and have no intention to change. *Intenders* are people who are not active, but who have the intention to change. *Maintaining Actors* are those who are active with regard to the target behavior and who wish to maintain this state. In contrast, *Changing Actors* are individuals who are active and have an intention to change their behavior (cf. Schwarzer, 2008; cf. Table 2).

Stage-Allocation Criteria

Very few health behaviors are naturally dichotomous, such as receiving a flu shot or testing one's house for radon, (cf. Weinstein, Lyon, Sandman, & Cuite, 1998). In this case, the question of labeling individuals as either active or inactive with regard to the behavior in question is unequivocal. However, most health behaviors are not that simple. The majority of preventive actions are not dichotomous but rather continuous. For example, individuals are not either physically active or sedentary; they are rather located on a continuum of physical activity behavior.

When dealing with continuous behaviors, the introduction of a well-defined cut-off score is useful, as it allows for grouping of individuals. This grouping is of practical relevance (e.g., in the field of stage-based health promotion interventions). Sometimes, the criteria used are of medical importance. That is their attainment is necessary to generate desired health outcomes. So the classification of individuals as active or inactive with regard to this indisputable criterion is absolutely legitimate.

Often, benchmarks or standards are advocated by renowned institutions and taken at face value by researchers and practitioners alike, even when empirical evidence for the medical exigency or expediency is lacking. For example, the World Health Organization (WHO), the National Health Service (NHS) and the German Robert Koch Institut (RKI) recommend engaging in physical activities (defined as somewhat exhausting) on four to seven days of the week, for at least 30 minutes (National Health Service, 2009; Robert Koch Institut & Statistisches Bundesamt, 2005; World Health Organization, 2009). Conducting experimental, long-term studies to test whether precisely this, and only this, particular amount of exercise is beneficial whereas ten minutes less will be worthless, is impossible. Thus, apart from all major health institutions having agreed on one criterion, there is great variety of criteria used in studies dealing with stages (most without evidence for the meaningfulness of the employed criterion).

If such unverified benchmarks are merely recommended standards, they can be a useful tool in guiding people in their goal setting. It is highly problematic, however, if they are used for dichotomization of continuous behaviors and rigorous classification of individuals as either active or inactive (and subsequently allocating them to stages). The following example (Table 1) shall illustrate this issue.

Table 1. Active or Inactive? Classification of five individuals with different behavior routines, according to three commonly used criteria.

		Classification Criteria		
		3 x 30 minutes	5 x 30 minutes	Total of 180 minutes
actual behavior routines	Individual A 3 x 30 minutes	✓		
	Individual B 5 x 30 minutes	✓	✓	
	Individual C 6 x 30 minutes	✓	✓	✓
	Individual D 1 x 85 minutes			
	Individual E 2 x 90 minutes			✓

Note. Checked boxed indicate that individuals are considered active with regard to the criterion. Take note that although the objective behavior does not change for each individual, the classification of active vs. inactive changes drastically, depending on the classification criteria used.

Though they have different behavioral routines, all individuals in the example are fairly active, exercising on average for about 130 minutes per week (Individuals A, B, C, D, and E). However, individual D, who exercises for 85 minutes every week, is regarded as completely inactive, because she does not reach the aforementioned criteria. Also, while the objective level of behavior remains the same in all three cases of classification criteria, the decision of whether or not individuals are active varies greatly. So in one case a person is considered physically active and in another that very same person is regarded as not active (cf. Lamb & Joshi, 2004).

This example illustrates that external stage-allocation criteria may easily lead to misclassification because stage-allocation might not be representative of individuals' behavior and mindset. Therefore, one aim of this study is to provide empirical evidence for this notion. Thus far, the idea that misclassification may be the cause of stage-specific treatments' failures to outperform generic interventions has been no concern for exercise and health psychologists and has not been addressed in previous studies.

In this study, a subjective, individual criterion is introduced as an alternative criterion for stage-allocation. For this, individuals are asked whether they believe they are physically active on a regular basis and whether or not they have an intention to change their previously performed behavior (also cf. Methods section). In the following, the individual criterion is compared to the most prominent and widely used external criterion for defining people as physically active today: moderate physical activity for at least thirty minutes on 4 to seven days of the week.

If the criterion employed for stage-allocation was irrelevant, then looking at a stage by stage matrix, the distribution of cell sizes should ideally be like the one displayed in Table 2. That is, individuals allotted to a stage assessed with one criterion should fall into the same category based on any other criterion (stage-congruence).

Table 2. Theoretical distribution of cell sizes, if stage allocation criteria did not matter.

		individual criterion			
		Non-Intender	Intender	Maintainer	Changer
≥ 30min on 4-7 days of the week - criterion	Non-Intender	100%			
	Intender		100%		
	Maintainer			100%	
	Changer				100%

Note. Maintainer, Maintaining Actor; Changer, Changing Actor
The one hundred percent figure in the table is an ideal used for illustrative purposes, not accounting for misclassification. Thus, the counterpart of zero percent is to be seen as interchangeable with frequencies expected based on mere chance.
Frequencies in cells surrounding the diagonal would have to be lower than or equal to frequencies expected based on mere chance.

In empirical terms, Table 3 visualizes the distribution of cell sizes as is expected when stage-allocation criteria matter. Some cells are a logical discrepancy and should therefore not be occupied. It is illogical to report a difference between one's performed and one's intended physical activity behavior, measured as frequency and duration, while at the same time denying having an intention to change one's behavior. It should, therefore, be impossible for self-declared Non-Intenders to be classified as Intenders or Changing Actors with regard to some external criterion. The same antilogy holds true for all six cells bearing a zero. Likewise, although theoretically possible, individuals perceiving themselves as inactive (either Non-Intenders or Intenders based on the individual criterion) will most likely not be considered active with regard to the external criterion. In the remaining eight cells, however, frequencies should be higher than what would be expected based on chance.

Table 3. Theoretical distribution of cell frequencies, if stage allocation criteria do have an effect.

		individual criterion			
		Non-Intender	Intender	Maintainer	Changer
≥ 30min on 4-7 days of the week –	Non-Intender	>	>	>	>
	Intender	0	>	0	>
	Maintainer	~	0	>	0
	Changer	0	~	0	>

Note. Maintainer, Maintaining Actor; Changer, Changing Actor.
Zeros, found frequencies should be significantly lower than or equal to frequencies expected based on mere chance; ~, although theoretically possible, it is highly unlikely that these cells will be occupied. Frequencies in the remaining eight cells (marked with >) are hypothesized to be higher than frequencies expected based on chance.

If the notion that criteria are highly important for stage-allocation (and thus have vast implications for further treatment of individuals) is confirmed, then the question of which criterion to employ for classification of people into stages arises.

Subjective evaluation (individual criterion) as an alternative to external criteria

The idea behind stage-allocation is that people are grouped according to behavior and mindset and then can receive treatments with appropriate content (matched to their specific behavioral and psychological baseline). Consequently, the question of which criterion to employ for classifying people is really a question of which criterion best represents individuals' behavior and state of mind.

Theoretically, Non-Intenders and Maintaining Actors have no intention to change their current level of physical activity. Thus, differences between their reported weekly minutes of activity and their intended total weekly minutes of physical activity should not be significant. Likewise, Intenders and Changing Actors theoretically have an intention to change their current level of physical activity. Thus, the differences between their reported and their intended total weekly minutes of physical activity should be significant. If the external criterion was reflective of this notion, then it should hold true for all individuals allocated to these stages based on the external criterion, regardless of other factors. However, if the individual criterion is truly representative of an individual's mindset,

differences between reported and intended total weekly minutes should become significant for all individuals labeled as Intenders and Changing Actors. Vice versa, non-significant differences are expected for Non-Intenders and Maintaining Actors regardless of the stage they are allocated to based on the external criterion.

With regard to the target behavior, Non-Intenders and Intenders are by definition inactive while Maintaining Actors and Changing Actors are active (Schwarzer, 2008). When employing an external criterion, individuals are judged based on reports of their actual and intended physical activity. So, differences between the total weekly minutes of physical activity reported by individuals in inactive stages and of individuals in active stages is artificially created and should not have any meaning read into. The question of interest is whether inactive and active stage groups, based on the individual criterion, differ with regard to their weekly minutes of physical activity within the inactive stage groups based on the external criterion. If those in the active stages show higher levels of physical activity, this is indicative of the notion that the individual criterion is the criterion that better reflects individual behavior.

One concern about not giving guidelines as to which frequency and duration of physical activity should be reached in order to consider one physically active, is that individuals with extremely low levels of physical activity will claim they are active. But just how reasonable is a person's self-perception (In other words, how active are individuals, who view themselves as being physically active?).

Hypotheses

Based on these theoretical considerations, we had the following hypotheses.

Hypothesis 1. Individuals' stage-allocation differs depending on the criterion used. Indicative are observed frequencies (number of individuals) in four out of 12 stage-incongruent cells that are higher than what is expected based on chance (Table 2 and Table 3).

Hypothesis 2. Stages based on the individual criterion reflect individuals' actual behavior and their intention to change better than the stages based on the external criterion.

Method

Procedure

Via eMail, 2,600 individuals who had partaken in a previous study conducted by the authors' home institution were invited to participate in an online study[1]. Additionally, a link was placed on the department's homepage. The study's webpage was accessible for two months. In total, 655 individuals gave informed consent and were included in the study. Of those, 76 were excluded from the analyses because they had more than 50% missing data and were subsequently judged to not have taken their participation seriously. Furthermore, 10 individuals were excluded because they had a missing value in either one of two variables necessary for stage-allocation based on the individual criterion. Of all cases in the sample, 97.7% had less than 5% missing data, which were missing completely at random: χ^2 (1, 563) = 1521.42, $p > .05$ and were imputed with the Expectation Maximization (EM) algorithm (Tabachnick & Fidell, 2007).

Participants

The final study sample consisted of 569 individuals. The majority was female (78.6%) and on average, partakers were $M = 38.43$ years old ($SD = 12.57$), highly educated (75.8% with a 12-year school education: 47.1% with a university degree), employed (62.9%) and living in a partnership (57.7%).

Measures

Physical Activity. Physical activity was defined and explained to participants as any intentionally performed, somewhat exhausting, physical activity. This is equivalent to domain unspecific physical activities (i.e., all leisure-time and household activities, as well as activities carried out as means of transport) that are performed deliberately and that are of at least moderate intensity (for the rationale behind this, cf. Biddle, Goudas, & Page, 1994). Following the validated Godin Scale (Godin & Shephard, 1985), items assessing individuals' aggregated physical activity read: "Think of the recent past. On average, (1)

[1] The questionnaire was programmed with the software dynQuest (Rademacher & Lippke, 2007).

on how many days of the week . . . and (2) for how many minutes per day… have you been physically active?" Answers were given in pull-down answer format (one to seven days of the week and zero to 300 minutes per day). In a yes/no answer format, participants were also asked: "Are you physically active on a regular basis?" No external criterion/standard was given as to which frequency and duration constitutes 'active' or 'regular'.

Intention. Intention was assessed analogously. Participants were asked to think about the near future and report the intended frequency (in days per week) and duration (in minutes per day) of their physical activity.

Intention to change. Intention to change was then defined as positive deviation of an individual's current (reported) physical activity from his/her intended physical activity. Participants were also asked to judge (in a yes/no format) whether or not they wanted to change something about their current physical activity (in the way of more often, longer, new activities).

Stage. Stage was then derived from individual reports about their behavior and intention to change, either based on the external criterion (cf. Table 4) or the individual criterion respectively (cf. Table 5).

Table 4. Stage allocation based on individuals' behavior and intention to change according to the external criterion.

External criterion: WHO standard of at least 30 minutes of physical activity on 4-7 days of the week			behavior		
			no		yes
			does not engage in physical activity for at least 30 minutes on 4-7 days of the week		engages in physical activity for at least 30 minutes on 4-7 days of the week
intention to change	no	intention to be physically active on < 4 days of the week **or** for < 30 minutes on 4-7 days of the week	Non-Intender	negative or no deviation of physical activity from intended physical activity	Maintaining Actor
	yes	intention to be physically active for at least 30 minutes on 4-7 days of the week	Intender	positive deviation of physical activity from intended physical activity	Changing Actor

Note. In this study, both days per week and duration per day were considered for stage allocation as opposed to simply calculating a week sum score (cf., WHO, 2009).

Table 5. Stage-allocation based on reported behavior and intention to change according to participants' subjective evaluation (the individual criterion).

individual criterion			behavior	
			"Are you physically active on a regular basis?"	
			no	yes
intention to change	"Do you intend to change anything about your physical activity?"	no	Non-Intender	Maintaining Actor
		yes	Intender	Changing Actor

Socio-demographics. Age, gender, educational status, marital status and employment status were assessed as well.

Statistical Procedures

The hypothesis that individuals do not belong in a stage per se, but rather that the employed criterion will affect stage allocation, is tested by investigating adjusted standardized residuals. These can be understood as deviations of observed frequencies

from those that are expected based on chance alone. In addition, these deviation values are adjusted to the cell size and are approximately normally distributed. So, if values exceed 1.96, 2.58, or 3.29, respectively, the number of cases found in that cell is significantly higher than would be expected based on mere chance; whereas values less than -1.96, -2.58, or -3.29, respectively, indicate that the number of cases in that cell is significantly lower than expected (Agresti, 2002).

Whether or not an individual's intended total weekly minutes of physical activity differed significantly from their rendered total weekly minutes, and whether or not individuals in inactive stages differed significantly from individuals in active stages with regard to their level of physical activity, is tested via *t*-tests. To account for the alpha error accumulation in multiple comparisons, the Bonferroni correction was applied (Abdi, 2007). All analyses were performed using SPSS 17.

Results

Descriptive Results

According to the external criterion, 249 individuals (43.8% of the total sample) were Non-Intenders, 114 participants (19.6%) were Intenders, 75 participants (13.2%) were Maintaining Actors, and 131 individuals (23.0%) were Changing Actors. Based on the individual criterion, only 14 individuals (2.5% of the total sample) were Non-Intenders, 136 participants (23.9%) were Intenders, 147 individuals (25.8%) were Maintaining Actors, and 272 participants (47.8%) were Changing Actors (Table 6). For the majority of participants, incongruence between their stages was found. 94.8% of all participants having been classified as Non-Intenders based on the external criterion (n = 249), were allocated to a different stage based on the individual criterion.

Results for Hypothesis 1

Looking at the adjusted standardized residuals, the above described incongruence in stage-allocations is tested statistically. The cell containing individuals who are Non-

Intenders based on the external criterion and Intenders based on the individual criterion represents such a stage-incongruence. A total of 96 individuals were observed, while 60 were expected (cf. Table 6). The adjusted standardized residual of 7.2 is significant at the $p < .001$ level. This indicates that the cell is occupied more frequently than would have been expected based on chance. Likewise, the cell containing individuals who are Intenders based on the external criterion and Changing Actors based on the individual criterion ($n_{observed} = 78$, $n_{expected} = 55$) is occupied significantly more frequently than expected based on chance (adjusted standardized residual = 4.9, $p < .001$). No individuals (i.e., no more than are expected based on chance), who were externally classified as Intenders, considered themselves Intenders as well (stage-congruence: $n_{observed} = 29$, $n_{expected} = 27.2$, adjusted standardized residual = 0.4, $p > .05$).

All cells that were hypothesized to be vacant were found to have frequencies equal to or lower than frequencies expected based on chance, and both cells anticipated to be vacant (although theoretically having the potential to be occupied) were showing frequencies equal to or lower than what would have been expected based on chance (see Table 6). Overall, for three out of four hypothesized stage-incongruent cells observed frequencies of individuals were higher than what was expected based on chance.

Table 6. Stage-allocation based on the external criterion: ≥ 30min on at least 4 days of the week by stage-allocation based on subjective evaluation (the individual criterion).

			individual criterion (ic)				
			NI	I	MA	CA	total
external criterion (ec) ≥ 30min on 4-7 days of the week	NI	observed *N*	13	96	58	82	249
		expected *N*	6	60	64	119	
		stand. adjusted residuals	3.7**	7.2**	-1.2	-6.3**	
		% within all based on ic	92.9%	70.6%	39.5%	30.1%	
		% within all based on ec	5.2%	38.6%	23.3%	32.9%	100
		% of total *N*	2.3%	16.9%	10.2%	14.4%	43.8
	I	observed *N*	0	29	7	78	114
		expected *N*	3	27	30	55	
		stand. adjusted residuals	-1.9	0.4	-5.4**	4.9**	
		% within all based on ic	0%	21.3%	4.8%	28.7%	20
		% within all based on ec	0%	25.4%	6.1%	68.4%	100
		% of total *N*	0%	5.1%	1.2%	13.7%	19.6
	MA	observed *N*	0	0	60	15	75
		expected *N*	2	18	19	36	
		stand. adjusted residuals	-1.5	-5.2**	11.5**	-5.2**	
		% within all based on ic	0%	0%	40.8%	5.5%	13.2
		% within all based on ec	0%	0%	80%	20%	100
		% of total *N*	0%	0%	10.5%	2.6%	13.2
	CA	observed *N*	1	11	22	97	131
		expected *N*	3	31	34	63	
		stand. adjusted residuals	-1.4	-4.7**	-2.7*	6.9**	
		% within all based on ic	7.1%	8.1%	15%	35.7%	23
		% within all based on ec	0.8%	8.4%	16.8%	74%	100
		% of total *N*	0.2%	1.9%	3.9%	17%	23.0
	total		14	136	147	272	569

Note. NI, Non-Intender; I, Intender; MA, Maintaining Actor; CA, Changing Actor; stand. adjusted residuals, standardized adjusted residuals; ic, individual criterion; ec, external criterion; expected *N*s are rounded up as frequencies refer to people; * $p < .01$, ** $p < .001$

Results for Hypothesis 2.

Intention to change. Table 7 displays the results of dependent *t*-tests for each stage-combination, comparing reported physical activity (in minutes per week) with intended physical activity (in minutes per week). A significant difference between performed and intended physical activity describes an intention to change.

Table 7. Physical activity and intended physical activity (displayed as total weekly minutes = twm)

			indivial criterion (ic)				
			NI	I	MA	CA	Total
		twm of pa	51.15	50.31	198.10	129.15	110.26
		intended twm of pa	58.08	130.31	191.98	191.46	161.04
		t-value	-1.90	**-9.96****	0.52	**-8.66****	
	NI	*N*	13	96	58	82	
		Df	12	95	57	81	
		R	.48	.71	.07	.69	
		Cohen's *d*	-1.10	-2.04	0.14	-1.92	
		twm of pa	n.a.	74.83	291.43	159.04	145.30
		intended t twm of pa		250.00	284.29	276.03	269.91
		t-value		**-6.73****	0.09	**-8.94****	
external	I	*N*		29	7	78	
criterion		*Df*		28	6	77	
(ec)		*R*		.79	.04	.72	
		Cohen's *d*		-2.54	0.07	-2.04	
≥ 30min		twm of pa	n.a.	n.a.	498.67	695.00	543.11
on at		intended twm of pa			386.17	551.00	419.13
least		*t*-value			**2.84***	1.91	
4 days of	MA	*N*			60	15	
the week		*Df*			59	14	
		R			.35	.45	
		Cohen's *d*			0.74	1.02	
		twm of pa	200.00	306.36	305.00	367.84	352.59
		intended twm of pa	450.00	490.36	363.18	543.92	508.32
		t-value	n.a.	**-3.07**$^{+}$	**-3.46***	**-7.29****	
	CA	*N*	1	11	22	97	
		Df		10	21	96	
		R		.70	.60	.60	
		Cohen's *d*		-1.94	-1.51	-1.49	
		twm of pa	61.79	76.25	341.22	229.35	
	total	intended twm of pa	86.07	184.93	301.26	361.23	
		N	14	136	147	272	

Note. NI, Non-Intender; I, Intender; MA, Maintaining Actor; CA, Changing Actor; twm of pa, total weekly minutes of physical activity; n.a., not available; ic, individual criterion; ec, external criterion; $^{+}p < .05$, $*p < .01$, $**p < .001$; a Cohen's *d* of 0.3 or $r \geq .1$ denotes a small effect, $d \geq 0.5$ or $r \geq .3$ shows a medium sized effect and $d \geq 0.8$ or $r \geq .5$ denotes a large effect.

Individuals, who are Non-Intenders based on the external criterion, and Intenders based on the individual criterion are, on average, physically active for 50.31 minutes per week. They intend to be physically active for 130.31 minutes per week. This difference is statistically significant (t = -9.96, p < .001, d = -2.04), indicating an intention to change. All other cells can be read in the same manner.

Overall, the result is that all Intenders and Changing Actors based on the individual criterion, regardless of their stage diagnosed with the external criterion, have an intention to change. For stages based on the external criterion, there is no consistent picture. For example, not all individuals, who are classified as Intenders were found to have an intention to change.

Behavior. Overall, participants having been allotted to either of the inactive stages based on their own evaluation (Non-Intenders and Intenders) reported to be physically active for M = 74.90 (SD = 120.83) total weekly minutes. Individuals allocated to either of the inactive stages based on the external criterion (Non-Intenders and Intenders) reported to be physically active for M = 121.74 (SD = 108.62) total weekly minutes. This difference is statistically significant: t = 4.30, p < .001, d = -0.4.

Overall, individuals allocated to either of the active stages based on the individual criterion (Maintaining Actors and Changing Actors) are on average physically active for M = 284.64 (SD = 278.39) total weekly minutes. The lowest reported amount of physical activity was 20 minutes per week, reported by n = 4 or 0.9% of all Actors based on the individual criterion.

Table 8. Physical activity (displayed as total weekly minutes) for the inactive and active groups based on the individual criterion within both inactive groups based on the external criterion.

			individual criterion (ic)		
			inactive	active	
external criterion (ec)	NI	twm of pa	50.41	157.71	110.26
		N	109	140	
		t-value	**-9.16****		
		Df	274		
		R	-.52		
		Cohen's *d*	-1.21		
≥ 30min on at least 4 days of the week	I	twm of pa	74.83	169.94	145.30
		N	29	85	
		t-value	**-4.28****		
		Df	111		
		R	-.47		
		Cohen's *d*	-1.07		
	all	twm of pa	74.90	284.64	

Note. NI, Non-Intender; I, Intender; all, over all four stages; twm of pa, total weekly minutes of physical activity; ***p* < .001; a Cohen's *d* of 0.3 or *r* ≥ .1 denotes a small effect, *d* ≥ 0.5 or *r* ≥ .3 shows a medium sized effect and *d* ≥ 0.8 or *r* ≥ .5 denotes a large effect.

Table 8 displays the total weekly minutes of physical activity of both inactive and active individuals based on the individual criterion within the groups of inactive individuals based on the external criterion. Drawing attention to the first row showing Non-Intenders based on the external criterion, the difference between the inactive and the active stage group's total weekly minutes is significant: $t = -9.16$, $p < .001$, $d = -1.21$. The same holds true for the row showing Intenders based on the external criterion: $t = -4.28$, $p < .001$, $d = -1.07$.

Discussion

The present study aimed at demonstrating how large an impact classification criteria have on stage-allocation. Secondly, an alternative to the external criteria commonly used was presented. Empirical evidence for the legitimacy of this individual criterion was given.

Fewer external Non-Intenders than expected, based on chance, considered themselves active. However, this can be seen as a support of the notion that self-perception is sensible and not unreasonable. Overall, frequencies in stage-incongruent cells that are theoretically plausible were higher than what would be expected based on chance. This result goes against the notion that criteria are irrelevant. Subsequently, these findings support the hypothesis that criteria do matter.

Individuals in the three gray cells (Table 7) have been misclassified (cf. Table 3) and will not be interpreted. Looking at the rows of the matrix, focusing on the Non-Intenders in particular, no significant differences between performed and intended total weekly minutes of physical activity should be found if the external criterion was truly reflective of intention to change. Instead, what is observed is where participants' subjective evaluation (the internal criterion) led to the classification of Intenders and Changing Actors, the differences between performed and intended total weekly minutes of physical activity did become significant, yielding large effect sizes. The fact that the difference between physical activity and intended behavior became significant for individuals allocated to the maintaining actor stage based on both criteria seems at odds, but can easily be explained. Changing Actors are defined as individuals intending to do more in terms of frequency or duration or intending to change something else about their behavior, such as the type of activity (e.g., seasonal changes, such as wanting to go swimming rather than running). Unexpectedly, some individuals in the sample reported wanting to engage in less activity in the future. While their wish to exercise less is obviously not the same as wanting to maintain a current level of activity, it is also not the same as intending to do more. In anticipation of the consequences that a given treatment might have (a Changing Actor treatment helping them plan less activity versus relapse prevention for Maintaining Actors), these individuals were classified as Maintaining Actors. This decision is debatable. In future research, an alternative approach might be taken. The given explanation makes clear why the found difference did become significant contrary to the assumption that it would not. The effect was only medium-sized. Those participants who

were considered being inactive based on the external criterion were significantly more active than those who were considered inactive based on the individual criterion.

The data of this study illustrates the extent of misclassification caused by the external criterion, i.e., allocating individuals to a stage that does not fit their self-perception and that reflects neither their physical activity nor their intention to change. The consequence, as has been stated before: individuals receive treatments that are supposedly matched to their stage. But in fact, they are not, and, as a consequence, the treatment will not be effective (Adams & White, 2005, Bridle et al., 2005; Conn, Hafdahl, Brown & Brown, 2008; Noar, Benac & Harris, 2007). Results also indicate that the individual criterion leads to stage-allocation which reflects an individual's behavior and their intention to change better than the external criterion.

One concern about letting people judge whether or not they are active was that a person's subjective evaluation may be unreasonable (in the way that even very low amounts of activity might lead to the conclusion that one is active). However, considering that Actors based on the individual criterion reported to be deliberately physically active for roughly five hours a week and that only less than 1% of these Actors reported rather low levels of physical activity (20 minutes per week), this concern can safely be discounted.

The message of this study is to focus on subjective evaluation rather than external standards when classifying people into stages as this might be a better reflection of individuals' mindset. Even if convincing evidence is presented that a recommended standard is necessary to promote and uphold health, such findings are no reason to use the external standard as classification criterion. Such evidence would only indicate that the standard is worth being advocated. Interventions should assist individuals so they will comply with this standard, while still acknowledging their self-perceived state and addressing variables that are in accordance with their mindset.

Limitations

Sensitivity and Specifity. When dealing with any diagnostic instrument, looking at the sensitivity (i.e., the proportion of individuals correctly classified as being active out of all individuals, who are active) and the specificity (i.e., the proportion of individuals correctly classified as being not active out of all individuals, who are not active) of the measure is useful (Lippke et al., 2009). In this study, looking at sensitivity and specificity with regard to the stage measure based on the individual criterion seemed inappropriate as some reference criterion is needed to determine both. Using any external standard to check stages based on the individual criterion against, however, is contradictory to the idea of subjective evaluation. Also, as individuals were classified into stages based on the external criterion by checking their performed and intended physical activity (cf. Table 4); there was no possibility of misclassification with regard to this external standard. In the future, a more objective behavioral measure and distal health outcomes may be employed to test misclassification in this sense. As was explained in the introduction, misclassification here was defined as not matching individuals' mindset rather than not fitting some external standard.

The validity of the stage measure based on the individual criterion. In this article, we took the attitude that stage-allocation needs to be representative of individuals' subjective mindsets (i.e., their reported intention and behavior) rather than their objective behavior. Therefore, the question as to what constitutes 'active' or 'regular' physical activity was deliberately defined as subjective and left to personal understanding of the participants. No external standard or criterion was given. This was done with the aim to not influence participants. It might be a problem that such subjective perceptions can neither be assessed with, nor validated against, objective measures. In the future, the employed stage measure based on the individual criterion will have to prove its legitimacy by showing predictive validity by demonstrating a better fit between stages and stage-matched treatments in terms of desired outcomes.

Self-report measures of physical activity. Self-report measures allow for data collection from a large number of people at very low cost, and evidence supports the validity of self-reports for physical activity (e.g., Miller, Freedson, & Kline, 1994). In future studies, employing objective measures might be considered. However, it should be kept in mind that these measures are rather limited in their scope of activity assessment. For example, pedometers cannot account for activities in water.

Generalizability to other stage-models. This study specifically employed a four stage model (derived from the Health Action Process Approach, Schwarzer, 2008) as basis for all analyses and only the two variables deemed most significant to stage definition were analyzed: intention to change and behavior. The answer to the question of which criterion to use for stage-allocation is inherently linked to these decisions.

Nonetheless, the basic principle that was laid out in this paper should hold true for all stage models no matter the differentiation made. That is, criteria do have an impact on stage-allocation, and the individual criterion better differentiates between active and inactive as well as intending to change and not intending to change.

Conclusion and Outlook

This study questioned researchers' current approach to stage-diagnostics (i.e., stage-allocation) and provides first empirical evidence for an alternative method. The paper offers an explanation for the lack of supportive results in the field of stage-matched interventions. By this means, it contributes to the advancement of exercise science and health psychology, as well as exercise and health promotion alike. Future research has to take on this issue and provide (further) evidence for the adequacy or superiority of any criterion used to classify people into stages. Investigating all common social-cognitive variables that make up a psychological mindset was beyond the scope of this article. However, such research is an endeavor, whose results would strengthen the statements made here. If a criterion truly reflected the theoretically proposed mindset, then interventions that are matched to stages obtained based on that criterion would be truly matched. Thus, they should be most effective. The ultimate test is the experimental one:

interventions matched to participants' stage based on one criterion are more effective than interventions matched to participants' stage based on another criterion.

References

Abdi, H. (2007). Bonferroni and Šidák corrections for multiple comparisons. In N.J. Salkind (Ed.), *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage. http://www.utdallas.edu/~herve/Abdi-Bonferroni2007-pretty.pdf.

Adams, J., & White, M. (2005). Why don't stage-based activity promotion interventions work? *Health Education Research*, *20*(2), 237-243.

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). New York: Wiley.

Armitage, C. J. (2009). Is there utility in the Transtheoretical Model? *British Journal of Health Psychology*, *14*, 195-210.

Biddle, S., Goudas, M., & Page, A. (1994). Social-psychological predictors of self-reported actual and intended physical activity in a university workforce sample. *British Journal of Sports Medicine, 8*, 160-163.

Bridle, C., Riemsma, R. P., Pattenden, J., Sowden, A. J., Mather, L., Watt, I. S., et al. (2005). Systematic review of the effectiveness of health behavior interventions based on the Transtheoretical model. *Psychology & Health*, *20*(3), 283-301.

Brug, J., Conner, M., Harré, N., Kremers, S., McKellar, S., & Whitelaw, S. (2005). The Transtheoretical Model and stages of change: A critique: observations by five commentators on the paper by Adams, J. and White, M. (2004). Why don't stage-based activity promotion interventions work? *Health Education Research, 20*(2), 244-258.

Conn, V., Hafdahl, A., Brown, S., & Brown, L. (2008). Meta-analysis of patient education interventions to increase physical activity among chronically ill adults. *Patient Education and Counseling*, *70*, 157-172.

Godin, G., Shephard, R. J. (1985). A simple method to assess exercise behavior in the community. *Canadian Journal of Applied Sports Science,* 10, 141-146.

Kreuter, F., & Beddo, F. (2004). A handbook of statistical analyses using SPSS, *Journal of Statistical Software, 11*(02). [Electronic Version]. Retrieved 4.12.2009 from http://www.jstatsoft.org/

Lamb, R., & Joshi, M. S. (2004). Active but not consistent: Dietary behavior and the Stage of Change Model. *Psychology & Health, 19*(5), 543-559.

Lippke, S., Ziegelmann, J. P., & Schwarzer, R. (2005). Stage-specific adoption and maintenance of physical activity: Testing a three-stage model. *Psychology of Sport & Exercise, 6*, 585-603.

Lippke, S., Ziegelmann, J. P., Schwarzer, R., & Velicer, W. F. (2009). Validity of stage assessment in the adoption and maintenance of physical activity and fruit and vegetable consumption. *Health Psychology, 28*(2), 183-193.

Miller, D.J., Freedson, P.S., & Kline, G.M. (1994). Comparison of activity levels using the Caltrac accelerometer and five questionnaires. *Medicine and Science in Sports and Exercise*, 26, 376-382.

National Health Service. (2009). 30 a day: Everyday exercise [Electronic Version]. Retrieved 1.12.2009 from http://www.nhs.uk/Livewell/Derby/Pages/30adayexercise.aspx.

Noar, S. M., Benac, C., & Harris, M. (2007). Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychological Bulletin*, *133*(4), 673-693.

Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology, 51*, 390-395.

Richert, J., Lippke, S., & Ziegelmann, J. P. (in press). Intervention-engagement and its role in the effectiveness of stage-matched interventions promoting physical exercise. *Research in Sports Medicine: An International Journal.*

Robert Koch Institut, & Statistisches Bundesamt. (2005). *Körperliche Aktivität* [Physical Activity] [Electronic Version]. Retrieved 1.12.2009 from http://edoc.rki.de/documents/rki_fv/ren4T3cctjHcA/PDF/29BFVzVHbIM_45.pdf.

Schwarzer, R. (1992). Self-efficacy in the adoption and maintenance of health behaviours: Theoretical approaches and a new model. In R. Schwarzer (Ed.), *Self-efficacy: Thought control of action* (pp. 217-243). Washington, DC: Hemisphere.

Schwarzer, R. (2008). Modeling health behavior change: How to predict and modify the adoption and maintenance of health behaviors. *Applied Psychology: An International Review, 57*(1), 1-29.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Allyn and Bacon.

Warner, L. M., & Lippke, S. (2008). Psychological stage models of physical exercise – Research Advances. In M. P. Simmons & L. A. Foster (Ed.), *Sport and exercise psychology - research advances* (pp. 19-51). Hauppauge, NY: Nova Science Publishers.

Weinstein, N. D., & Sandman P. M. (1992). A model of the precaution adoption process: evidence from home radon testing. *Health Psychology*, *11*, 170-180.

Weinstein, N. D., Lyon, J. E., Sandman, P. M., & Cuite, C. L. (1998). Experimental evidence for stages of health behavior change: The precaution adoption process model applied to home radon testing. *Health Psychology*, *17*(5), 445-453.

Weinstein, N. D., Rothman A., & Sutton S. R. (1998). Stage theories of health behavior. *Health Psychology*, *17*, 1-10.

World Health Organization. (2009). *Recommended amount of physical activity* [Electronic Version]. Retrieved 1.12.2009 from http://www.who.int/dietphysicalactivity/factsheet_recommendations/en/index.html.

3

Stages of Health Behavior Change and Mindsets:

A Latent Class Approach

Jana Richert[1], Natalie Schüz[1], Sonia Lippke[1,2],
& Benjamin Schüz[3]

[1]Freie Universität Berlin, [2]Maastricht University,
[3]German Centre of Gerontology

Richert, J., Schüz, N., Lippke, S., & Schüz, B. (submitted). Stages of Health Behavior Change and Mindsets: A Latent Class Appraoch, Manuscript submitted for publication.

Abstract

Objective. Stage theories of health behavior are popular and of high practical relevance. Tests of the validity of these theories provide limited evidence because of validity and reliability problems. This study provides a bottom-up approach to identify behavioral stages from examining differences in underlying mindsets. We examine the concurrent validity of a latent-class based approach and a commonly used stage-algorithm based on self-reports about intentions and behavior in order to identify possible strengths and shortcomings.

Methods. Social-cognitive variables and individuals' stages were assessed in a sample of 2219 internet users. Latent class analysis (LCA) was used to identify distinct groups with similar patterns of social-cognitive predictors. Convergent validity of the LCA solution and stage algorithms was tested by examining adjusted standardized residuals.

Results. The LCA identified four distinct profiles – not intending to change, intending to change (no action), intending to change with action, and maintaining. Convergent validity with a stage algorithm was low, in particular in the non-intending and maintaining stages.

Conclusion. Stages as assigned by the stage-algorithm did not correspond well with the extracted mindsets: This indicates that commonly used stage-algorithms might not be effective in assigning individuals to stages that represent mindsets, undermining the possibility for stage-matched interventions.

Keywords: Latent class analysis, Validity, Stage-algorithm, Stage-theories, Self report measure

Stages of Health Behavior Change and Mindsets: A Latent Class Approach

In recent years, particularly in applied fields, stage theories of health behavior change have become increasingly popular. The idea that people pass through an ordered set of qualitatively different stages on their way to a new health behavior is intuitively appealing and a *motif* for the description of many change processes (Brug et al., 2005). It also is highly attractive for practical application, as it implies targeting specific intervention components for individuals in different stages, and suggests that such interventions are more effective than one-size-fits-all measures (Prochaska et al., 2004).

However, a crucial question beyond this attractiveness is the question about the construct validity of stages of health behavior change. In this article, we propose an alternative to current procedures examining the validity of stages, which heavily rely on the validity and reliability of the algorithms used for the measurement of stages. We argue that subgroups of individuals with a particular mindset towards health behaviors (as inherently assumed by the stage construct; Weinstein, Rothman, & Sutton, 1998) can be more reliably inferred from the data using a latent class analysis approach. In a second step, we examine the convergent validity of a commonly used stage measure and these mindsets.

Discontinuity and Mindsets

Current tests of the construct validity of stage theories rely on the identification of discontinuity in the means or effects of relevant factors across the stages of the theory (Weinstein, Rothman, & Sutton, 1998). The rationale underlying this quest for discontinuity is the assumption that during the process of change, individuals can have different mindsets towards behavior, and that these different mindsets manifest themselves in different cognitions. Mindset theory (Heckhausen, 1991; Heckhausen & Gollwitzer, 1987) assumes that the mindset of an individual towards behavior changes as a result of cognitive and behavioral processes. One of the most basic differences between mindsets, which might serve as an example here, is the difference between a *deliberative* and an

implemental mindset. Individuals in a deliberative mindset weigh the advantages and disadvantages of a specific behavior, finally resulting in a decision for or against the behavior. Individuals in an implemental mindset focus on executing this decision. Experimental studies support this assumption: For example, individuals in a deliberative mindset had a better memory for words describing both positive and negative aspects of a behavior, while individuals in an implemental mindset remembered positive attributes better, indicating that they no longer searched for a balanced decision (Fujita, Gollwitzer, & Oettingen, 2007). Another study documented that individuals in an implemental mindset were more likely to generate success-oriented scenarios resulting from behavior than individuals in a deliberative mindset (Taylor & Gollwitzer, 1995). These studies suggest that deliberative and implemental mindsets are characterized by fundamentally different cognitions and cognitive processes.

The Stage Construct in Stage Theories of Health Behavior Change

Stage theories of health behavior adopt this idea of different mindsets by construing the process of health behavior change as progressing through different stages with differential mindsets. As the stages/ mindsets are defined by different cognitions and processes, they should be affected differentially by specific treatment content. Most stage theories however assume a more fine-graded stage distinction than just a deliberative and implemental mindset. For example, the Transtheoretical Model (TTM) (Prochaska, DiClemente, & Norcross, 1992) assumes that individuals pass through five distinct stages from precontemplation to maintenance. The Precaution Adoption Process Model (Weinstein, 1988) assumes no less than six stages from unaware of the issue to maintenance, and the Health Action Process Approach (Schwarzer, 1992) assumes two meta-stages and a number of finer-graded stages. All approaches, however, share transitions from non-intending to change behavior to intending to change behavior, from intending to change behavior to actually changing behavior, and from changing behavior to maintaining or habituation (Schüz, Sniehotta, Mallach, Wiedemann, & Schwarzer, 2009;

Schwarzer, 2008). This idea of qualitative differences between the stages is the logic underlying the tests for the validity of the stages.

Validity of Stage Theories – the Quest for Discontinuity

Most tests of the validity of stage theories rely on the idea that individuals in the same stage have a mindset more similar within this stage than to individuals in another stage. As a consequence of this idea, various tests of the validity of stage theories can be formulated (Sutton, 2000; Weinstein, Rothman, & Sutton, 1998). These tests share the idea that the effects of a particular factor on the likelihood of subsequent stage transitions follow a discontinuous (non-linear) pattern across the stages. For example, for cross-sectional data, Sutton (2000) requires that the means of stage-specific factors should follow a pattern that does not fit a linear trajectory across the stages but rather a quadratic, cubic or any other non-linear trend. A number of studies have examined such discontinuity patterns of means across stages and interpret these to support the validity of the underlying stage construct (Armitage, Povey, & Arden, 2003; Sniehotta, Luszczynska, Scholz, & Lippke, 2005). As Weinstein and colleagues (1998) point out, such tests constitute the lowest level of evidence for a stage theory, since alternative explanations for the discontinuity patterns are possible, such as non-linear increases across the stages or reverse causality. Stronger evidence, according to Weinstein et al. (1998), is constituted by discontinuous predictors of stage transitions in longitudinal settings. Accordingly, a factor predicting transitions from an earlier stage to a later stage should be specific solely for this transition if the underlying stage construct were true. This idea has been examined in a range of studies with moderate evidence strength for various predictors from risk perceptions over specific self-efficacy beliefs to social support in a range of health behaviors (Armitage, Sheeran, Conner, & Arden, 2004; Schüz, Sniehotta, Mallach, Wiedemann, & Schwarzer, 2009; Wiedemann et al., 2009). The strongest evidence for the validity of stage theories comes from experimental matched-mismatched intervention studies, in which the idea of discontinuity is evident in the test for differential effects of the intervention according to the stage a person is in (Richert, Lippke, & Ziegelmann, in press;

Schüz, Sniehotta, & Schwarzer, 2007; Weinstein, Lyon, Sandman, & Cuite, 1998; Weinstein, Rothman, & Sutton, 1998).

Validity of the Validity Tests Revisited

A crucial issue in this domain is the way stages are operationalized in studies aiming to test the construct validity of the stages concept. Most often, stage assessments are based on algorithms consisting of the answers to a number of questions with regard to the studied behavior (Godin, Lambert, Owen, Nolin, & Prud'homme, 2004), which can be more or less successful compared to other assessments of intentions or behavior (Lippke, Ziegelmann, Schwarzer, & Velicer, 2009). However, while such approaches are useful in examining whether individuals assigned to specific stages differ in the effects or means of variables deemed important, they rely on the limited reliability and sometimes limited validity of the underlying stage algorithm and stage theory. There are both statistical and theoretical problems with such tests: Algorithms based on single items or combinations of single items can face a serious problem resulting from limited reliability as measurement errors cannot be corrected for in such assessment. In addition, ANOVA-based tests for discontinuity such as fitting linear or quadratic trends to mean differences or planned contrasts rely on the statistical assumption that there is equidistance or at least a monotonous increase or decrease between the stages. This statistical requirement however is not inherent in the stage definition of stage theories (see for example the arbitrary sequential order of the *decided to act* / *decided not to act* stages in the PAPM (Weinstein, 1988)). Coming from a theoretical viewpoint, an examination of the idea that individuals differ in mindsets, i.e., differ in cognitions and cognitive processes, would require that differences in these cognitions are used to assign individuals to mindsets. This test is what this article aims at – provide a bottom-up assignment of individuals to similar groups, based on the similarity of their cognitions, and examine whether these groups match the predictions and allocations offered by current stage allocation procedures. Furthermore, as stages of health behavior change are a social and scientific construct rather than an empirical entity (Schwarzer, 2008), a nomothetic approach prescribing a stage

differentiation and sequence might be an oversimplification of the complex nature of human behavior change processes. Applying confirmatory approaches by examining discontinuity between arbitrary or at least a-priori-defined stages might therefore not be appropriate to examine mindset differences—an exploratory approach is better suited to examine the assumption of qualitative differences between stages or mindsets. In this article, we propose an alternative to such nomothetic top-down approaches by applying a bottom-up based approach, i.e., inferring differential mindsets from differences in cognitions and cognitive processes.

Inferring mindsets from social cognitions – a bottom-up-approach

As outlined above, the idea of qualitatively different stages of health behavior is based on the assumption of differential mindsets in stages (Heckhausen, 1991; Weinstein, Rothman, & Sutton, 1998). Our approach takes these differential mindsets as starting point. Unfortunately, most stage theories are not very precise with regard to the factors assumed differentially important in the stages. The TTM (Prochaska, DiClemente, & Norcross, 1992) proposes ten different processes of change, but so far, tests have provided no evidence for the assumed stage-specific effects (Herzog, 2008). The PAPM makes differential assumptions for the effects of risk perception, which should be more important for stage transitions in early stages such as unaware of the issue, and self-efficacy, which should be more important in later stages, such as decided to act. The HAPA makes differential assumptions for the transitions from not intending to intending to change, and from intending to change to changing behavior: For transitions from the first stage, *risk perceptions*, *outcome expectations*, and *motivational self-efficacy* are assumed important, whereas for the transition from intending to acting, *coping self-efficacy*, *planning* and cognitive *action control* are assumed important, while for the transition from acting to maintaining in particular *recovery self-efficacy* is deemed effective (Schwarzer & Luszczynska, 2008). These factors are also inherent in most theories delineating the determinants of intention formation (Schüz, Sniehotta, Mallach, Wiedemann, & Schwarzer, 2009), and the predictors of behavior change are similarly shared between

various theories (Sniehotta, 2009). Assuming that these predictors comprise the most relevant factors for explaining behavior change, it should be possible to characterize individuals in different mindsets by a different combination of values on these factors—such as that individuals in an initial mindset before committing to a behavioral intention should have rather low perceptions of risk and low levels of proximal behavior predictors such as planning or action control, while individuals in a mindset aimed at pursuing a behavioral intention should score higher on such proximal factors and lower on risk perception or negative outcome expectations. Our approach therefore aims at identifying subgroups of individuals with similar patterns of the social cognitions inherent in most theories of behavior change. We aim at testing the convergent validity of this approach with a standard stage-algorithm by examining whether these subgroups match those predicted by an algorithm based on self-reports of intentions and behavior (Lippke, Ziegelmann, Schwarzer, & Velicer, 2009; Richert, Lippke, & Schwarzer, 2010).

Matching extracted profiles to stages predicted by a stage-algorithm

A valid and reliable stage measure (or algorithm) should assign individuals to stages that are an accurate reflection of individuals' mindsets (or in other words their social-cognitive patterns). Subsequently, if a stage measure is indeed valid, the classes that represent the social cognitive patterns found in the data should correspond well with the stages as assigned by a stage algorithm.

Research Questions

In this study, we aim at examining whether mindsets inferred from patterns of social cognitions in homogeneous subgroups of individuals match the predictions of stage allocation of a current stage algorithm based on intentions and behavior.

This approach might overcome problems of current tests of discontinuity between behavioral stages or mindsets due to assumptions of equidistance between stages, sequence of stages and limited validity of stage assessments. These homogeneous subgroups of individuals will be identified using latent class analysis, a statistical technique assessing a

categorical latent variable (e.g., the latent stage) in a data set constituting groups of individuals with maximally homogeneous patterns of predictors. Using this approach, it might be possible to overcome problems of limited reliability inherent in current stage algorithms, because it accounts for measurement error in its latent variable framework. It may also overcome problems of validity limitations as it goes beyond a nomothetic top-down approach by identifying differential mindsets from a bottom-up perspective.

Method

Procedure

The protocol for this study has received approval by the Internal Ethics Review Board of the Freie Universität Berlin[2]. Individuals were recruited for a web-based intervention study on fruit and vegetable consumption by press releases (radio, newspaper, TV) and advertisements posted on the university website. Participants visited a starting web page, and, after giving informed consent, were directed to a baseline questionnaire. As incentive, participants could take part in a raffle for online shop gift certificates. After the baseline questionnaire, which the current study is based on, participants were randomly allocated to one of four experimental groups for an intervention study (not reported here).

Participants

The study sample consists of N=2220 individuals, who were on average M=38.22 years old (range=13-79, SD=12.64) and mostly women (80.8%). The majority of the sample was highly educated (43.6% College degree), employed (63.5%) and in a steady relationship (59.3%).

Measures

Unless otherwise noted, all items were rated on a 6-point Likert scale ranging from (1) not at all true to (6) exactly true. Scale means were computed, and scale values were

[2] Approval Number: Gespsy_2009-03-13

dichotomized at the theoretical mean of 3.5 in order to facilitate interpretation of the latent classes.

Risk Perception was measured with three items adapted from Schüz, Sniehotta, Wiedemann, Mallach, & Schwarzer (2009): "If I continue to live this way, there is a high probability of me… (1) having a heart attack or stroke, (2) having diabetes, and (3) being obese." Cronbach's Alpha was .86; $M = 3.35$, $SD = 1.33$.

Positive Outcome Expectancies were assessed with four items adapted from Schüz, Sniehotta, Wiedemann, Mallach, & Schwarzer (2009): "If I eat sufficient amounts of fruits and vegetables every day, then… (1) I feel good and content, (2) I am doing something for my health, (3) I have good mental functioning, and (4) it has positive effects on my physical appearance." Cronbach's Alpha was .86; $M = 4.85$, $SD = .70$.

Negative Outcome Expectancies were assessed using the same item stem followed by three statements: "(1) my food does not taste as good, (2) it will be a financial burden, and (3) then I will have to invest a lot of time and effort (e.g., grocery shopping, food preparation)." (cf., Schüz, Sniehotta, Wiedemann, Mallach, & Schwarzer, 2009). Cronbach's Alpha was low with $\alpha = .54$, which indicates that the scale assesses diverse outcome expectancies; $M = 2.7$, $SD = .98$.

Motivational Self-efficacy was measured with the two items: "I am confident that I can/ could eat sufficient amounts of fruits and vegetables…(1) even if it is difficult for me, and (2) even if there are few convenient shopping possibilities." (Schüz, Sniehotta, Wiedemann, Mallach, & Schwarzer, 2009). Items correlated significantly with $r = .55$, $p < .01$; $M = 4.61$, $SD = .95$.

Action Planning was assessed with three items based on Sniehotta, Schwarzer, Scholz, & Schüz (2005): "I have planned precisely… (1) which fruits and vegetables I will eat, (2) at which occasions (in which situations) I will eat fruits and vegetables, and (3) how I will eat my fruits and vegetables (e.g., cooked, cut up)." Cronbach's Alpha was .88; $M = 3.32$, $SD = 1.20$.

Coping Planning was assessed with two items: "I have planned precisely… (1) in which situations I need to be especially careful so as to succeed in eating sufficient

amounts of fruit and vegetables and (2) what I can do in difficult situations so as to succeed in eating sufficient amounts of fruits and vegetables." (Wiedemann, Lippke, Reuter, Schüz, Ziegelmann, & Schwarzer, 2009). Items correlated significantly with r=.68, p<.01; M = 2.81, SD = 1.20.

Coping Self-efficacy was measured with two items: "I am confident that I can keep eating sufficient amounts of fruits and vegetables…(1) even if I have to overcome obstacles (e.g., that there is no fruit or vegetable available at the grocery store I usually go to), and (2) even if have problems or worries." (Wiedemann, Lippke, Reuter, Schüz, Ziegelmann, & Schwarzer, 2009). Items correlated significantly with r = .62, p < .01; M = 4.15, SD = 1.04.

Recovery Self-efficacy was measured with two items: "I am confident that I can eat sufficient amounts of fruits and vegetables again…(1) even if I have failed to do so for a few days, and (2) even if I haven't done so for quite some time." (Wiedemann, Lippke, Reuter, Schüz, Ziegelmann, & Schwarzer, 2009). Items correlated significantly with r = .78, p < .01; M = 4.86, SD = .84.

Action Control was assessed with three items based on Sniehotta, Scholz, & Schwarzer (2005): "I am aware of how many portions of fruits and vegetables I want to eat daily.", "I check whether or not I have eaten as many fruits and vegetables as I had intended.", and "I am trying hard to eat as many fruits and vegetables as I had intended." Homogeneity of the items was high with Cronbach's Alpha being .87; M = 2.79, SD = 1.16.

Stage assessment. Adapted from previous studies (Lippke, Ziegelmann, Schwarzer, & Velicer, 2009; Richert, Lippke, & Schwarzer, 2010), participants responded to the following two items: (1) "In the past week, have you eaten *enough* fruits and vegetables *per day*?" as well as (2) "In the near future, do you intend to eat more fruits and vegetables than you are eating now?". Answers were given in a closed yes/ no format. Participants were coded as *Non-Intenders* if they answered 'no' to both questions, as *Intenders* if they answered 'no' to the first and 'yes' to the second question, as *Maintaining Actors* if they responded 'yes' to the first and 'no' to the second question, and as *Changing Actors* if they

answered ‘yes’ to both questions (stage labels are in accordance with Richert, Lippke, & Schwarzer, 2010).

Statistical Analyses

Latent class analysis (LCA) was used to identify mutually distinct subpopulations of individuals sharing profiles or response patterns on the social-cognitive predictors of behavior outlined above, thus identifying groups of individuals in qualitatively different mindsets. Latent class indicators consist of distinct categories, and the latent class variable is assumed to be measurement-error-free (Magidson & Vermunt, 2004). The result of LCA is a reduced number of latent profiles that can explain all existing response patterns in the data. For each participant, probabilities of class membership are available. Based on this information, individuals are assigned to the one latent class for which their assignment probability is highest. Class membership is mutually exclusive, so that each individual is assigned to one class only. The model fit is evaluated by the Akaike information criterion (AIC) and Bayesian information criterion (BIC). Class solutions for different numbers of classes can be tested against each other based on the Lo-Mendell-adjusted bootstrapped likelihood ratio (LR) test, which compares an estimated model to a model of one less class than the estimated model (Lo, Mendell, & Rubin, 2001).

The research question of whether the LCA classes are convergent valid with the stage algorithm is tested by examining adjusted standardized residuals, which are deviations of observed frequencies from frequencies that are expected based on chance. Additionally, these deviation values are adjusted to the cell size and are approximately normally distributed. Thus, if values exceed the critical values +/- 1.96, 2.58, or 3.29 respectively, the number of cases found in that cell is significantly higher or lower than would be expected based on mere chance (Agresti, 2002). If the stage measure and the LCA classes are convergent valid, the frequencies in correspondence cells (i.e., cells in which classes are matched against their stage counterpart) should be significantly higher than what is expected based on mere chance. Likewise, frequencies in non-correspondence cells should be lower than or equal to frequencies expected based on chance.

Results

Results from Latent Class Analysis

The AIC indicates that a solution with four (AIC = 18362.47) or five classes (AIC = 18361.31) fits the data better than the three (AIC = 18383.15) or the six-classes (AIC = 22331.13) solutions. The BIC, however, is lowest for a 3 class solution (BIC = 18548.59 as compared to 18757.55 for 2 classes or 18584.96 for four classes). The LR test for the four-class model yielded a significant result (LR test value = -9162.58, $p < .001$), indicating that the four-class solution fits the data better than a three class model. For the five-class model, the LR test was non-significant (LR = -9142.234, $p = .37$), which indicated that the four-class model fits better than the five-class model.

We therefore decided to extract four latent classes, and figure 1 shows the profiles in terms of conditional solution probabilities. A solution probability close to 1 indicates a high likelihood of scoring high on the respective scale. Individuals were assigned to the class for which they had the highest probability scores. In this study, these probabilities were exceptionally high: class 1 = .91, class 2 = 1.0, class 3 = .97, class 4 = .93, indicating a high reliability of class assignment.

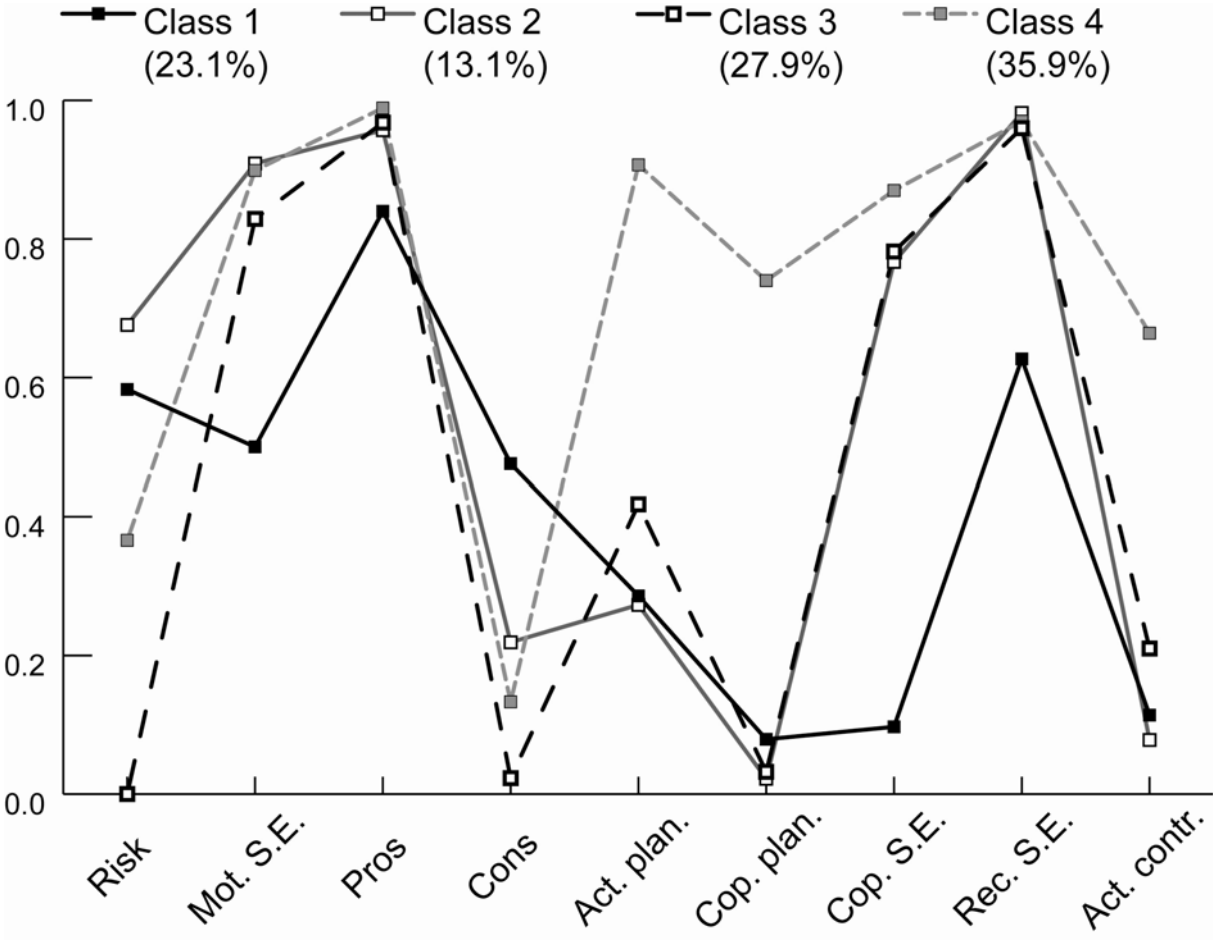

Figure 1. Latent class profiles for the four-class model.
Note. Risk = risk perception, Mot. S.E. = motivational self-efficacy, Pros = positive outcome expectancies, Cons = negative outcome expectancies, Act. Plan. = Action Planning, Cop. plan. = Coping Planning, Cop. S.E. = Coping Self-efficacy, Rec. S.E. = Recovery self-efficacy, Act. contr. = Action Control.

Members of class 1 (23.1% of all participants) had a moderate likelihood (.58) to score high on risk perception, the lowest likelihood to have high levels of motivational self-efficacy, and they were the least likely to perceive advantages of fruit and vegetable intake, while at the same time being the most likely to perceive disadvantages of the behavior compared to all other classes. Furthermore, their likelihood to score high on volitional scales (i.e., Action Planning, Coping Planning, Coping- and Recovery Self-efficacy, and Action Control) was relatively low. This pattern resembles individuals in a deliberative mindset, i.e., *not intending to change*.

Individuals in class 2 (13.1%) were most likely to score high on risk perception (.68). Members of this class were also very likely to have high levels of motivational self-efficacy as well as to perceive advantages of fruit and vegetable intake. They had low

likelihoods to perceive disadvantages of behavior. Their likelihood to score high on Action Planning, Coping Planning and Action Control was rather low, while their likelihood of scoring high on Coping- and Recovery Self-efficacy was high. This pattern resembles individuals who are not deliberating anymore, that is, individuals who are *intending to change*.

Class 3 (27.9% of participants) had a zero likelihood of scoring high on risk perception. Individuals in this class were likely to have high levels of motivational self-efficacy as well as to perceive benefits of fruit and vegetable consumption. At the same time, they were very unlikely to perceive disadvantages of the behavior. Their likelihood of scoring high on Action Planning, Coping Planning and Action Control was low. Their likelihood to score high on Coping- and Recovery Self-efficacy was rather high. This pattern most resembles individuals *maintaining* behavior.

Members of class 4 (27.9%) had a moderate likelihood (.37) to score high on risk perception and to have high levels of motivational self-efficacy. They were very likely to perceive advantages of fruit and vegetable intake and very unlikely to perceive disadvantages of the behavior. Furthermore, their likelihood to score high on volitional scales (i.e., Action Planning, Coping Planning, Coping- and Recovery Self-efficacy, and Action Control) was high. This pattern resembles individuals with an *intention to change* something about their behavior.

Test of Class–Stage Correspondence

The stage algorithm appointed 43 individuals (1.9% of the total sample) to the Non-Intenders stage, the majority of participants (n = 1591, 71.7%) were categorized as Intenders, 235 individuals (10.6%) were labeled as Maintaining Actors and 350 people (15.8%) were classified as Changing Actors.

Table 1. Frequencies for LCA classes and for stages as obtained by a stage measure.

			Latent Classes extracted by the LCA				
			Class 1	Class 2	Class 3	Class 4	total
Stages assigned by the stage measure	NI	observed *N*	**7**	15	6	15	43
		expected *N*	**10**	6	12	15	
		stand. adj. residuals	**-1.1**	4.3***	-2.1*	-0.1	
		% within LCA classes	**1.4%**	5.2%	1.0%	1.9%	
		% within stages	**16.3%**	34.9%	14.0%	34.9%	
		% of total *N*	**0.3%**	16.9%	0.3%	0.7%	1.9%
	I	observed *N*	329	**242**	353	667	1591
		expected *N*	368	**208**	444	571	
		stand. adj. residuals	-4.3***	**4.8*** **	-9.5***	9.4***	
		% within LCA classes	64.1%	**83.4%**	57.0%	83.7%	
		% within stages	20.7%	**15.2%**	22.2%	41.9%	
		% of total *N*	14.8%	**10.9%**	15.9%	30.1%	71.7%
	MA	observed *N*	79	10	**111**	35	235
		expected *N*	54	31	**66**	84	
		stand. adj. residuals	4.0***	-4.2***	**7.0*** **	-7.1***	
		% within LCA classes	15.4%	3.4%	**17.9%**	4.4%	
		% within stages	33.6%	4.3%	**47.2%**	14.9%	
		% of total *N*	3.6%	0.5%	**5.0%**	1.6%	10.6%
	CA	observed *N*	98	23	149	**80**	350
		expected *N*	81	46	98	**125.7**	
		stand. adj. residuals	2.4*	-3.9***	6.7***	**-5.5*** **	
		% within LCA classes	19.1%	7.9%	24.1%	**10.0%**	
		% within stages	28.0%	6.6%	42.6%	**22.9%**	
		% of total *N*	4.4%	1.0%	6.7%	**3.6%**	15.8%
	total	observed *N*	513	290	619	797	2219
		% of total *N*	23.1%	13.1%	27.9%	35.9%	

Note. NI, Non-Intender; I, Intender; MA, Maintaining Actor; CA, Changing Actor; stand. adj. residuals, standardized adjusted residuals; expected *N*s are rounded; * $p<.05$, ** $p<.01$, *** $p<.001$

Of the individuals placed in class 1, only 1.4% of individuals were classified as Non-Intenders based on the stage-algorithm. This number was not higher than what would be expected based on mere chance (standardized adjusted residual = -1.1, $p > .05$), indicating that there was no correspondence between class 1 and the Non-Intender stage. The majority of individuals assigned to class 2 (83.4%) were classified as Intenders based on the stage measure. The frequency observed in this cell ($n = 242$) was significantly higher than the frequency expected based on chance ($n = 208$, standardized adjusted residual = 4.8, $p < .001$). This indicates a high correspondence between class 2 and the Intender stage. Only 17.9% of the individuals, who were placed in class 3 were classified

as Maintaining Actors. The standardized adjusted residual of 7.0 ($p < .001$) shows that the observed frequency was significantly higher than what would have been expected based on chance. This indicates a good match. However, the frequency found in the cell pinning class 3 against the Changing Actor stage was also significantly higher than what would be expected based on chance (standardized adjusted residual = 6.7, $p < .001$). Here, 24% of the individuals, who were placed in class 3 were classified as Changing Actors, indicating an equally good match between these categories. The majority of individuals assigned to Class 4 (83.7%) were classified as Intenders based on the stage measure. The frequency observed in this cell ($n = 667$) was significantly higher than the frequency expected based on chance ($n = 571$, standardized adjusted residual = 9.4, $p < .001$). This indicates a high correspondence between class 4 and the Intender stage. Only 10% of individuals assigned to class 4 were classified as Changing Actors (correspondence cell). The standardized adjusted residual of -5.5 ($p > .05$) revealed that the observed frequency was significantly lower than what would be expected based on chance, indicating that there was no correspondence between these categories.

Discussion

This study aimed at providing a bottom-up approach to test stages of behavior change. Furthermore, we examined whether the assumption of stage theories, namely qualitatively different mindsets of individuals in different stages of behavior change (Weinstein, Rothman, & Sutton, 1998), can be held if allocation of individuals to mindsets based on patterns of social cognitive variables using Latent class analysis are compared to the allocation of individuals to stages based on a stage algorithm.

Mindsets and behavioral stages

Our study followed a bottom-up-approach, that is, we did not rely on somewhat arbitrary temporal (Sutton, 2001) or sequential criteria to assign individuals to stages, but followed the basic idea of stage theories that individuals in qualitatively different stages of

behavior change should differ in what their cognitions about specific behaviors (Weinstein, Rothman, & Sutton, 1998). This idea implies different mindsets of individuals in different stages. This should be evident in greater similarity of cognitions between individuals *within* one stage than *between* different stages. Evidence from research on mindsets support this notion (Gollwitzer, Heckhausen, & Steller, 1990). Our analysis design accounts for this demand as latent class analysis infers latent classes and membership to these classes from similarities within and dissimilarities between classes (Magidson & Vermunt, 2004). In contrast to other group-identifying techniques such as cluster analysis, LCA, by way of the Chi²-difference-test, allows for statistically testing the number of latent classes.

This approach also allows for overcoming potential problems of confirmatory approaches that rely on limited reliability and validity of stage algorithms, as it is a latent variable procedure allowing for measurement-error-free assessment of latent classes. The bottom-up nature of our approach, that is, inferring latent classes representing different mindsets or stages from cognitions towards behavior, is also closer to the logic inherent in stage theories. There are qualitative differences with regard to cognitions between individuals in different stages (Weinstein, Rothman, & Sutton, 1998). In contrast to tests for discontinuity of means or effects of specific variables between behavioral stages in order to test the validity of the stages construct (Sutton, 2000), our approach is not dependent on assumptions of equidistance or monotonous linear relations between stages. Although an ANOVA itself of course does not require equidistance or ordinal characteristics of the levels of the factor, the interpretation of statistical tests for trends between the levels of the independent factor does. A reordering of the levels of a factor might turn a linear trend into a quadratic one and vice versa. This ordinal or even equidistant requirement for tests however is not inherent in the concept of behavioral stages. Our approach does not rely on such assumptions, but instead infers behavioral stages as qualitatively different mindsets from the data and as such might help future research on the validity of behavioral stages to overcome these limitations.

Additionally, as latent class analysis displays probabilities rather than certainties to score highly on a respective scale, it accounts for the possibility of individuals belonging in a particular class to score differently on that scale. This might be beneficial when dealing with interventions, because behavior change interventions are developed based on group means, i.e., they target variables that have been identified as predictors of behavior change in the majority of individuals in a particular class.

Behavioral Stages Identified by Latent Class Analysis

Our analysis identified four latent classes that can be matched unto the stages defined in most stage theories (Schüz, Sniehotta, Mallach, Wiedemann, & Schwarzer, 2009): a stage *before* individuals have formed an intention, a stage *with* intentions but *without* behavior, a stage with maintained behavior and one with intended changes in current behavior (Lippke, Ziegelmann, Schwarzer, & Velicer, 2009).

Individuals classified into the first latent class match individuals in a stage *before* an intention for behavior change in that they have lower levels of risk perception, motivational self-efficacy, expected positive outcomes of fruit and vegetable consumption, plans, coping or recovery self-efficacy and action control, but higher levels of negative outcome expectancies than individuals in other latent classes. This matches a *deliberative* mindset, in which individuals are more open to positive and negative information about behavior (Gollwitzer, Heckhausen, & Steller, 1990). The low levels of all post-intentional factors suggest that these individuals have engaged in little reasoning about behavior change.

Individuals classified in the second latent class match individuals in a stage *after* intention formation, but before actual behavior change in that they have the highest level of risk perception, an indicator of their perceived relevance of behavior change (Weinstein, 2003). In contrast to individuals in the first latent class, they perceive high levels of positive outcome expectancies and low levels of negative outcome expectancies, which indicates an *implemental* mindset (Gollwitzer, Heckhausen, & Steller, 1990), and could also serve the purpose to reduce discrepancy once a behavioral decision has been made.

With regard to volitional factors, individuals in this latent class have low levels of action planning and coping planning, which may be explained by the fact that they have not initiated behavior change so far, but relatively high levels of coping and recovery self-efficacy. While especially this latter result might seem to contradict predictions made e.g., by the HAPA (Scholz, Sniehotta, & Schwarzer, 2005), tenets of self-efficacy theory can help in interpreting this result: An optimistic belief in one's abilities to overcome setbacks and to recover from behavioral lapses can also be an important precondition of reasoning about behavior change, and only high levels of self-efficacy in these domains will help to commit to the goal of adapting a new behavior such as increased fruit and vegetable consumption.

Individuals classified in the third latent class show a profile of cognitions that matches individuals *maintaining* behavior. They score lowest on risk perception, reflecting the fact that their risk for diseases due to absence of nutritional health risk behavior is very low (Renner, Schüz, & Sniehotta, 2008; Weinstein, Rothman, & Nicolich, 1998). They also report low levels of cognitive action control, which might reflect the fact that fruit and vegetable consumption is habitual for them and requires little to no conscious effort (Wood, Quinn, & Kashy, 2002). In contrast to individuals assigned to class 4, individuals in this stage have a very low chance of scoring high on coping planning; possibly reflecting that due to habituation of behavior, no cognitive efforts for overcoming critical situations is needed.

Individuals allotted to latent class 4 match individuals in a stage with *some* behavior but *intended changes*. These individuals have especially high levels of action- and coping planning, and action control – cognitive indicators of ongoing behavior change processes and effective strategies to initiate and maintain behavior change (Sniehotta, Scholz, & Schwarzer, 2005). This finding might be due to their increased efforts to change behavior and also because of better recall of such strategies due to the relative recency of their behavior change.

Correspondence between mindsets and stages as measured by a stage-algorithm

As the LCA extracts social-cognitive profiles from the data (rather than confirming *a priori* set classes), it might be justified to infer that these profiles are an accurate reflection of qualitatively distinct mindsets. The reliability of the LCA solution (cf. probabilities in the results section) suggests that if stages and classes do not correspond well, the stage-algorithm might not be valid and reliable in assessing stages that are reflective of mindsets. The consequence of such misclassification would ultimately be the receipt of a mismatched and subsequently ineffective intervention. To investigate this question of correspondence, we matched LCA mindsets to the predictions of stage allocation made by a current stage measure based on intentions to change and behavior. To understand the results, it is necessary to consider the frequencies that were observed as well as their relationship to the frequencies that were expected based on chance. For example, by looking at observed frequencies alone, it looks as though the match between class 3 and the Intender stage was better than the match between class 3 and any other stage (the majority of all individuals placed in class 3 were classified as Intenders). However, the frequency was significantly lower than what would have been expected based on chance, suggesting that the visible match is not tenable.

Although the profile of class 2 clearly resembles individuals with a post-deliberative mindset (cf. Figure 1 as well as the results section), it corresponds equally well with the Non-intender Stage and the Intender Stage. This suggests that the algorithm used in this study is not successful in assigning individuals to stages that represent a definite mindset. From a practical view, this implies that if an individual has a post-deliberative mindset but is assigned to the Non-Intender Stage when using the stage algorithm, an intervention matched to a deliberative mindset (as is appropriate for a Non-Intender) will be unsuited. The same holds true for class 3. The correspondence is equally high between this class and the Maintaining Actor and Changing Actor Stage. This ambiguity might result in incorrect classifications of individuals with the mindset of a maintaining actor as changing actors. Similar results were found for classes 1 and 4. Class 1 best corresponded with the Maintaining Actor stage and class 4 best matched the Intender Stage. This data

suggests that this commonly used stage-algorithm based on self-reports about intentions and behavior contains a high risk of misclassification. It thereby offers an explanation for the sometimes found lack of supportive results in the field of stage-matched interventions (Adams & White, 2005; Bridle et al., 2005; Conn, Hafdahl, Brown & Brown, 2008; Noar, Benac & Harris, 2007).

Limitations

A potential limitation relates to the fact that we have relied on self-reports of social cognitions. Research on mindsets has shown that differences between mindsets are also evident on the level of cognitive performance (Fujita, Gollwitzer, & Oettingen, 2007) and implicit cognitions (Custers & Aarts, 2007). Future research should consider this to distinguish mindsets. Also, the latent classes we extracted are based on social-cognitions, excluding intentions to change and behavior. Although future research may consider including both variables, we decided against it as we matched the social-cognitive profiles against stages as predicted by a measure relying on self-reports about intentions to change and behavior. We thereby avoided circular reasoning. Additionally, our sample was self-selected. This limits the generalizability of our results. Finally, although our data allows assessing homogeneous subsets of individuals with similar cognitions (i.e., mindsets), the cross-sectional nature of our data set allows no conclusions with regard to predictors of behavior change. Future studies might want to use Latent Transition Analysis to analyze transitions between latent class patterns indicative of changes in intention and behavior (cf., Guo, Aveyard, Fielding, & Sutton, 2009; Velicer, Martin, & Collins, 1996).

Outlook

Our study relied on the identification of latent classes from profiles of social-cognitive variables. These variables do not necessarily drive stage transitions. For example, both Non-intenders and Maintaining actors have low risk perceptions. It does not follow however, that both groups need an intervention addressing risk perception as Maintaining Actors might e.g., have low levels of risk perception, because they take their preventive

health behavior into account when estimating their risks. Future Research should further investigate if the process of change is different for different classes characterized by different social-cognitive profiles. Only experimental research, where interventions that are tailored to classes are pitted against interventions that are mismatched, can really answer this question (Sutton, 2000).

Implications and Conclusion

The bottom-up approach used in this study has considerable advantages for the identification of distinct stages of health behavior change. Inferring stages from mindsets found in the data does not require equidistance or monotonous increase and still allows for statistical goodness-of-fit tests. Second, potential statistical problems of stage algorithms such as limited reliability are accounted for by the latent variable approach, and finally, a bottom-up-approach might reveal more information about the actual mindsets of individuals in different stages of health behavior change than a confirmatory top-down approach such as comparing means of variables in individuals across a-priori defined stages.

It is not feasible to assess a wide range of variables for an individual and to allocate them to stages based on their profile. Researchers and practitioners who wish to administer stage-matched or tailored interventions need to be able to diagnose an individual's stage fast and accurately, that is, assign them a stage that is a good reflection of their mindset. It follows that it is necessary to continuously ensure the validity of stage measures used (Lippke, Ziegelmann, Schwarzer, & Velicer, 2009; Richert, Lippke, & Schwarzer, 2010).

References

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). New York: Wiley.

Armitage, C. J., Povey, R., & Arden, M. A. (2003). Evidence for discontinuity patterns across the stages of change: A role for attitudinal ambivalence. *Psychology and Health, 18*(3), 373-386. doi:10.1080/0887044031000066553

Armitage, C. J., Sheeran, P., Conner, M., & Arden, M. A. (2004). Stages of Change or Changes of Stage? Predicting Transitions in Transtheoretical Model Stages in Relation to Healthy Food Choice. *Journal of Consulting and Clinical Psychology, 72*(3), 491-499. doi:10.1037/0022-006X.72.3.491

Brug, J., Conner, M., Harre, N., Kremers, S., McKellar, S., & Whitelaw, S. (2005). The Transtheoretical Model and stages of change: A critique - Observations by five commentators on the paper by Adams, J. and White, M. (2004) Why don't stage-based activity promotion interventions work? *Health Education Research, 20*(2), 244-258. doi:10.1093/her/cyg105

Custers, R., & Aarts, H. (2007). In search of the nonconscious sources of goal pursuit: Accessibility and positive affective valence of the goal state. *Journal of Experimental Social Psychology, 43*(2), 312-318. doi:10.1016/j.jesp.2006.02.005

Fujita, K., Gollwitzer, P. M., & Oettingen, G. (2007). Mindsets and pre-conscious open-mindedness to incidental information. *Journal of Experimental Social Psychology, 43*(1), 48-61. doi:10.1016/j.jesp.2005.12.004

Godin, G., Lambert, L.-D., Owen, N., Nolin, B., & Prud'homme, D. (2004). Stages of motivational readiness for physical activity: A comparison of different algorithms of classification. *British Journal Of Health Psychology, 9*, 253-267. doi:10.1348/135910704773891087

Gollwitzer, P. M., Heckhausen, H., & Steller, B. (1990). Deliberative and implemental mind-sets: Cognitive tuning toward congruous thoughts and information. *Journal of Personality and Social Psychology, 59*(6), 1119-1127. doi:10.1037/0022-3514.59.6.1119

Guo, B., Aveyard, P., Fielding, A., & Sutton, S. (2009). Using Latent Class and Latent Transition Analysis to Examine the Transtheoretical Model Staging Algorithm and Sequential Stage Transition in Adolescent Smoking. *Substance Use & Misuse, 44:* 14, 2028 — 2042. doi:10.3109/10826080902848665

Heckhausen, H. (1991). *Motivation and action*. Berlin, Heidelberg, New York: Springer.

Heckhausen, H., & Gollwitzer, P. M. (1987). Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion, 11*(2), 101-120. doi:10.1007/BF00992338

Herzog, T. A. (2008). Analyzing the transtheoretical model using the framework of Weinstein, Rothman, and Sutton (1998): The example of smoking cessation. *Health Psychology, 27*(5), 548-556. doi:10.1037/0278-6133.27.5.548

Lippke, S., Ziegelmann, J. P., Schwarzer, R., & Velicer, W. F. (2009). Validity of Stage Assessment in the Adoption and Maintenance of Physical Activity and Fruit and Vegetable Consumption. *Health Psychology, 28*(2), 183-193. doi:10.1037/a0012983

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika, 83*(3), 767-778. doi:10.1093/biomet/88.3.767

Magidson, J., & Vermunt, J. K. (2004). Latent Class Models. In D. Kaplan (Ed.), *The SAGE Handbook of Quantitative Methodology for the Social Sciences* (pp. 175-198). Thousand Oaks: Sage.

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47*(9), 1102-1114. doi:10.1037//0003-066X.47.9.1102

Prochaska, J. O., Velicer, W. F., Rossi, J. S., Redding, C. A., Greene, G. W., Rossi, S. R., et al. (2004). Multiple risk expert systems interventions: Impact of simultaneous stage-matched expert system interventions for smoking, high-fat diet, and sun exposure in a population of parents. *Health Psychology, 23*(5), 503-516. doi:10.1037/0278-6133.23.5.503

Renner, B., Schüz, B., & Sniehotta, F. F. (2008). Preventive Health Behavior and Adaptive Accuracy of Risk Perceptions. *Risk Analysis, 28*(3), 741-748. doi:10.1111/j.1539-6924.2008.01047.x

Richert, J., Lippke, S., & Schwarzer, R. (2010). Comparison of individual criteria and externally imposed criteria for stage allocation: Findings from an internet study

addressing physical activity. *Measurement in Physical Education and Exercise Science, 14*(4), 225 – 240. doi: 10.1080/1091367X.2010.520242

Richert, J., Lippke, S., & Ziegelmann, J. P. (in press). Intervention-engagement and its role in the effectiveness of stage-matched interventions promoting physical exercise. *Research in Sports Medicine: An International Journal.*

Richert, J., Reuter, T., Wiedemann, A., Lippke, S., Ziegelmann, J.P., & Schwarzer, R. (2010). Differential Effects of Planning and Self-Efficacy on Fruit and Vegetable Consumption, *Appetite, 54(3),* 611-614. doi:10.1016/j.appet.2010.03.006

Scholz, U., Sniehotta, F. F., & Schwarzer, R. (2005). Predicting Physical Exercise in Cardiac Rehabilitation: The Role of Phase-Specific Self-Efficacy Beliefs. *Journal of Sport and Exercise Psychology, 27*(2), 135-151. doi:10.1080/08870440512331317670

Schüz, B., Sniehotta, F. F., Mallach, N., Wiedemann, A. U., & Schwarzer, R. (2009). Predicting transitions from preintentional, intentional and actional stages of change. *Health Educ. Res., 24*(1), 64-75. doi: 10.1093/her/cym092

Schüz, B., Sniehotta, F. F., & Schwarzer, R. (2007). Stage-specific effects of an action control intervention on dental flossing. *Health Education Research, 22*(3), :332-341. doi: 10.1093/her/cyl084

Schwarzer, R. (1992). Self-efficacy in the adoption and maintenance of health behaviors: Theoretical approaches and a new model. In R. Schwarzer (Ed.), *Self-Efficacy: Thought Control of Action* (pp. 217-242). Washington, DC: Hemisphere.

Schwarzer, R. (2008). Some burning issues in research on health behavior change. *Applied Psychology: An International Review, 57*(1), 84-93. doi: 10.1111/j.1464-0597.2007.00324.x

Schwarzer, R., & Luszczynska, A. (2008). How to overcome health-compromising behaviors: The health action process approach. *European Psychologist, 13*(2), 141-151. doi:10.1027/1016-9040.13.2.141

Sniehotta, F. F. (2009). Towards a theory of intentional behavior change: plans, planning, and self-regulation. *British Journal Of Health Psychology, 14*(Pt 2), 261-273. doi:10.1348/135910708X389042

Sniehotta, F. F., Luszczynska, A., Scholz, U., & Lippke, S. (2005). Discontinuity patterns in stages of the precaution adoption process model: Meat consumption during a livestock epidemic. *British Journal Of Health Psychology, 10*(2), 221-235. doi:10.1348/135910705X26137

Sniehotta, F. F., Scholz, U., & Schwarzer, R. (2005). Bridging the intention-behavior gap: Planning, self-efficacy, and action control in the adoption and maintenance of physical exercise. *Psychology And Health, 20s*(2), 143-160. doi:10.1080/08870440512331317670

Sniehotta, F. F., Schwarzer, R., Scholz, U., & Schüz, B. (2005). Action Planning and Coping Planning for Long-Term Lifestyle Change: Theory and Assessment. *European Journal of Social Psychology, 35*(4), 565-576. doi:10.1002/ejsp.258

Sutton, S. (2000). Interpreting cross-sectional data on stages of change. *Psychology & Health, 15*(2), 163-171. doi:10.1080/08870440008400298

Sutton, S. (2001). Back to the drawing board? A review of applications of the transtheoretical model to substance use. *Addiction, 96*(1), 175-186. doi:10.1046/j.1360-0443.2001.96117513.x

Taylor, S. E., & Gollwitzer, P. M. (1995). The effects of mindsets on positive illusions. *Journal of Personality and Social Psychology, 69*, 213-226. doi:10.1037//0022-3514.69.2.213

Velicer, W. F., Martin, R. A., Collins, L. M. (1996). Latent transition analysis for longitudinal data. *Addiction, 91*:197–209. doi:10.1080/09652149638926

Weinstein, N. D. (1988). The precaution adoption process. *Health Psychology, 7*(4), 355-386. doi:10.1037/0278-6133.7.4.355

Weinstein, N. D. (2003). Exploring the links between risk perceptions and preventive health behavior. In J. Suls & K. A. Wallston (Eds.), *Social psychological foundations of health and illness.* (pp. 22-53). Malden, MA: Blackwell Publishers.

Weinstein, N. D., Lyon, J. E., Sandman, P. M., & Cuite, C. L. (1998). Experimental evidence for stages of health behavior change: The precaution adoption process model applied to home radon testing. *Health Psychology, 17*(5), 445-453. doi:10.1037/0278-6133.17.5.445

Weinstein, N. D., Rothman, A. J., & Nicolich, M. (1998). Use of correlational data to examine the effects of risk perceptions on precautionary behavior. *Psychology & Health, 13*(3), 479-501. doi:10.1080/08870449808407305

Weinstein, N. D., Rothman, A. J., & Sutton, S. R. (1998). Stage theories of health behavior: Conceptual and methodological issues. *Health Psychology, 17*(3), 290-299. doi:10.1037/0278-6133.17.3.290

Wiedemann, A. U., Lippke, S., Reuter, T., Schüz, B., Ziegelmann, J. P., & Schwarzer, R. (2009). Prediction of Stage Transitions in Fruit and Vegetable Intake. *Health Education Research, 24*, 596-607. doi:10.1093/her/cyn061

Wood, W., Quinn, J. M., & Kashy, D. A. (2002). Habits in everyday life: Thought, emotion, and action. *Journal Of Personality And Social Psychology, 83*(6), 1281-1297. doi:10.1037//0022-3514.83.6.1281

4

Assessment and Validation of the Task-Engagement Scale (TES)

Jana Richert, &Sonia Lippke

Freie Universität Berlin, Germany

Richert, J., & Lippke, S. (2010). Assessment and Validation of the Task-Engagemen Scale (TES). Manuscript submitted for publication.

Abstract

Background: Engagement is a key determinant of any intervention's effectiveness. Yet a comprehensive definion and measure of engagement are lacking.

Aims: In this paper, a comprehensive definition of engagement is provided and a new measure, the Task Engagement Scale (TES), proposed and tested.

Sample: The study was conducted in the domain of health behavior interventions within a sample of *N*=142 college students.

Methods: Factorial structure and construct validity of the scale were assessed with Mplus.

Results: Main results are that the TES can be used as a composite measure of engagement. Also, any of the four subscales task-compliance, effort, undivided attention and absorption can be measured as individual constructs. Furthermore, the information content of the TES captured more information than the objective measures time-on-task and the completion rate of intervention materials. In addition, associations between theoretically related constructs and the TES score indicate that engagement is also sufficiently distinct.

Conclusions: With the existence of a validated self-report measure of all main components of engagement, researchers and practitioners are able to assess whether or not their messages are received and can subsequently judge interventions' effectiveness (or lack theiroff) more adequately.

Keywords: Intervention-Engagement, CFA, participation, compliance, attention, effort

Assessment and Validation of the Task-Engagement Scale (TES)

Engagement is a key determinant of success (e.g., Celio, Winzelberg, Dev, & Taylor, 2002; Manwaring et al., 2008; e.g., Richert, Lippke, & Ziegelmann, 2009). Students need to be engaged with their schoolwork in order to have academic success (Hawkins & Axelrod, 2008), employees need to be engaged in their job to cause productivity (e.g., Maslach & Leiter, 2008), patients/ clients need to engage with their therapy in order for it to be effective (Strecher et al., 2008) and study participants need to be engaged with the study material for an effect to occur. Thus, the investigation of engagement is of relevance to experts of all kinds of areas, e.g., teachers, coaches, managers, physicians and scientists alike.

Conceptualization of the construct engagement

There is no consistent terminology or established definition of engagement in the literature (Guthrie et al., 2004), which is owing to the multifaceted nature of the construct. In this paper, a comprehensive yet integrative approach is taken to capture the concept and subsequently develop an exhaustive yet parsimonious measure of engagement.

Engagement has three components: a behavioral, a cognitive and an emotional one. The behavioral component comprises, e.g., attendance, participation or *task compliance*, which is the lowest form of engagement as it refers to doing the bare minimum. The cognitive component comprises aspects, such as undivided attention, thoughtfulness and effort to complete the task(s) as well as, e.g., depth of information processing. More specifically, *undivided attention* refers to the avoidance and the disregard of distractions and the restraint from engagement in task-unrelated activities. *Effort* is engagement in the narrowest sense as it shows that individuals are not merely "going through the motions". The emotional component denotes, e.g., *absorption* (the term is adopted from the Utrecht Work Engagement Scale, cf., Schaufeli & Bakker, 2003), which expresses a state of immersion and being engrossed in the task, usually accompanied by feelings of pleasure (c.f., Fredricks, Blumenfeld, & Alison, 2004; Hallberg & Schaufeli, 2006; cf., Matthews et

al., 2002; Shernoff & Vandell, 2007). While in this framework absorption is explicitly understood as facet of engagement, it may also be seen as a consequence of the interplay between compliance, effort and undivided attention.

As people may engage in various domain-specific activities, such as school-, work-, leisure time-, or intervention program activities, the broad term task-engagement will be used to account for this diverse applicability. Thus, task-engagement refers to engagement in a particular task or a set of tasks that is administered in one session (e.g., doing one's homework, reading a text, watching a video, generating and writing down ideas). Intervention programs may be seen as a series of different tasks, so intervention-engagement subsequently is a domain specific form of task-engagement.

Rationale for the development of a self-report measure of task-engagement

Engagement has been assessed with a variety of measures ranging from objective means such as direct observation (e.g., Amato-Zech, Hoff, & Doepke, 2006; Chafouleas, Riley-Tillman, Sassu, LaFrance, & Patwa, 2007), behavioral indicators such as time spent on task (e.g., Richert, Lippke, & Ziegelmann, 2010) and various self-report measures such as the Motivation and Engagement Scale (MES, Martin, 2007) and the Utrecht work engagement scale (UWES, Schaufeli, Martinez, Pinto, Salanova, & Bakker, 2002). However, there are several issues related to these assessment techniques.

Direct observation of individuals. Firstly, the very resource demanding observation of individuals is not always possible (e.g., if a task is administered online). Secondly, if people (for ethical reasons) know that they are being monitored, it is likely that data is going to be skewed one way or another, e.g., due to either the Hawthorne effect (Parsons, 1974) or mock participation (owing to social desirability). Thirdly, ambiguous behaviors, such as staring off into space (suggesting that one is being lost in thought), can be interpreted as indicative of both high and little engagement, because the observer does not know whether those thoughts are task related or not.

Behavioral indicators. Some behavioral indicators, such as the number of generated ideas or products, the frequency of breaks etc., are useful as additional measures of

engagement (Manwaring et al., 2008). As standalone measures, these indicators, such as time spent on task, are unreliable. For example, time in minutes does not allow for any conclusions as to how this time was spent and whether or not it was actually spent on the task (Richert, Lippke, & Ziegelmann, in press). In addition, both types of measures are lacking the ability to assess cognitions or emotions, such as attention and absorption.

Self-report measures. The few validated self-report measures that do exist are domain-specific and do not assess all main components of engagement (cf. the MES-HS, Martin, 2007; Matthews et al., 2002; the UWES, Schaufeli, Martinez, Pinto, Salanova, & Bakker, 2002). However, instruments for intervention-engagement are lacking altogether. Thus, the need for a domain-unspecific, easily applicable, exhaustive yet parsimonious self-report measure of all main components of engagement is apparent.

Relationship with other constructs

As the conceptualization of engagement as is presented in this paper is original, assumptions about associations with other constructs are based on similarities according to the constructs' definitions rather than previous empirical findings.

Social desirability is defined as an individuals' tendency to answer questions in a way that is viewed as favorably by others (cf., Furnham, 1986). It seems reasonable that individuals with a tendency for socially desired behaviors would also comply (or at least want to appear to comply) with a given task. *Conscientiousness* is defined as the tendency to behave in accordance with the dictate of one's conscience (cf., Costa & McCrae, 1992). It seems plausible that conscientious individuals will more likely be compliant. *General self-efficacy* can be understood as individuals' beliefs in their ability to cope with new challenges (Bandura, 1997). If participants believe that they are not able to complete a given task, it is unlikely that they will even try to do the task or put effort into it. *Concentration* is defined as individuals' ability to focus on a task at hand (cf., Schwarzer, 2000) and as such is linked to the construct of undivided attention. *Intrinsic motivation* refers to behaviors that are done simply for the sake of doing them, i.e., the reward lies in the behavior itself rather than some external reward (cf. self-determination theory, Ryan &

Deci, 2000). *Flow* is defined as a state in which an individual is focused on and fully engrossed in a task (cf. Csikszentmihalyi, 1988). So while flow and absorption are very similar, absorption is restricted to the immersion aspect. *General affinity to working on tasks* and *perceived importance* of physical activity goal (Wigfield & Eccles, 2000) as well as *positive outcome expectancies* with regard to the task at hand, i.e., the *perceived value of the intervention* should also have some association with components of engagement.

Research Aim and Hypotheses

The aim of this study was to confirm the hypothesized factorial structure of the construct and to validate the German version of the scale in terms of construct validity (cf., Cadiz, Sawyer, & Griffith, 2009). This was done in the domain of health behavior change interventions.

Hypothesis 1. A second-order factor model (see Figure 1) with all items loading on four first-order factors (the hypothesized subscales) and those four latent factors in turn loading on one second-order g-factor (i.e. engagement) …

a) …fits the data well in terms of the Root Mean Square Error Approximation (RSMEA), the Standardized Root Mean Square Residual (SRMSR), the Comparative Fit Index (CFI) and Tucker-Lewis Index (TFI) and

b) …fits the data significantly better than a one-factor model with all items loading on a single underlying factor (i.e. engagement) and a four-factor model with the items loading on four latent factors (the hypothesized subscales) with regard to the value of the Akaike Information Criterion (AIC), the Tucker-Lewis Index (TLI), as well as the Chi-square difference test.

Hypothesis 2. Engagement is positively associated with the behavioral indicators *time spent on task* and the *completion rate for the intervention material*.

Hypothesis 3. The task-engagement scale and/or its subscales show construct validity. In terms of convergent validity, we expect that…

…social desirability, conscientiousness and general self-efficacy are positively related to task compliance.

…concentration is positively related to undivided attention.

…affinity to working on tasks as well as general self-efficacy are positively associated with effort.

… flow (cf. Csikszentmihalyi, 1988) and intrinsic motivation are positively correlated with absorption.

…perceived importance of one's physical activity goal(s) as well as positive outcome expectancies of participation in the intervention are related positively to each facet of engagement as well as the overall engagement scale.

In terms of divergent validity, we expect that…

…work-engagement will not be related to intervention-engagement.

Method

Procedure

First year undergrad-psychology students were informed of the possibility to earn credits that are required for admission to a mandatory empirical science course by participating in this study. Interested students were asked to schedule a time to come to the computer lab. Over the course of two weeks, students in groups of 15 came to our lab for the first measurement point. After having given informed consent, participants filled in the first part of a paper-pencil questionnaire followed by an online-intervention targeting physical activity and then answering the second part of the paper-pencil questionnaire.

In total, N=142 individuals gave informed consent and were included in the study. All participants in the sample had less than 5% missing data, which were missing completely at random: χ^2 (1, 1553) =1598.44, p>.05 and were imputed with the Expectation Maximization (EM) algorithm (Tabachnick & Fidell, 2007). No missing values occurred in the engagement scale.

Participants

The sample consists of N=142 students, who were on average M=25.91 years old

(Range=19-57, *SD*=6.83) and mostly female (72.3%).

Intervention

An online-intervention targeting physical activity behavior was administered. All materials were developed using the Intervention Mapping approach (Bartholomew et al., 2006; Michie, Johnston, Francis, Hardeman, & Eccles, 2008) and are based on the Health Action Process Approach (Schwarzer, 2008). The intervention content was tailored to participants' behavioral and motivational baseline. In particular, individuals' personal goals with regard to increasing or maintaining physical activity levels (frequency/ duration) were considered. A detailed description of the intervention material, the rationale behind it as well as results on the effectiveness of the treatment will be reported elsewhere.

Measures

Behavioral indicators.

Time spent on task. The time that participants started working on the intervention as well as the time participants finished working on the intervention was recorded automatically by the online system. The intervention was unique for different groups of participants (see section below on intervention) and subsequently the time necessary to complete the intervention differed slightly. In order to control for this difference, the obtained time spent on task (in minutes) was relativized on the average time (in minutes) that participants having received the same intervention needed. The result is a standardized unit less value for time spent on task with a mean of 1.00 and a standard deviation of 0.37 (cf. Richert, Lippke, & Ziegelmann, 2010).

Completion rate for the intervention material. The completion rate was assessed via number of generated ideas and number of answers in the intervention. The obtained value was then relativized on the number of opportunities given in the intervention (as was explained above, interventions were slightly different for groups of individuals). For example, if participants were prompted to identify up to five personal barriers and an individual generated two, their completion rate was 0.4 (5/2).

Scales.

All items are translated from German. Please refer to the cited literature for original wording.

Task-engagement. An item pool reflecting the components of engagement was generated on basis of the literature. The items were refined by a panel of experts. All items are listed in Table 1 (the German originals can be obtained from the authors).

Table 1. Task-engagement Scale (TES): Subscales (in italics) and Items

Participation/ task compliance
(1) "I completed all set tasks."
(2) "I did everything that was required of me within the programme."
Undivided Attention
(1) "I made sure that I could not be distracted (e.g., I turned off my cell phone) or I did not let myself be distracted (e.g., by an incoming call)."
(2) "I was preoccupied with other things, my mind was not in it." (reverse recoded)
(3) "I took a lot of breaks (e.g., to go to the restroom, to prepare food)." (reverse recoded)
Effort
(1) "I thoroughly thought about the tasks."
(2) "I put a lot of effort into completing the tasks."
Absorption
(1) "I was so immersed, I completely forgot everything else around me."
(2) "Time flew by."

Items are to be rated on a 6-point Likert scale ranging from (1) not at all true to (6) exactly true. The psychometric properties are reported in the results section.

Social desirability. Individuals' tendency to answer questions in a way that is viewed as favorably by others was assessed with a shortened version of the SES-17 (Stoeber, 1999). An example item reads as follows: "In a dispute, I am always impartial and factual." Answers to the four items were given in a 'true' or 'false' format. The scale is rather heterogeneous (Cronbach's α = .26), however, previous studies have demonstrated good psychometric properties, which is why a mean over all items was computed (cf., Stoeber, 1999).

Conscientiousness. This construct was assessed with the twelve item subscale of the NEO-FFI (Borkenau &Ostendorf, 1993). An example item is: "When I have committed myself to something, people can rely on me." Answers were given on a five point Likert scale with responses ranging from 'strong disagreement' to 'strong agreement'. Cronbach's Alpha as an indicator of internal consistency was .82.

General self-efficacy. Individuals' beliefs in their ability to cope with new challenges was assessed with four items (Cronbach's α = .54) taken from Schwarzer's general self-efficacy scale (Schwarzer, 1999). Example items are "I can solve difficult problems when I really try." and "I have no trouble achieving my goals.". Answers were given on a four point Likert scale ranging from 'not true' to 'very true'.

Concentration. Individuals' ability to focus on a task at hand was measured with four items (Cronbach's α = .59) taken from Schwarzer's self-regulation scale (Schwarzer, 2000). An example item reads as follows: "When necessary, I can concentrate on something for a long time." Responses were given on a four point Likert scale ranging from 'not true' to 'very true'.

General affinity to working on tasks. Participants were asked to respond to the statement: "In general, I find working on tasks…" with answers on a four point Likert scale ranging from 'very boring' to 'very interesting' (cf. subscale on interest, Wigfield & Eccles, 2000).

Flow. This construct was assessed with six items taken from the flow scale (Rheinberg et al., 2007). Students were asked to apply the following statements to the program they just completed. Item examples read: "I am lost in thought." and "I don't

notice how time goes by." Cronbach's Alpha was sufficiently high with $\alpha = .72$. Responses were given on a six-point Likert scale ranging from 'not at all true' to 'exactly true'.

Intrinsic motivation. Following the question "Why did you participate in this program?", students were asked to rate four statements that are adapted from the work extrinsic and intrinsic motivation scale (Tremblay et al., 2009) on a seven point Likert scale ranging from 'not at all true' to 'exactly true' (Cronbach's $\alpha = .62$). Statement examples are: "Because I was asked to." (reverse recoded) or "Because I enjoy learning new things.".

Importance of physical activity goal. On a four-point Likert scale, students rated "achieving (their own) physical activity goal(s)" as being 'not important' to 'very important' and their own physical activity goal(s) as 'very low' to 'very high up' in their hierarchy of personal goals. Both items correlate with $r = .67$, $p < .001$ (cf., Wigfield & Eccles, 2000).

Positive outcome expectancy. The perceived value of the intervention was assessed with the item: "How useful are the things you learned in this program?" rated on a four-point Likert scale ranging from 'not at all useful' to 'very useful' (cf., Wigfield & Eccles, 2000).

Work-engagement. This construct was assessed with the 17 items of the Utrecht Work Engagement Scale (Schaufeli & Bakker, 2003). The scale has high internal consistency with $\alpha = .92$. The three subscales *vigor* ("I am full of energy when I am working."), *dedication* ("My work inspires me.") and *absorption* ("When I am working, I forget everything around me.") were sufficiently homogeneous with Cronbach's Alphas ranging from .80 to .83. All items were rated on a seven-point Likert scale ranging from 'never' to 'always'.

Analyses

All analyses were conducted with Mplus 5.

Results

Descriptives

Normality. All scales deviate significantly from normality (see Table 2) in the way that values are clustered at the higher end of the distribution.

Reliability/ Internal Consistency. Despite its heterogeneous nature, the engagement scale (consisting of all nine items) demonstrated good internal consistency, with Cronbach's $\alpha = .80$. The great majority (75%) of the inter-item correlations clustered around $r = .12$ - .38, indicating that the items are sufficiently differentiating and not redundant with one another. There were no negative inter-item correlations. Table 2 displays descriptive values including inter-item correlations and Cronbach's Alphas for the engagement scale and its four subscales.

Table 2. Desciptive Statistics, Cronbach's Alphas or Inter-Item Correlations for the TES and its Four Subscales, respectively.

	min.	max.	M	SD	r	α	Skewness	Kurtosis	KS-D
Engagement	2.00	5.89	4.58	0.70		.80	-1.10	1.67	.13**
Compliance	1.00	6.00	5.16	1.15	.86**		-1.58	1.97	.26**
Effort	2.00	6.00	4.23	0.89	.70**		-0.41	0.04	.13**
Attention	2.33	6.00	4.99	0.79		.60	-0.87	0.37	.16**
Absorption	1.00	6.00	3.75	1.12	.57**		-0.54	-0.35	.20**

Note. Attention, undivided attention; min., minimum; max., maximum; r, inter-item correlation; α, Crohnbach's Alpha; KS-D, Kolmogorov-Smirnov D; **p<.01, *p<.05.

Table 3. Intercorrelations between the four subscales of the TES

	Effort	Attention	Absorption
Compliance	.33**	.34**	.70**
Effort		.31**	.44**
Attention			.38**

Note. ** p<.01..

In a regression model with compliance, effort and attention as predictors and absorption as outcome, 27% of variance were explained by the model**.**

Hypothesis 1

Confirmatory Factor Analysis. Overall, the hypothesized second-order factor model with all items loading on four first-order factors and those four latent factors in turn loading on one second-order *g*-factor fit the data very well when evaluated in terms of the recommended cut-offs (cf., Hu & Bentler, 1998). For purposes of comparison, we contrasted the hypothesized model with a one-factor model, in which all of the items were set to load on a single underlying factor and also a four-factor model, in which the items were set to load on four latent factors (our hypothesized subscales). All results are reported in Table 4.

Table 4. Measurement Model Fit Statistics

Model	χ^2	*df*	RMSEA (90% CI)	SRMR	CFI	TLI	AIC
One-factor model	218.77**	27	.22 (.20 - .25)	.15	.59	.45	3600.78
Four-factor model	33.77	22	.06 (<.01 - .10)	.10	.98	.96	3425.15
2nd-order factor model	26.00	23	.03 (<.01 - .08)	.04	.99	.99	3415.32

Note: $N = 142$; χ^2, chi-square statistic; RMSEA, root mean square error approximation; SRMR, standardized root mean square residual; CFI, comparative fit index; TLI, Tucker–Lewis index; AIC, Akaike Information Criterion; ** $p<.001$

One factor loading for each latent construct was fixed to one for model identification.

The hypothesized model fit the data slightly better than the four-factor model (Chi² difference test: $\Delta_{\chi 2}(1) = 7.13$, $p < .05$). Considerable correlations between the four latent factors ($r = .36$ - .57) underscore the result that while working well on their own, the four factors are also associated with each other in the way that they share an underlying second-order *g*-factor.

Both models fit the data significantly better than the alternative one-factor model in terms of Chi² difference test results: $\Delta_{\chi 2}(5) = 185.64$ and $\Delta_{\chi 2}(4) = 192.77$, respectively (*p*s < .001) as well as in terms of the fit statistics and AIC (cf. Table 4). The result for the hypothesized model is visualized in Figure 1.

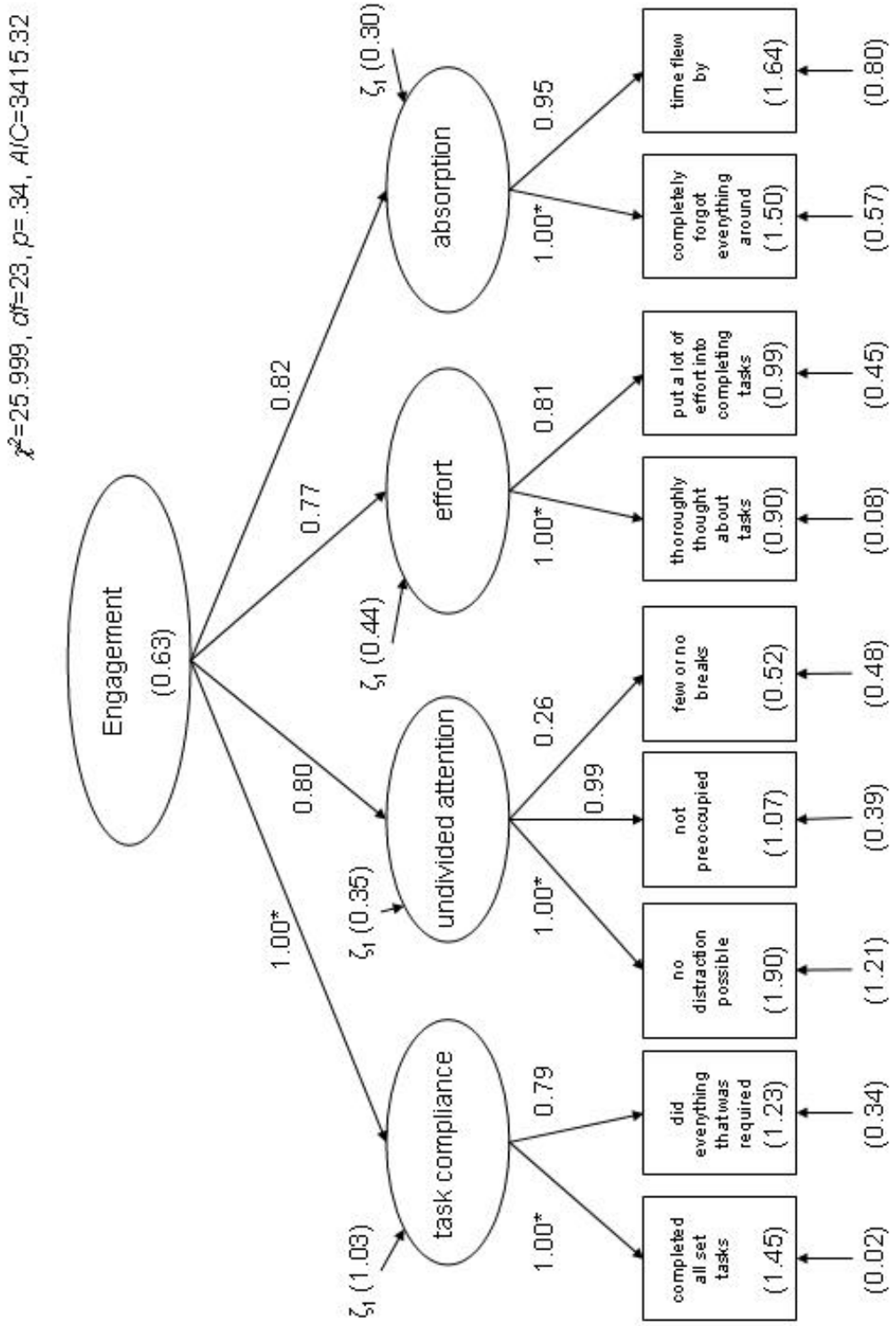

Figure 1. Hypothesized model with the four first-order factors task compliance, undivided attention, effort and absorption loading on the second-order g-factor intervention-engagement. The displayed values are unstandardized Mplus parameter estimations. The values marked with an asterix are loadings that are fixed to 1 for indentification purposes. All other loadings are statistically significant at the p<.001 level. Values in brackets are estimated variances and residual variances, respectively.

Hypothesis 2

The correlation between time on task and the completion rate for the intervention material is fairly high with $r = .59$, $p<.01$. Engagement is positively associated with time spent on task as well as with the completion rate for the intervention material (see Table 5). No association was found for time spent on task and undivided attention as well as for the completion rate for the intervention material and absorption.

Table 5. Correlations Between the TES and Its Four Subscales and All Other Measures.

	TES total	Subscale compliance	Subscale effort	Subscale attention	Subscale absorption
Behavioral Indicators					
Time spent on task	**.27****	**.22****	**.34****	.08	**.19***
Completion rate	**.22****	**.17***	**.25****	**.15***	.08
Psychometric Scales					
Social desirablity	.10	.12	.11	.01	.07
Conscientiousness	.11	**.14***	.02	.10	.03
General self-efficacy	**.23****	**.19***	**.16***	.08	**.24****
Concentration	.10	.10	-.01	.03	**.17***
Affinity to working on tasks	**.40****	**.23****	**.31****	**.23****	**.39****
Flow	**.58****	**.35****	**.37****	**.35****	**.59****
Intrinsic motivation	**.37****	**.23****	**.33****	**.17***	**.35****
Importance of physical activity goal	**.14***	.12	**.15***	.02	**.15***
Positive outcome expectancy	**.36****	.10	**.45****	**.19***	**.34****
Work-engagement	**.29****	**.24****	**.17***	**.22****	**.19***
Dedication	**.25****	**.17***	**.15***	**.26****	**.17***
Vigor	**.31****	**.27****	**.18****	**.26****	**.17***
Absorption	**.23****	**.21****	**.15***	.13	**.18***

Note. completion rate, completion rate for the intervention material; attention, undivided attention; **$p<.01$, *$p<.05$

Hypothesis 3

Table 5 shows all correlations between the TES as well as its four subscales and all other measures. Associations were low to moderately high (correlations ranging from .14 for conscientiousness and compliance to .59 for flow and absorption). Social desirability was not associated with any component of engagement.

Discussion

The aim of this study was to confirm the hypothesized factorial structure of the task-engagement construct and to validate the German version of the TES in the domain of health behavior change interventions.

Hypothesis 1

We hypothesized that all items would load on four distinct latent factors, namely task compliance, undivided attention, effort and absorption and that these factors would share a common *g*-factor, namely engagement. The four factor model fit the data very well demonstrating that each subscale can be used on its own. The found inter-factor correlations indicate that while the four factors are sufficiently distinct, they are also related to one another and the confirmatory factor analysis for the second-order factor model gives evidence that the four factors indeed share a common second order *g*-factor: engagement.

Hypothesis 2

Engagement is positively associated with time spent on task as well as with the completion rate for the intervention material, illustrating that both behavioral indicators may be used as additional measures of engagement. As a standalone measure, time spent on task is a rather weak measure of task-engagement: A conclusions cannot easily be drawn as to how this time was spent (Richert, Lippke, & Ziegelmann, in press).

In this study, no association was found between time spent on task and undivided attention, e.g., indicating that while working on the tasks, individuals may have been preoccupied with task-unrelated thoughts, which in turn denotes less engagement. Likewise, absorption was not related to the completion rate, which shows that generating a few ideas (e.g., a detailed action plan for one's physical activity behavior) may be sufficient to lead to absorption. In that sense, the notion of "the more the better" does not

hold true universally. These results strengthen the notion that behavioral indicators are valuable add-on measures, but no sufficient indicator of all components of engagement.

Hypothesis 3

Based on similarities in definitions, we expected engagement and/or it's subscales to be associated with other constructs. We assumed these correlations to be of medium size, as all constructs, while being related, were supposed to be distinct from each other nonetheless. In the following, each result is discussed in more detail.

It seemed reasonable to expect that individuals with a tendency for socially desired behaviors would also comply (or at least want to appear to comply) with a given task (cf., Furnham, 1986). However, this assumption is only reasonable, if participants evaluate engaging in the intervention as something that is favored by others. This is something we did not control for. Thus it can only be assumed that no association between social desirability and compliance indicate that study participants did not consider engaging in the intervention something that would be favored by others.

As expected, conscientiousness (cf., Costa & McCrae, 1992) was related to task-compliance. The fact that the association was small emphasizes that there are additional factors that explain individuals' compliance in an intervention.

As was anticipated, general self-efficacy was positively associated with individuals' task-compliance and subsequent effort in the intervention. Surprisingly, it was also associated with participants' absorption. This is explicable with the notion that absorption is not only a further and distinct component of engagement (cf. results and discussion of hypothesis 1), but can also be seen as a byproduct of the interplay between compliance, effort and undivided attention. The substantial intercorrelations and the fact that together, compliance, effort and undivided attention were able to explain 27% of the variance in absorption support this idea. However, as all scales were assessed at the same point in time, no causal inferences can be made.

In this sample, no association between trait concentration and attention in the intervention was found. This may be due to the fact that undivided attention was extremely

high among all participants. This might have been because of the setting the study was conducted in (see also section above on normality).

As expected, individuals' general affinity to work on tasks, regardless of the topic, was associated with their task-engagement. Again, the fact that correlations are low to moderate indicates that other factors explain variation in task engagement over and above one's general affinity to work on tasks. Absorption and flow share about one third of their variance, indicating that both scales are related, while being sufficiently distinct from one another (cf. Csikszentmihalyi, 1988).

If the topic of the intervention and the tasks at hand are irrelevant to the participants and if they do not expect participation in the intervention to be useful to them, it is unlikely that they will engage with the tasks. For this reason, it was expected that perceived importance of participants' physical activity goals (i.e. of the topic of the intervention) as well as positive outcome expectancies with regard to participating in the intervention are associated with engagement and its components. Compliance was not related to either construct, which indicates that while effort, attention, absorption and overall engagement are associated with something like an overall sense of purpose of the intervention, compliance is not affected by it. Whether or not individuals comply with a given task is affected by other factors.

It is reasonable to assume that engagement in one area of one's life (e.g., school work) is not associated with engagement in any other distinct area of one's life (e.g., a physical activity intervention). Accordingly, the work engagement scale and the TES may not be correlated (cf., Schaufeli & Bakker, 2003). However, students received credits for participation in this study and thus, the study may have been understood as part of their school work. This may explain why both scales were associated. Also, all correlations were low, clustering around .17, indicating that while there is a small relationship between both types of engagement, and they are clearly distinct from one another. Concluding, it can be argued that engagement and its four individual components show sufficient construct validity in terms of convergent but also divergent validity.

Normality

Participants anticipated receiving credits for completing the study. In addition, the fact that the study was conducted in a group setting may have caused students to be more thorough in their completion of tasks than they would have been in a more anonymous context. Also, distractions were highly unlikely in this setting. These circumstances may explain why all engagement subscales were skewed in the direction of higher values. This should not be seen as assault of the psychometric characteristics of the scale. Rather, it may be a quality indicator. However, the study setting influenced the results and they should be replicated in different settings where engagement is likely to be more evenly distributed (e.g., students could be given a lengthy homework task and might be told that the homework would be submitted anonymously).

Missing values in the task-engagement scale

While in this study, no missing values occurred in the TES (cf. section on normality for explanation), it is expected that in a more representative sample, missing values might be a problem. Task-engagement is related to questionnaire-engagement (i.e. the response rate in the engagement scale) and the association will be the stronger the more similar both tasks are.

This raises the question of how to deal with missing values in the engagement scale in general. If a significant association between both types of engagement (in this case assessed via percentage of unplanned missing values in the intervention material and the scale) is found, then missing values in the engagement scale are not at random and thus imputation is no legitimate treatment option (Tabachnik & Fidell, 2006). Another common way of dealing with unplanned missings is the exclusion of cases. This, however, will lead to analyzing a positively skewed sample (i.e. only engaged individuals are included in the analyses) thus distorting the results. So alternatively, unplanned missing values in the engagement scale should be used as an indicator of low task engagement, i.e. instead of coding unplanned missing as missing values, they should be coded as zeros.

On a critical note, it may be argued that missing values may not be unplanned, in which case coding them as zeros would lead to distortion of the data as well. However, planned missing values in the questionnaire are highly unlikely as (1) all participants do receive the scale, (2) comprehension of the items was tested in a pilot study and (3) the items are not addressing intimate or sensitive matter.

Reliability

The engagement scale as well as each subscale demonstrated good internal consistency. In the future, retest-reliability might be assessed.

Conclusion and Outlook

The task-engagement scale (TES) may be used as a composite measure of all four components of engagement. Further, depending on the research question at hand, any of the four facets may be measured as individual constructs. Also, the information content of the TES goes beyond that which objective measures such as time-on-task or the completion rate of intervention materials can offer. Furthermore, initial data on construct validity of the measure and its subscales are promising. It could be demonstrated that while there is an association between theoretically related constructs, engagement is also sufficiently distinct.

The current literature indicates that useful interventions are not effective if participants do not engage (e.g., Richert, Lippke, & Ziegelmann, 2009). Unfortunately, psychologists often restrict their research to the content of the message (i.e. testing theories of health behavior change) and depending on the treatment's effectiveness draw conclusions that may be inappropriate (e.g., that the intervention did not address the right variables) when it may just be a matter of lack of participants' engagement.

With the existence of a validated self-report measure of all main components of engagement, researchers and practitioners are now able to assess whether or not their messages are "being heard". Also, they may now (better) examine the relationship between engagement and desired outcomes.

In the domain of behavior change interventions, questions that are of high relevance to theory development and intervention design are: Does participants' engagement mediate the association between treatment and behavior change? Does engagement moderate the treatment – changes in cognitions – changes in behavior mediation? Which variables predict engagement and how can we manipulate these predictors to bring about changes in engagement? Future research needs to address these questions to further knowledge on health behavior change and with the presented measure this endeavor is now feasible. With the current scale such research might be conducted effectively.

References

Amato-Zech, N. A., Hoff, K. E., & Doepke, K. J. (2006). Increasing on-task behavior in the classroom: Extension of self-monitoring strategies. *Psychology in the Schools, 43(2),* 211-221.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Borkenau, P., Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI) nach Costa und McCrae.* Göttingen: Hogrefe.

Cadiz D., Sawyer, J. E., & Griffith, T. L. (2009). Developing and Validating Field Measurement Scales for Absorptive Capacity and Experienced Community of Practice. *Educational and Psychological Measurement*, *69*, 1035 - 1058.

Celio, A. A., Winzelberg, A. J., Dev, P., & Taylor, C. B. (2002). Improving Compliance in Online, Structured Self-help Programs: Evaluation of an Eating Disorder Prevention Program. *Journal of Psychiatric Practice, 8(1)*, 14-20.

Chafouleas, S. M., Riley-Tillman, T. C., Sassu, K. A., LaFrance, M. J., & Patwa, S. S. (2007). Daily Behavior Report Cards: An Investigation of the Consistency of On-Task Data Across Raters and Methods. *Journal of Positive Behavior Interventions, 9(1)*, 30-37.

Costa, P.T., Jr. & McCrae, R.R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual.* Odessa, FL: Psychological Assessment Resources.

Csikszentmihalyi, M. (1988). The flow experience and its significance for human psychology. In M. Csikszentmihalyi (Ed.) *Optimal experience: psychological studies of flow in consciousness* (pp. 15-35). Cambridge, UK: Cambridge University Press.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68-78.

Evers, A., Klusmann, V., Schwarzer, R., Dimeo, F., Reischies, F. M., & Heuser, I. (2009). Time effort of older women in a 6-month mental or physical intervention: Who benefits? *Psychology & Health, 24(1),* 166.

Fredricks, J. A., Blumenfeld, P. C., & Alison, H. P. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research, 74(1)*, 59-109.

Furnham, A. (1986). Response bias, social desirability, and dissimulation, *Personality and Individual Differences, 7,* 385–400.

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., et al. (2004). Increasing Reading Comprehension and Engagement Through Concept-Oriented Reading Instruction. *Journal of Educational Psychology, 96(3)*, 403-423.

Hallberg, U. E., & Schaufeli, W. B. (2006). 'Same Same' But Different? Can Work Engagement Be Discriminated from Job Involvement and Organizational Commitment? *European Psychologist, 11(2)*, 119-127.

Hawkins, R. O., & Axelrod, M. I. (2008). Increasing the on-task homework behavior of youth with behavior disorders using functional behavioral assessment. *Behavior Modification, 32(6)*, 840-859.

Hu, L. and Bentler, P.M. (1998). Fit indices in covariance structure modeling: sensitivity to under parameterized model misspecification. *Psychological Methods, 3*, 424–453.

Manwaring, J. L., Bryson, S. W., Goldschmidt, A. B., Winzelberg, A. J., Luce, K. H., Cunning, D., et al. (2008). Do adherence variables predict outcome in an online program for the prevention of eating disorders? *Journal of Consulting and Clinical Psychology, 76(2),* 341-346.

Martin, A. J. (2007). Examining a multidimensional model of student motivation and engagement using a construct validation approach. *British Journal of Educational Psychology, 77(2),* 413-440.

Maslach, C., & Leiter, M. P. (2008). Early predictors of job burnout and engagement. *Journal of Applied Psychology, 93(3),* 498-512.

Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., et al. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion, 2(4),* 315-340.

Parsons, H. M. (1974). What happened at Hawthorne? *Science, 183*, 922-932.

Rheinberg, F., Manig, Y., Kliegl, R., Engeser, S., Vollmeyer, R. (2007). Flow bei der Arbeit, doch Glück in der Freizeit: Zielausrichtung, Flow und Glücksgefühle. *Zeitschrift für Arbeits- u. Organisationspsychologie, 51(3),* 105 – 115.

Richert, J., Lippke, S., & Ziegelmann, J. P. (in press). Intervention-Engagement and its role in the effectiveness of stage-matched interventions: An online-study on physical activity. *Research in Sports Medicine: An International Journal*.

Schaufeli, W. B., Martinez, I. M., Pinto, A. M., Salanova, M., & Bakker, A. B. (2002). Burnout and Engagement in University Students: A Cross-National Study. *Journal of Cross-Cultural Psychology, 33(5)*, 464-481.

Shernoff, D. J., & Vandell, D. L. (2007). Engagement in after-school program activities: Quality of experience from the perspective of participants. *Journal of Youth and Adolescence, 36(7),* 891-903.

Stoeber, J. (1999). Die Soziale-Erwünschtheitsskala-17 (SES 17): Entwicklung und erste Befunde zu Reliabilität und Validität. *Diagnostica, 45(4)*, 173-177.)

Strecher, V. J., McClure, J., Alexander, G., Chakraborty, B., Nair, V., Konkel, J., et al. (2008). The role of engagement in a tailored web-based smoking cessation

program: randomized controlled trial. *Journal of Medical Internet Research, 10(5),* 23-32.

Schwarzer, R., & Jerusalem, M. (Eds.) (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*. Berlin: Freie Universität Berlin.

Schwarzer, R. (2000). *Streß, Angst und Handlungsregulation* (4th rev. ed.). Stuttgart: Kohlhammer. bzw. Schwarzer, R. (1996). *Psychologie des Gesundheitsverhaltens* (2nd rev. ed.). Göttingen: Hogrefe.

Tabachnik, B., & Fidell, L. S. (2006). *Using Multivariate Statistics* (5th ed.): Allyn & Bacon.

Tremblay, M. A., Blanchard, C. M., Taylor, S., Pelletier, L. G., & Villeneuve, M. (2009). A Work Extrinsic and Intrinsic Motivation Scale: Its Value for Organizational Psychology Research. *Canadian Journal of Behavioural Science, 41(4),* 213–226.

Wigfield, A., Eccles, J. S. (2000). Expectancy–Value Theory of Achievement Motivation. *Contemporary Educational Psychology, 25,* 68–81.

5

Intervention Engagement in Behaviour Change: How to Plan Fruit and Vegetable Consumption

Jana Richert, Sonia Lippke, Daniela Lange,
Linda Parschau, & Ralf Schwarzer

Freie Universität Berlin

Richert, J., Lippke, S., Lange, D., Parschau, L., & Schwarzer, R. (2010). Intervention Engagement in Behavior Change: How to Plan Fruit and Vegetable Consumption, Manuscript submitted for publication.

Abstract

Planning interventions might lead to changes in fruit and vegetable consumption through changes in planning cognitions. However, this mechanism might only hold true for those participants who are engaged in the intervention (i.e., who show compliance, effort, attention, and absorption). This study investigated the role of engagement in the process of behavior change.

A sample of 701 participants (*M*=38.71 years old, 81% women) was randomly assigned to a web-based planning intervention or to a web-based control condition. Intervention engagement, action- and coping planning as well as fruit and vegetable consumption were assessed with self-report measures. Multiple mediation and moderated mediation analyses were carried out.

The effect of intervention condition (T1) on changes in fruit and vegetable consumption (T3) was fully mediated by changes in action- and coping planning (T2). Engagement emerged as a moderator in the way that intervention led to changes in planning, when intervention engagement was at a moderate level. Examining participants' intervention engagement allows for a more careful evaluation of why some interventions work and others do not.

Keywords: Moderated Mediation, Randomized Controlled Trial, Intentions, Planning, Nutrition

Intervention Engagement in Behavior Change:

How to Plan Fruit and Vegetable Consumption

The leading causes of morbidity and mortality in the developed world are modifiable risk behaviors, such as physical inactivity or an unhealthy diet. Developing effective behavior change interventions is therefore one of the key endeavors of health psychologists (cf., Dombrowski, Sniehotta, Avenell, & Coyne, 2007). One target behavior that is considered crucial for preventive health interventions is a diet consisting of mainly low energy dense foods, such as fruits and vegetables (WHO, 2002).

Behavior Change Interventions

Behaviour change interventions (or treatments) are complex affairs comprised of many different components that can be clustered into the two categories *material* and *implementation*. Intervention material consists of some content (i.e., the variables that are addressed in the material, such as self-efficacy), various change techniques (e.g., model learning) as well as mode(s) of delivery (e.g., puppets, video, brochures), and it is theories of behavior change that propose and accumulate knowledge about these topics (cf., Michie, Johnston, Francis, Hardeman, & Eccles, 2008). The treatment implementation process is comprised of delivery fidelity (e.g., compliance with protocol, cf., Hardeman, Michie, Fanshawe, Prevost, McLoughlin, & Kinmonth, 2008), successful dissemination among the target population (cf., Crutzen, 2010) and intervention engagement (Richert, Lippke, & Ziegelmann, in press).

Often, when evaluating the effectiveness of an intervention, researchers restrict their investigation to the treatment content (a question concerning intervention material only) and conclude that depending on the outcome of the intervention study, they either addressed the right or the wrong determinants of behavior change (Richert, & Lippke, under review). Looking at factors concerning both the material and its implementation would allow for more sound conclusions as to why one intervention is effective in modifying behavior and another is not. Such an evaluation approach also allows for the development of more consistent and comprehensive causal models of behavior change.

Planning as the Target of Behavior Change Interventions

Behavior change interventions aim at modifying risk behaviors via addressing predictors of behavior change. One predictor that has been identified as facilitator of health behavior change is planning (Sniehotta, 2009; also cf., implementation intentions, Gollwitzer, 1999). Through planning, individuals notice and create opportunities for engaging in intended behaviors. As such, planning is a prospective self-regulatory strategy. In the Health Action Process Approach (Schwarzer, 1992, 2008), two types of planning are differentiated. *Action planning* refers to the precise specification of the situation, in which the behavior will be performed. *Coping planning* refers to the anticipation of barriers that may hinder behavior performance and subsequently identifying strategies that help the individual cope with said barriers (Sniehotta et al., 2006). Action and coping planning work best in orchestration (cf., Wiedemann, Schüz, Sniehotta, Scholz, & Schwarzer, 2009). In practice, this finding needs to result in the implementation of combined planning interventions. Statistically, their simultaneous influence needs to be accounted for by multiple mediation models rather than single mediation models (cf. Wiedemann et al., 2009).

For a wide range of behaviors, it has been demonstrated that interventions prompting planning lead to changes in behavior (cf., Armitage, 2008; Chapman, Armitage, & Norman, 2009; Gollwitzer, & Sheeran, 2006; Kellar, & Abraham, 2005; Kwak, Kremers, van Baak, & Brug, 2007; Lippke, Ziegelmann, & Schwarzer, 2004; Luszczynska, 2006; Sniehotta et al., 2006; Van Osch, Lechner, Reubsaet, Wigger, & de Vries, 2008). The proposed mechanism hereby is that a planning intervention exerts its influence on changes in behavior via changes in planning rather than via changes in other social-cognitive variables such as self-efficacy (cf., Sniehotta, 2009).

This mediation mechanism might differ in subgroups of participants (e.g., the mediation might hold true for older adults but not for children) which may account for some studies' failure to support the usefulness of a planning intervention for behavior change (Jackson et al., 2005; Michie, Dormandy, & Marteau, 2004; Rutter, Steadman, & Quine, 2006; Skår, Sniehotta, Molloy, Prestwich, & Araújo-Soares, in press). The given

explanation represents a case of *moderated mediation* (cf., MacKinnon, 2008; Preacher, Rucker, & Hayes, 2007). If a moderator variable is dichotomous (e.g., gender), then a mediation in one group (e.g., in the group of women) and lack of mediation in the other (e.g., in the group of men) reflects a moderated mediation. If a moderator variable is continuous, then a moderated mediation is expressed as an interaction between the moderator and the predictor or between the moderator and the mediator, respectively (MacKinnon, 2008). A variable that constitutes a putative moderator of the effect that a planning intervention exerts on changes in behavior via changes in planning is intervention engagement.

Intervention Engagement as Moderator

Over and above compliance with an assignment, the concept *engagement* additionally comprises undivided attention, effort and, to some degree, absorption in a given task, such as a health promotion intervention. As such, it is more than just being physically present or going through the motions; it is to employ the provided time for its designated purpose and means to avoid distractions as well as engagement in goal unrelated actions (cf. Richert, & Lippke, under review).

There is some evidence that engagement is predictive of intervention success (e.g., Celio, Winzelberg, Dev, & Taylor, 2002; Manwaring et al., 2008; Richert, Lippke, & Ziegelmann, 2010; Ruiter, Kessels, Jansma, & Brug, 2006; Strecher et al., 2008;). However, terms, conceptualization and assessment methods have been inconsistent in previous research (cf., Richert, & Lippke, under review) and generalization of results should be made with caution.

Engagement is expected to moderate the effect that a health promotion intervention exerts on changes in behavior via changes in planning, because the intervention should not lead to changes in cognitions (e.g., action- and coping planning) unless participants engage in it (Richert, Lippke, & Ziegelmann, in press). This represents a case of *stage 1* moderated mediation, whereby the moderator interacts with the predictor (cf., Preacher, Rucker, & Hayes, 2007).

Aim of Study and Hypotheses

Over and above replicating evidence for the planning mediation mechanism (i.e., a question concerning intervention *material*), the aim of the current study was to examine the role of intervention engagement in this process (i.e., a question concerning intervention implementation). By this integrative approach, we aimed at an evaluation of a planning intervention's usefulness in increasing fruit and vegetable consumption that goes beyond that which has been done in previous research. Our specific hypotheses were:
(1) The effect that an intervention prompting action- and coping planning has on changes in fruit and vegetable consumption is fully mediated via changes in action- and coping planning (multiple mediation).
(2) The mediation effect of intervention on changes in behavior via changes in (a) action planning and (b) coping planning varies for different values of intervention engagement (moderated mediation).

Method

Procedure

Study participants of this online study were recruited by personal invitations, press releases (radio, newspaper, TV) and advertisements posted on the university website. After having given informed consent, participants followed a link to a baseline questionnaire. They were then randomly allocated to a waiting control group (questionnaires only), a experimental group, or a control condition (T1). As part of the baseline questionnaire, participants were asked for their personal goal concerning fruit and vegetable intake (in portions per day) and to specify a date by which they wanted to have achieved this goal (participants were able to specify any date within two months of their baseline assessment). One week (T2) (and one month, respectively (T3)) after this personal deadline, participants received an eMail invitation for the post- (and respectively follow-up) assessment. As incentive for study participation, individuals were able to participate in an optional raffle in which they could win attractive gift certificates for an online book store.

Participants

Inclusion criteria for this study were the completion of all three assessments and the participation in an intervention (as opposed to being part of the waiting control group). The final study sample consists of *N*=701 individuals, who were on average *M*=38.71 years old (range=13-79, *SD*=12.84) and mostly women (80.8%). The majority of the sample was highly educated (43.6% College degree), employed (64%) and in a steady relationship (59%).

Interventions

An online-intervention targeting fruit and vegetable consumption was administered. All materials were developed using the Intervention Mapping approach (Bartholomew, Parcel, Kok, & Gottlieb, 2006) and are based on the Health Action Process Approach (Schwarzer, 1992, 2008).

Experimental groups

Participants in the experimental group received an online intervention prompting action- and coping planning. In particular, participants were asked to commit to a specific personal goal with regard to fruit and vegetable consumption and write it down (e.g., to eat 5 portions of fruits and vegetables daily by next month). Individuals were then prompted to specify opportunities (where and when) for a smaller initial sub goal, such as one piece of fruit a day by the end of this week. If participants' end goal was rather simple, they specified opportunities for that instead. Additionally, individuals were asked to identify opportunities for preparatory behaviors (such as buying and preparing foods) and to write it all in a calendar that they could print out if desired. Participants were encouraged to try out their planned behavior and gain experience and to then review and potentially revise their set goals (empty calendars were provided). In little vignettes, role models identified five common situations that may pose a challenge and provided solutions to overcome these obstacles. Subsequently, individuals were prompted to indentify up to three personal barriers and to find strategies to overcome them.

Control condition

Individuals in the control condition received another kind of intervention that also targeted fruit and vegetable consumption, was comparable in length, employed similar strategies (e.g., generating and writing down of ideas, reading about role models' experiences) and modes of delivery (web-based texts and pictures) but targeted different constructs that are typically addressed in standard care interventions on fruit and vegetable consumption (e.g., risk perception and outcome expectancies).

Measures

Intervention Engagement. Intervention Engagement was measured right after participants finished their intervention. The construct was assessed with the Task-Engagement Scale (TES, cf., Richert, & Lippke, under review). Example items are: "I completed all set tasks.", "I was preoccupied with other things; my mind was not in it." (reverse recoded), "I put a lot of effort into completing the tasks.", and "I was so immersed, I completely forgot everything else around me.". All statements were rated on a 6-point Likert scale ranging from (1) *not at all true* to (6) *exactly true*. Cronbach's Alpha was satisfactory with .73. Missing values in this scale were replaced with zero (for rational, please cf., Richert, & Lippke, under review).

Action Planning. Action planning was assessed with three items at baseline (T_1) and at post-test (T_2). Items read: "I have planned precisely… (1)…which fruits and vegetables I will eat, (2)… at which occasions (in which situations) I will eat fruits and vegetables, and (3)… how I will eat my fruits and vegetables (e.g., cooked, cut up)." Participants rated these statements on a 6-point Likert scale ranging from (1) not at all true to (6) exactly true. Homogeneity of the items was high with Cronbach's Alpha being .88. Scale means for T_1 and T_2 were calculated and subsequently difference scores (i.e., post-test – baseline) obtained.

Coping Planning. Coping planning was assessed with two items both at baseline (T_1) and at post-test (T_2). Items read: "I have planned precisely… (1)…in which situations I need to be especially careful so as to succeed in eating sufficient amounts of fruit and

vegetables and (2)… what I can do in difficult situations so as to succeed in eating sufficient amounts of fruits and vegetables." Statements were rated on a 6-point Likert scale ranging from (1) not at all true to (6) exactly true and items correlated significantly with r=.68, p<.01. Scale means for T_1 and T_2 were calculated and subsequently difference scores (i.e., post-test – baseline) obtained.

Fruit and vegetable consumption. Fruit and vegetable consumption was measured at baseline (T_1) and again at the follow-up assessment (t_3). In an open answer format, participants were asked: "How many servings of (a) fruit… and (b) vegetables… …do you eat on average per day?" One serving was defined and visualized as a 'handful'. Sum scores for T_1 and T_2 were calculated and subsequently difference scores (i.e., follow-up – baseline) obtained.

All items were used and validated in previous studies (e.g., Lippke, Ziegelmann, & Schwarzer, 2004; Richert, & Lippke, under review). Means, standard deviations, and intercorrelations are displayed in Table 1.

Table 1. Means (*M*), Standard Deviations (*SD*), and Intercorrelations for changes in action- and coping planning, changes in fruit and vegetable consumption and engagement in *N* = 701 participants

	changes in action planning (T_2)	changes in coping planning (T_2)	changes in fruit and vegetable consumption (T_3)	Engagement (T_1)
M	0.29	0.45	0.64	4.26
SD	1.21	1.26	1.57	0.68
changes in action planning (T_2)	1.00	.51**	.16**	.06*
changes in coping planning (T_2)	.51**	1.00	.15**	.14**
changes in fruit and vegetable consumption (T_3)	.16**	.15**	1.00	.01

Note: ** p < .001, *p < .01.

Analytical procedure

As preliminary analysis, using the INDIRECT macro by Preacher and Hayes (2008), it was tested whether the effect of intervention on changes in fruit and vegetable consumption was mediated via changes in action- and coping planning (multiple mediation).

Subsequently, for each mediator separately, it was tested whether engagement functions as moderator of these mediations, using the MODMED macro (Version 1.1; Model 2) by Preacher et al. (2007). To test the interactions, variables were z-standardized (Aiken & West, 1991). First stage moderated mediation is expressed by an interaction between engagement and intervention on changes in planning (MacKinnon, 2008).

Results

(1) Multiple Mediation

The effect of intervention on changes in fruit and vegetable consumption was fully mediated by changes in action- and coping planning. That is, the direct effect of intervention on changes in fruit and vegetable consumption was reduced from β=.21, p<.001 to β=.16, p>.05 after controlling for changes in action- and coping planning. The contrast test of the specific indirect effects for changes in action planning (β=.03, p<.05) and coping planning (β=.03, p<.05) did not become significant (contrast<.01, p>.05), suggesting that both variables are equally important mediators. The multiple mediator model accounted for 3% of the variance in changes in fruit and vegetable consumption (p < .05).

(2) Moderated Mediation

(a) Engagement moderates the mediation effect of intervention on changes in behavior via changes in action planning. Having received a planning intervention at time 1 emerged as significant predictor of changes in action planning at time 2 (β = .31, p < .001). Changes in action planning in turn predicted changes in fruit and vegetable consumption (β = .15, p < .001). There was a direct effect of the planning intervention on changes in fruit

and vegetable consumption (β = .27, p < .001), indicating that in addition to eliciting changes in action planning, there were other mechanisms via which the intervention led to changes in behavior (see results on multiple mediation above). There was a significant interaction between intervention and engagement on changes in action planning (β = .19, p < .02), indicating that the partial mediation of intervention on changes in fruit and vegetable consumption via changes in action planning was moderated by engagement. Participants needed an engagement value between 3.2 and 5 on the 1-6 point scale to allow for a significant mediation effect. Figure 1 visualizes the results.

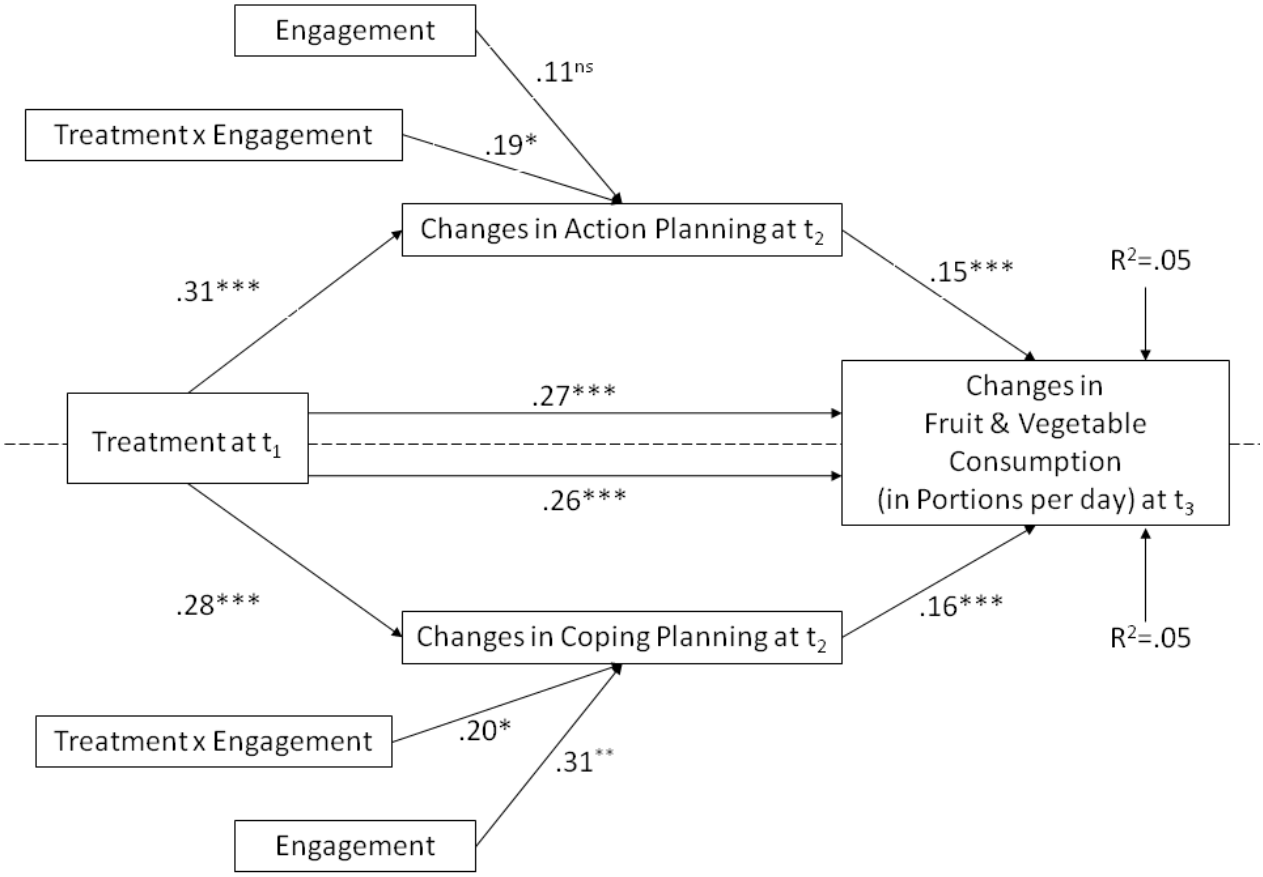

Figure 1. Two Separate Moderated Mediation Models for Changes in Fruit and Vegetable Consumption in N=701 Individuals. The Top Part Shows the Model for Changes in Action Planning as Mediator and the Bottom Part of the Figure Shows the Model for Changes in Coping Planning as Mediator.

Note: *** p < .001, **p < .01, *p < .05

(b) Engagement moderates the mediation effect of intervention on changes in behavior via changes in coping planning. Having received a planning intervention also emerged as significant predictor of changes in coping planning ($\beta = .28$, $p < .001$). Changes in coping planning in turn predicted changes in fruit and vegetable consumption ($\beta = .16$, $p < .001$). There was a significant interaction between intervention and engagement on changes in coping planning ($\beta = .20$, $p < .02$), indicating that the partial mediation of intervention on changes in fruit and vegetable consumption via changes in coping planning is moderated by engagement (see Figure 1). Participants needed an engagement value between 2 and 4.5 on the 1-6 point scale to allow for a significant mediation effect.

To summarize, the effect of intervention on changes in fruit and vegetable consumption was fully mediated by changes in action- and coping planning. Engagement emerged as moderator in the way that intervention led to changes in action- and coping planning, resp., only, if engagement in the intervention was at a certain level.

Discussion

In the present study, we experimentally manipulated action- and coping planning to investigate whether the effect of a health promotion on changes in fruit and vegetable consumption could be accounted for by changes in planning. Moreover, we examined the role of intervention engagement in the process of behavior change to find out if this mediation mechanism varies for different levels of intervention engagement. By this, we aimed at integrating aspects of intervention material as well as intervention implementation. This allows for a more sound evaluation of a behavior change intervention and subsequently provides a possible explanation for ambivalent research findings concerning the role of planning in the behavior change process (cf., Armitage, 2008; Chapman et al., 2009; Gollwitzer, & Sheeran, 2006; Jackson et al., 2005; Kwak et al., 2007; Luszczynska, 2006; Rutter et al., 2006; Skår, et al., in press; Van Osch et al., 2008).

To our knowledge, this is the first study that investigated this research question by looking at causal changes rather than at cross-sectional or non-experimental longitudinal data.

The causal effect of intervention on changes in fruit and vegetable consumption was fully mediated by changes in action- and coping planning. The finding corroborates previous research (Chapman et al., 2009; Gollwitzer & Sheeran, 2006; Wiedemann et al., 2009). The result supports the value of simultaneously addressing action planning and coping planning in health promotion interventions addressing fruit and vegetable consumption.

Moreover, engagement in the intervention emerged as moderator of this effect in the way that intervention led to changes in action- and coping planning, respectively, only, if engagement in the intervention was at a certain level. This model accounted for five percent of the variance in behavior change.

While previous research has suggested a positive linear relationship between engagement and intervention effects (cf., Richert, Lippke, & Ziegelmann, 2010; Ruiter et al., 2006), results of this study point towards a non-linear relationship. In particular, we found that the mediation mechanism only held true for medium levels of engagement. That is, when participants were too little or too highly engaged in the intervention, the effects of intervention on changes in fruit and vegetable consumption were not mediated via changes in action- and coping planning, respectively. This result is in agreement with a study that suggests a similar inverted u-shaped relationship between number of plans and changes in behavior, arguing that cognitive demands increase with the number of plans that are generated in an intervention (Koring, Wiedemann, & Richert, 2009). While making plans is not equal to engagement (cf., Richert, & Lippke, under review), it may be seen as an indicator of effort and explain the mediation mechanism did not work why for high levels of engagement.

Study Limitations

In this study, only one aspect of self-regulation was focused on: namely planning. Research suggests, however, that other social-cognitive variables, such as self-efficacy and

social support, are also useful for behavior change (Araújo-Soares, McIntyre, & Sniehotta, 2009). This is in line with our finding that only five percent of the variance in changes in fruit and vegetable consumption were explained. Subsequently, models accounting for additional predictors may be favored in future research.

Although recruitment was broad so as to obtain a representative sample, study participants were predominantly highly educated women. In addition, the time frame that individuals chose for reaching their goal (and inevitably the time of their follow-up dates) might be confounded with the difficulty of their goal. As only those participants, who had already completed all three assessments (cf., procedure), were included in this study, this may have resulted in a sample that is skewed towards easier goals. Also, for reasons of the study design, it was not possible to test for systematic dropout of participants over time. At any rate, the results of this study should be generalized with caution and replication should be attempted in a more representative sample.

Fruit and vegetable consumption was assessed with a self-report measure, because the online format of the study did not allow for more objective measures. However, intentional misreporting (owing, e.g., to social desirability) was limited due to the anonymity provided by the Internet. Unintentional over- and underestimation due to memory and recall skill limitations was unlikely as the sample was not mentally impaired and misreports due to a misunderstanding of the term “portion” was held low by providing a clear definition along with pictures portraying a *handful*. In addition, systematic misreports still allow for analyses of changes in behavior.

Analyzing difference scores was the method of choice in this study as we were interested in predicting *change* rather than discrete follow-up scores. Our causal assumption was that a health intervention would cause changes in cognitions and that those changes in cognitions would subsequently cause changes in behavior. Commonly, alternative measures, such as baseline controlled follow-up scores or residualized change scores are favored. However, Willet (1997) has pointed out that under ordinary circumstances, reliability of difference scores can actually be greater than the reliabilities of the individual baseline and follow-up assessments.

Outlook

Integrating aspects of both intervention material and intervention implementation seems to be fruitful in obtaining a more sound evaluation of intervention effects. In the future, this approach should be continued and extended. Additionally, questions concerning the nature of intervention engagement should be addressed, e.g., 'Of which nature is the relationship between engagement and intervention effects (linear vs. non-linear)?', 'What predicts engagement?' and 'How can we manipulate engagement so as to optimize intervention effects?'. The more we know about the process of behavior change the better are our chances to develop highly effective behavior change interventions and subsequently to reduce behavior-induced morbidity and mortality.

References

Aiken, L. S., & West, S. G. (1991). Multiple Regression: Testing and interpreting interactions. Newbury Park, CA: Sage.

Araújo-Soares, V., McIntyre, T., & Sniehotta, F. F. (2009). Predicting changes in physical activity amongst adolescents: The role of self-efficacy, intention, action planning and coping planning. Health Education Research, 24 (1), 128-139.

Armitage, C. J. (2008). A volitional help sheet to encourage smoking cessation: A randomized exploratory trial. Health Psychology, 27, 557-566.

Bartholomew, L. K., Parcel, G. S., Kok, G., & Gottlieb, N. H. (2006). Planning Health Promotion Programs. An Intervention Mapping Approach. San Francisco: Jossey-Bass.

Celio, A. A., Winzelberg, A. J., Dev, P., & Taylor, C. B. (2002). Improving compliance in on-line, structured self-help programs: Evaluation of an eating disorder prevention program. Journal of Psychiatric Practice, 8(1), 14-20.

Chapman, J., Armitage, C. J., & Norman, P. (2009). Comparing implementation intention interventions in relation to young adults' intake of fruit and vegetables.Psychology & Health, 24, 317-332.

Crutzen, R. (2010). When is dissemination of an internet-delivered intervention successful? In E. Olsen (Ed.), Behaviour Change & Technology. Oslo: Changetech.

Dombrowski, S. U., Sniehotta, F. F., Avenell, A., & Coyne, J. C. (2007). Current issues and future directions in psychology and health: Towards a cumulative science of behavior change: Do current conduct and reporting of behavioural interventions fall short of best practice? Psychology and Health, 22(8), 869–874.

Hardeman, W., Michie, S., Fanshawe, T., Prevost, A. T., McLoughlin, K., & Kinmonth, A. L. (2008). Fidelity of delivery of a physical activity intervention: Predictors and consequences. Psychology and Health, 23(1), 11-24.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects on simple plans. American Psychologist, 54, 493-503.

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. Advances in Experimental SocialPsychology, 38, 70-110.

Jackson, C., Lawton, R., Knapp, P. R., Raynor, D. K., Conner, M. T., Lowe, C. J., & Closs, S. J. (2005). Beyond intention: Do specific plans increase health behaviours in patients in primary care? A study of fruit and vegetable consumption. Social Science & Medicine, 60, 2383-2391.

Kellar, I ., & Abraham, C. (2005). Randomised controlled trial of a brief research-based intervention promoting fruit and vegetable consumption. British Journal of HealthPsychology, 10, 543-558.

Koring, M., Wiedemann, A., & Richert, J.(2009). Who benefits from planning? Moderators of intervention effectiveness. 23rd Annual Conference of the European Health Psychology Society. Poster session presented at the 23rd conference of the European Health Psychology Society, Pisa, Italy, 24th September 2009

Kwak, L., Kremers, S. P. J., van Baak, M. A., & Brug, J. (2007). A poster-based intervention to promote stair use in blue- and white collar worksites. Preventive Medicine, 45, 177-181.

Lippke, S., Ziegelmann, J. P., & Schwarzer, R. (2004). Initiation and maintenance of

physical exercise: Stage-specific effects of a planning intervention. Research in Sports Medicine, 12, 221-240.

Luszczynska, A. (2006). An implementation intention intervention, the use of a planning strategy, and physical activity after myocardial infarction. Social Science and Medicine, 62, 900-908.

MacKinnon, D. P. (2008). Introduction to Statistical Mediation Analysis. New York: LEA.

Manwaring, J. L., Bryson, S. W., Goldschmidt, A. B., Winzelberg, A. J., Luce, K. H., Cunning, D., et al. (2008). Do adherence variables predict outcome in an online program for the prevention of eating disorders? Journal of Consulting and Clinical Psychology, 76(2), 341-346.

Michie, S., Dormandy, E., & Marteau, T. M. (2004). Increasing screening uptake amongst those intending to be screened: The use of action plans. Patient Education and Counseling, 55, 218-222.

Michie, S., Johnston, M., Francis, J., Hardeman, W., & Eccles, M. (2008). From theory to intervention: Mapping theoretically derived behavioural determinants to behavior change techniques. Applied Psychology: An International Review, 57(4), 660-680.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods, 40, 879-891

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. Multivariate Behavioral Research, 42,185-227.

Richert, J., & Lippke, S. (under review). Validation of a scale measuring Intervention-engagement, Journal of Educational Psychology.

Richert, J., Lippke, S., & Ziegelmann, J. P. (in press). Intervention-engagement and its role in the effectiveness of stage-matched interventions promoting physical exercise. Research in Sports Medicine: An International Journal.

Ruiter, R. A. C., Kessels, L. T. E., Jansma, B. M., & Brug, J. (2006). Increased attention for computer-tailored health communications: An event-related potential study.

Health Psychology, 25(3), 300-306.

Rutter, D. R., Steadman, L., & Quine, L. (2006). An implementaion intention intervention to increase uptake of mammography. Annals of Behavioral Medicine, 32, 127-134.

Schwarzer, R. (1992). Self-efficacy in the adoption and maintenance of health behaviors: Theoretical approaches and a new model. In R. Schwarzer (Ed.), Self-efficacy: Thought control of action (pp. 217-243). Washington, DC: Hemisphere.

Schwarzer, R. (2008). Modeling health behavior change: How to predict and modify the adoption and maintenance of health behaviors. Applied Psychology: An International Review, 57(1), 1-29.

Skår, S., Sniehotta, F. F., Molloy, G. J., Prestwich, A., Araújo-Soares, V. (in press). Do brief online planning interventions increase physical activity amonst university students? A randomized controlled trial. Psychology and Health.

Sniehotta, F. F. (2009). Towards a theory of intentional behavior change: Plans, planning, and self-regulation. British Journal of Health Psychology, 14, 261-273.

Sniehotta, F. F, Scholz, U., & Schwarzer R. (2006). Action plans and coping plans for physical exercise: A longitudinal intervention study in cardiac-rehabilitation. British Journal of Health Psychology, 11, 23-37.

Strecher, V. J., McClure, J., Alexander, G., Chakraborty, B., Nair, V., Konkel, J., Greene, S., Couper, M., Carlier, C., Wiese, C., Little, R., Pomerleau, C., Pomerleau, O.(2008). The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. Journal of Medical Internet Research, 10(5), 23-32.

Van Osch, L., Lechner, L., Reubsaet, A., Wigger, S., & de Vries, H. (2008). Relapse prevention in a national smoking cessation contest: Effects of coping planning. British Journal of Health Psychology, 13, 525-535.

Wiedemann, A. U., Schüz, B., Sniehotta, F. F., Scholz, U., & Schwarzer, R. (2009).Disentangling the relation between intentions, planning, and behavior: A moderated mediation analysis. Psychology & Health, 24, 67-79.

Willett, J. B. (1997). Measuring change: What individual growth modeling buys you. In E.

Amsel, & K. A. Reninger (Eds.), Change and development (pp. 213-243). Mahwah, NJ: Lawrence Erlbaum.

World Health Organization. (2002). The World Health Report, Reducing risks, promoting health, Geneva.

6

General Discussion

General Discussion

In the following general discussion, the central research questions of this thesis are laid out and embedded in a rationale, thereby highlighting this dissertation's contribution to health psychology. Each chapter is briefly reiterated and reflected individually. Following, the main findings of each chapter are integrated and discussed in terms of implications for future research and practice. Table 1, which is located at the end of the chapter, provides an overview of the aims, findings and conclusions and may function as a guide and reference to this discussion.

Understanding what drives health behavior change

Lifestyle related risk behaviors, such as a poor diet and physical inactivity are the leading cause of morbidity and mortality in the developed world (Khaw et al., 2008). Understanding what drives *health behavior change* is therefore crucial (Leventhal, Weinman, Leventhal, & Phillips, 2008). *Theories of behavior and behavior change* provide insight on the various constructs (e.g., social cognitions) that are of importance in the process of behavior change. These theories also explain the mechanisms (e.g., mediation and moderation) via which the factors exert their influence on behavior and behavior change, respectively. Thus, they can and should inform the *content of an intervention* (cf., Michie, Johnston, Francis, Hardeman, & Eccles, 2008). A popular (Brug, Conner, Harré, Kremers, McKellar, & Whitelaw, 2005) type of theories of behavior change is subsumed under the term *stage theories*.

Stage theories of health behavior change assume that individuals differ in their mindsets (manifested in social cognitions) towards behavior change. Those who have maximally similar cognitions are said to be in the same *stage*. It is suggested that predictors driving transitions towards behavior change are differential in each stage. It follows then that interventions need to be stage appropriate, addressing only those factors that are relevant for a particular stage and mindset (Sutton, 2000; Weinstein, Rothman, & Sutton, 1998).

Validity of stage allocation

To ensure the receipt of appropriate and consequently effective health promotion interventions, stages need to adequately represent individuals' mindsets towards behavior change. In other words, stage *allocation* needs to be valid and reliable (Lippke, Ziegelmann, Schwarzer, & Velicer, 2009). The first part of this thesis (*Chapter 2* and *Chapter 3*) deals with this question of the validity of stage allocation. Each chapter addresses the topic from a different angle, focusing on unique aspects. In the following, the chapters will be described individually and their findings subsequently integrated.

Criteria in stage assessments

Measures used to allocate individuals into stages often rely on self-reports about intentions (to change) and behavior (for an example measure cf., Lippke, Ziegelmann, Schwarzer, & Velicer, 2009). Frequently, so as to allow for grouping of individuals, a standard for behavior and intention that individuals can be judged against (such as the guidelines on physical activity advocated by the World Health Organization) is embedded in the stage measure (Richert, Lippke, & Schwarzer, 2010).

Chapter 2 therefore takes on a question that has been overlooked in research thus far, namely whether the standards or criteria for behaviors that are embedded in stage measures impact stage allocation. If they do, the resulting question then is: which criterion yields the best reflection of individuals' mindsets? In *Chapter 2*, a commonly used external standard was compared with a novel alternative: individuals' subjective evaluation of their behavior and intention to change. Results indicate that the different criteria indeed affected stage-allocation. Staging based on an individual criterion seemed to reflect individuals' mindsets better than staging based on the external standard of 30 minutes of physical activity on at least 4 days of the week. The finding suggests that misclassification (defined as an inaccurate reflection of individuals' mindsets) is likely if an arbitrary standard that individuals are measured against is employed and that it may be more beneficial to rely on individuals' subjective evaluation when it comes to assessing their stage. However, in this

chapter, an accurate reflection of individuals' mindsets was defined as high correspondence between assigned stages and individuals' behavior and intention to change rather than all their social cognitions. This was grounded in the implicit notion that individuals' behavior and intentions are a good representation of the social cognitions that amount to a mindset (cf., Lippke, & Plotnikoff, 2006). In *Chapter 3,* this belief is not taken at face value, but rather tested empirically. The research thereby builds on and progresses from the research presented in *Chapter 2.*

Mindsets and Stages based on behavior and intention

Stage theories differ in their decision on how many mindsets/ stages exist (Warner, & Lippke, 2008). While Heckhausen (1991), for example, only differentiates between a deliberative and an implemental mindset, Weinstein's Precaution Adoption Process Model (PAPM) assumes no less than six different stages (Weinstein, 1988). Rather than selecting one theory and relying on the accuracy of the 'a priory' specified stages (*theory driven*), in *Chapter 3*, mindsets (i.e., distinct social-cognitive profiles) are extracted directly from the data using latent class analysis (LCA) (*empiricism driven*). This data-based approach avoids theoretical and methodological issues inherent in confirmatory approaches and subsequently allows for a more accurate consideration of mindsets. Thus, it extends and improves on the research presented in *Chapter 2*. The LCA identified four distinct mindsets in the data. Stages as assigned by a common stage-algorithm based on self-reports about intentions and behavior relying on the subjective criterion (cf., *Chapter 2*) were then pinned against the latent class based stages. This was done to determine how well the stages derived from the algorithm match the in the data identified mindsets. Correspondence was low, suggesting that the stage-algorithm might not be effective in assigning individuals to stages that reflect their mindsets towards behavior change accurately. In other words, misclassification is likely when a stage-algorithm based on reports about intentions and behavior is used.

Integrated conclusions and implications of Chapter 2 and 3

The findings of both *Chapters 2* and *3* suggest that common and widely used algorithms employed to assign individuals to stages contain the risk of misclassification. That is, these measures lead to an allocation to stages that might not accurately reflect individuals' mindsets, thus undermining the possibility of stage appropriate interventions. Both studies thereby contribute to the advancement of health psychology by offering an explanation for the lack of supportive results in the field of stage based health promotion (Adams & White, 2005; Bridle et al., 2005; Conn, Hafdahl, Brown & Brown, 2008; Noar, Benac & Harris, 2007) that does not dispute the notion of stage theories. Rather, this research raises awareness for the shortcomings of the measures that are necessary to test and apply these theories.

In *Chapter 2,* it is suggested that external standards are obsolete in stage assessment. Does it follow that they are irrelevant altogether? External standards are useful when they serve as recommendations, that is, as objectives that individuals can strive for. Standards may also be used to assess how individuals measure up, that is, how far from or close to a recommended goal they are. Such information when fed back may operate as motivator for an individual. When employed as cut-off criteria in stage assessments, however, such standards become a rigid tool that disregards individuals' perceptions.

The findings of *Chapter 3* suggest that diagnosing a stage based on behavior and intention to change, even when relying on subjective evaluations rather than external criteria, does not lead to a stage assignment that is reflective of individuals' mindsets. Again, this undermines the possibility of stage appropriate and consequently effective health promotion interventions. This seems to suggest that an assessment of all social cognitions comprising a mindset towards behavior change is necessary when assigning a stage. But is this feasible?

Wherever technology permits highly individualized tailoring, as is possible with computer-based interventions for example (Krebs, Prochaska, & Rossi, 2010), one idea that may be explored further is *module based interventions*. Rather than presenting a

limited range of interventions addressing a set of predictors as is done in stage based interventions, treatment components addressing individual social cognitions could be combined in every way imaginable to create a highly individualized and effective intervention. This concept is adopted in computer-tailoring (Kroeze, Werkman, & Brug, 2006).

In some cases, however, tailored interventions are not viable, for example because the target population is technology illiterate. In others, it is not desired, because it compromises standardization of interventions and impedes treatment evaluation for example. There, an assessment of all social cognitions assumed to make up a mindset towards behavior change is not practical. Grouping of individuals, as is done in a stage assessment, needs to be accurate and fast. Future research should aim at finding practical solutions for researchers and practitioners in the field of health psychology. This was beyond the scope of this dissertation.

The focus of the first part of this dissertation lies within the research area of *what drives behavior change* and informs the *content of an intervention*. The focus of the second part of this dissertation (*Chapter 4*) goes beyond this matter and addresses an issue that arises *after* the content of an intervention has been developed. Namely, how the intervention is *received* by the target population. In the following, it is laid out why this topic is addressed and how the research presented here contributes to the advancement of the field.

Investigating the receiving end of an intervention

While theories of behavior and behavior change inform the *content of an intervention* (cf., Michie, Johnston, Francis, Hardeman, & Eccles, 2008), they seldom specify the change techniques that are effective in modifying the proposed determinants in an intervention addressing behavior change (Ashford, French, Sniehotta, Bishop, & Michie, 2009). Nor do they address the questions of mode of delivery (e.g., puppets, video, brochures), delivery fidelity (e.g., compliance with protocol, cf., Hardeman, Michie,

Fanshawe, Prevost, McLoughlin, & Kinmonth, 2008), successful dissemination among the target population (cf., Crutzen, 2010) or intervention-engagement (Richert, Lippke, & Ziegelmann, in press).

As these factors affect the final *form and success of an intervention*, it is astounding how little attention they receive in health psychology research. Researchers often restrict their treatment evaluations to the predictors specified in theories of behavior change and conclude that depending on the outcome of the intervention study, they either addressed the right or the wrong determinants of behavior change (Richert, & Lippke, 2010). Looking at factors concerning the form of the intervention (e.g., change techniques, modes of delivery etc.) as well as if and to which degree the intervention is received by the target population, would allow for a more comprehensive and sounder evaluation of health promotion. Therefore, the research of the second part of this dissertation (*Chapter 4*) targets the receiving end of behavior change intervention.

Intervention Engagement

In particular, *Chapter 4* deals with the construct of *engagement* in the context of health behavior change interventions. Specifically, it demonstrates the shortcomings of previous research focused on the construct: absence of consistent terminology and established definition as well as lack of comprehensive and valid measures. Following, a comprehensive definition of the construct is proposed and a new self-report measure, the Task Engagement Scale (TES), suggested in accordance with the definition of the construct. In the presented study, the factorial structure of the scale is examined and data on its relationship both with objective indicators of engagement and with theoretically related constructs provided. The items of the TES loaded on four subscales: task-compliance, effort, undivided attention and absorption, which in turn loaded on the higher order factor engagement. The TES captured more information than the objective indicators time-on-task and the completion rate of intervention materials. Associations between theoretically related constructs and the TES score indicate that engagement is a sufficiently

distinct construct.

Beyond presenting a new measure of engagement that may be used to assess participants' engagement in a behavior change intervention, the novelty of this chapter and its main contribution to the research field lies in the recognition that it is not sufficient to restrict intervention research to theories of behavior and behavior change but that it is worthwhile to broaden our research to elements affecting intervention success beyond its theory-based content.

Generating an integrated understanding on what leads to health behavior change

The third and final part of this dissertation (*Chapter 5*) aims at *integrating* the two approaches to health intervention research: the focus on what informs intervention content (that is, determinants and mechanisms of behavior change) and the focus on the degree to which an intervention is received by the target population. Thus, it builds on and extends the research presented in *Chapters 2, 3* and *4* and helps generate an integrated understanding on what leads to health behavior change.

In particular, *Chapter 5* examines the role of intervention engagement in the behavior change process. The mechanism via which the intervention exerted its influence on changes in fruit and vegetable consumption was through changes in planning cognitions. This mediation was moderated by participants' engagement in the way that the treatment led to changes in cognitions only when participants' engagement in the treatment was at a moderate level. Furthering the notion presented in *Chapter 4*, this result demonstrates that a theory- and evidence based intervention does not invariably lead to changes in the cognitions it targets. The implication is that when developing health promotion interventions, researchers and practitioners have to give consideration to factors that affect participants' engagement in the intervention.

Integrated outlook

One such factor that might affect participants' intervention engagement favorably is *tailoring.* In *Chapters 2* and *3*, it is recognized that stage appropriate interventions are tailored, i.e., they address only those factors that are relevant for a particular stage and mindset (Sutton, 2000; Weinstein, Rothman, & Sutton, 1998). Research suggests that tailoring increases the effectiveness of an intervention due to greater *personal relevance* (Kreuter & Olevitch, 2000). Personally relevant information is suggested to receive more attention (e.g., Ruiter, Kessels, Jansma, & Brug, 2006) and even be processed in more depth (cf., Elaboration Likelihood Model, Petty, & Cacioppo, 1986).

Engagement (cf., *Chapter 4*) is a construct that captures this increased focus and attention. It is therefore plausible to assume that stage tailoring will be reflected in higher intervention-engagement. Research findings are ambivalent, however. While Richert, Lippke, & Ziegelmann (in press) could demonstrate that engagement increased the effectiveness of an intervention (this result is in line with findings presented in *Chapter 5*), they could not find evidence for the notion that stage appropriateness (tailoring) of an intervention affected engagement. A probable explanation may be found in the *form* of tailored interventions. This will be explained in the following.

Personalization refers to the method of *implying personal relevance* by (a) addressing participants by name, (b) customization (i.e. "This message was designed particularly for you") and (c) presenting information in a meaningful context to the participant (cf. Kreuter, & Olevitch, 2000). So, regardless of the actual relevance of materials, participants perceive information to be relevant for them. Similar to a placebo effect, this perception may be more significant in the behavior change process than the actual relevance of the information (Webb, Simmons, & Brandon, 2005). Taking this information into consideration, it may be understood why despite the lack of tailoring (as is the case in generic 'one size fits all' interventions), treatments are found to be effective (perhaps because the treatment was personalized and subsequently perceived as personally relevant), and why some interventions despite tailoring are not (cf., Brug Conner, Harré, Kremers, McKellar, & Whitelaw, 2005, Noar, Benac, & Harris, 2007).

Future research might address this question. In particular, a randomized controlled trial in which the effects of four types of interventions are compared may be fruitful. The interventions should contain either (1) generic information, (2) tailored information, (3) personalized generic information, or (4) personalized tailored information. Perceived personal relevance and engagement should be investigated as successive mediators.

Furthermore, this general discussion drew attention to elements that also shape the *form of an intervention*, such as the mode of delivery (e.g., puppets, video, and brochures). The form of a treatment likely influences intervention-engagement and subsequently affects the outcome of an intervention. Future research should further address this area of intervention research and build on the research presented in this dissertation. Moreover, the development of a meta-model on how interventions work might be useful in generating an integrated and comprehensive understanding of how to modify health risk behaviors.

Shortcomings of the studies presented in this dissertation limit the implications of the findings to some degree (cf., *Chapters 1-4*). Study findings need to be replicated in different settings with different health behaviors and with stronger experimental tests. Possible directions for future research have been pointed out.

However, while many questions are left unanswered, this thesis has made a contribution to the advancement of the field. This dissertation recognized and addressed two issues in health psychology and offered a response. First, it was suggested that the lack of supportive results in the field of stage based health promotion research may be accounted for by the risk of misclassification inherent in algorithms employed to assign individuals to stages. It thereby provides an alternative to interpretations that challenge the very notion of stage theories. Secondly, this dissertation recognized an area of intervention research that had this far been neglected in health psychology, the receiving end of a treatment: intervention-engagement. This thesis not only provided a comprehensive definition and measure of the construct, it also provided an example on how to integrate this aspect into the examination of the behavior change process. To conclude, this

dissertation was a first step in gaining a more comprehensive understanding of what leads to behavior change.

References

Adams, J., & White, M. (2005). Why don't stage-based activity promotion interventions work? *Health Education Research*, *20*(2), 237-243. doi:10.1093/her/cyg105

Ashford, S., French, D. P., Sniehotta, F. F., Bishop, A., & Michie, S. (2009, 9th - 11th September 2009). *The refinement of a taxonomy of techniques to change beahviour.* Paper presented at the 12th Annual Conference of the BPS Division of Health Psychology conference, Aston University.

Bridle, C., Riemsma, R. P., Pattenden, J., Sowden, A. J., Mather, L., Watt, I. S., et al. (2005). Systematic review of the effectiveness of health behavior interventions based on the Transtheoretical model. *Psychology & Health*, *20*(3), 283-301. doi: 10.1080/08870440512331333997

Brug, J., Conner, M., Harré, N., Kremers, S., McKellar, S., & Whitelaw, S. (2005). The Transtheoretical Model and stages of change: A critique: observations by five commentators on the paper by Adams, J. and White, M. (2004). Why don't stage-based activity promotion interventions work? *Health Education Research, 20*(2), 244-258. doi:10.1093/her/cyg105

Conn, V., Hafdahl, A., Brown, S., & Brown, L. (2008). Meta-analysis of patient education interventions to increase physical activity among chronically ill adults. *Patient Education and Counseling*, *70*, 157-172. doi:10.1016/j.pec.2007.10.004

Crutzen, R. (2010). When is dissemination of an internet-delivered intervention successful? In E. Olsen (Ed.), *Behaviour Change & Technology*. Oslo: Changetech.

Hardeman, W., Michie, S., Fanshawe, T., Prevost, A. T., McLoughlin, K., & Kinmonth, A. L. (2008). Fidelity of delivery of a physical activity intervention: Predictors and consequences. *Psychology and Health, 23(1),* 11-24. doi: 10.1080/08870440701615948

Heckhausen, H. (1991). *Motivation and action*. Berlin, Heidelberg, New York: Springer.

Khaw, K. T., Wareham, N., Bingham, S., Welch, A., Luben, R., & Day, N. (2008). Combined impact of health behaviours and mortality in men and women: the EPIC-Norfolk prospective population study. *PLoS Med, 5*(1), e12. doi:10.1371/journal.pmed.0050012

Krebs P., Prochaska, J. O., & Rossi, J.S. (2010). A meta-analysis of computer-tailored interventions for health behavior change. *Preventive Medicine, 51,* 214-221. doi:10.1016/j.ypmed.2010.06.004

Kreuter, M., Farrell, D.W., & Olevitch, L.R. (2000). *Tailoring health messages: customizing communication with computer technology.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc**.**

Kroeze, W., Werkman, A., & Brug, J. (2006). A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. *Annals of Behavioral Medicine*, 31, 205-223. doi: 10.1207/s15324796abm3103_2

Leventhal, H., Weinman, J., Leventhal, E. A., & Phillips, L. A. (2008). Health psychology: The search for pathways between behavior and health. *Annual Review of Psychology, 59*, 477-505. doi:10.1146/annurev.psych.59.103006.093643

Lippke, S., & Plotnikoff, R. C. (2006). Stages of change in physical exercise: A test of stage discrimination and non-linearity. *American Journal of Health Behavior, 30*(3), 290-301.

Lippke, S., Ziegelmann, J. P., Schwarzer, R., & Velicer, W. F. (2009). Validity of stage assessment in the adoption and maintenance of physical activity and fruit and vegetable consumption. *Health Psychology, 28*(2), 183-193. doi: 10.1037/a0012983

Michie, S., Johnston, M., Francis, J., Hardeman, W., & Eccles, M. (2008). From Theory to Intervention: Mapping Theoretically Derived Behavioural Determinants to Behaviour Change Techniques. *Applied Psychology: An International Review, 57*(4), 660-680. doi:10.1111/j.1464-0597.2008.00341.x

Noar, S. M., Benac, C., & Harris, M. (2007). Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychological Bulletin*, *133*(4), 673-693. doi:10.1037/0033-2909.133.4.673

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

Richert, J., & Lippke, S. (2010). Assessment and Validation of the Task-Engagement Scale (TES). Manuscript submitted for publication.

Richert, J., Lippke, S., & Schwarzer, R. (2010). Comparison of individual criteria and externally imposed criteria for stage allocation: Findings from an internet study addressing physical activity, *Measurement in Physical Education and Exercise Science, 14,* 1–16. doi: 10.1080/1091367X.2010.520242

Richert, J., Lippke, S., & Ziegelmann, J. P. (in press). Intervention-engagement and its role in the effectiveness of stage-matched interventions promoting physical exercise. *Research in Sports Medicine: An International Journal.*

Ruiter, R. A. C., Kessels, L. T. T., Jansma, B. M. & Brug, J. (2006). Increased Attention for Computer-Tailored Health Communications: An Event-Related Potential Study. Health Psychology, 25 (3), 300-306. doi:10.1037/0278-6133.25.3.300

Sutton, S. (2000). Interpreting cross-sectional data on stages of change. *Psychology & Health, 15*(2), 163-171. doi: 10.1080/08870440008400298

Warner, L. M., & Lippke, S. (2008). Psychological stage models of physical exercise – Research Advances. In M. P. Simmons & L. A. Foster (Ed.), *Sport and exercise psychology - research advances* (pp. 19-51). Hauppauge, NY: Nova Science Publishers.

Webb, M. S., Simmons, V. N. & Brandon, T. H. (2005). Tailored Interventions for Motivating Smoking Cessation: Using Placebo Tailoring to Examine the Influence of Expectancies and Personalization. Health Psychology, 24 (2), 179-188. doi:10.1037/0278-6133.24.2.179

Weinstein, N. D. (1988). The precaution adoption process. *Health Psychology, 7*(4), 355-386. doi:10.1037/0278-6133.7.4.355

Weinstein, N. D., Rothman, A. J., & Sutton, S. R. (1998). Stage theories of health behavior: Conceptual and methodological issues. *Health Psychology, 17*(3), 290-299. doi:10.1037/0278-6133.17.3.290

Table 1. Summary of dissertation aims, findings, and conclusions.

Aims	Findings	Conclusions
Chapter 2 deals with the question of the validity of stage allocation. The aim was to demonstrate the impact of different classification criteria on stage-allocation, and to test which criterion (individual vs. externally imposed) yields the best reflection of individuals' mindsets towards behavior change (manifested in intentions to change and behavior).	As predicted, the different criteria affected stage-allocation, and staging based on the individual criterion seemed to reflect individuals' intention to change and behavior better than staging based on the external standard of 30 minutes of physical activity on at least 4 days of the week.	To ensure the receipt of appropriate and consequently effective health promotion interventions, stages need to adequately represent individuals' mindsets towards behavior change. In other words, stage measures need to be valid and reliable. As the standards, which are embedded in stage measures or criteria that individuals are measured against have an impact on stage-allocation, the risk of misclassification is given if an arbitrary standard is used. Relying on an individual criterion seems to lead to stage-allocation that corresponds better with individuals' mindsets towards behavior change (manifested in intentions and behavior) than stage-allocation based on an external standard.
Chapter 3 extends the question of the validity of stage allocation. The aim was to first identify stages bases on underlying mindsets (i.e., based on distinct social-cognitive profiles extracted from the data) and to then examine the concurrent validity of this latent-class based approach and a commonly used stage-algorithm based on intentions to change and behavior.	The LCA identified four distinct mindsets – *not intending to change, intending to change (no action), intending to change (with action), and maintaining behavior.* Convergent validity with a stage-algorithm based on self-reports about intentions to change and behavior was low.	Stages as assigned by the stage-algorithm did not correspond well with the extracted mindsets (as manifested in social cognitions). This indicates that commonly used stage-algorithms based on self-reports about intentions and behavior might not be effective in assigning individuals to stages that represent mindsets towards behavior change, undermining the possibility for stage-matched interventions.
Chapter 4 deals with the construct of engagement in the context of health behavior change interventions. The aim was to propose a new, comprehensive measure, the Task Engagement Scale (TES) and confirm the hypothesized factorial structure of the scale. Concurrent and discriminant validity was tested by examining its relationship with objective indicators of engagement as well as with related constructs.	The items of the TES loaded on four subscales: *task-compliance*, *effort*, *undivided attention* and *absorption*, which in turn loaded on the higher order factor *engagement*. The TES captured more information than the objective indicators time-on-task and the completion rate of intervention materials. Associations between theoretically related constructs and the TES score indicate that engagement is sufficiently distinct.	With the existence of a valid self-report measure of all main components of engagement, researchers and practitioners are able to assess to which extend participants engage with intervention materials and can subsequently judge interventions' effectiveness (or lack thereof) more adequately.
Chapter 5 examines the role of engagement in the specific context of a planning intervention targeting fruit and vegetable consumption. In particular, the aim was to test the moderator effect of engagement on the treatment - changes in planning cognitions - changes in behavior mediation chain.	The effect of intervention condition (T1) on changes in fruit & vegetable consumption (T3) was fully mediated by changes in action- and coping planning cognitions (T2). Engagement emerged as moderator in the way that intervention led to changes in planning, when engagement was at a moderate level.	A theory- and evidence based intervention does not invariably lead to changes in the cognitions that it targets. Study findings suggest that only when participants engage in the intervention (that is, give their full attention to the material; put effort into completing given tasks etc.), will it lead to changes in behavior via changes in cognitions. This result demonstrates that when developing interventions, researchers and practitioners should also give consideration to factors that affect participants' engagement in an intervention.

Curriculum Vitae

For reasons of data protection,
the curriculum vitae is not included in the online version

Der Lebenslauf ist in der
Online-Version aus Gründen des Datenschutzes nicht enthalten

List of publications

Manuscript submitted for publication

Koring, M., Richert, J., Lippke, S., Parschau, L., Reuter, T., & Schwarzer, R. (submitted). Synergistic Effects of Planning and Self-Efficacy on Physical Activity, Manuscript submitted for publication

Koring, M., Richert, J., Parschau, L., Ernsting, A., Lippke, S. & Schwarzer, R. (submitted). A Combined Planning and Self-Efficacy Intervention to Promote Physical Activity: Effectiveness and Working Mechanisms?, Manuscript submitted for publication

Parschau, L., Richert, J., Koring, M., Lippke, S., Schwarzer, R. (submitted). Stages of Change in Physical Activity are Associated with Changes in Self-Efficacy, Outcome Expectancies, and Planning, Manuscript submitted for publication

Lippke, S., Ernsting, A., Richert, J., Parschau, L., & Schwarzer, R. (submitted). Nicht-lineare Zusammenhänge zwischen Intention und Handeln: Eine Längsschnittstudie zu körperlicher Aktivität und sozial-kognitiven Prädiktoren [Non-linear relationships between intention and action: A longitudinal study on physical activity and social-cognitive predictors], Manuscript submitted for publication

Richert, J., Lippke, S., & Schwarzer, R. (submitted). Assessment and Validation of the Task-Engagement Scale (TES), Manuscript submitted for publication

Richert, J., Lippke, S., Lange, D., Parschau, L., & Schwarzer, R. (submitted). Intervention engagement in behavior change – how to plan fruit and vegetable consumption, Manuscript submitted for publication

Richert, J., Schüz, N., Lippke, S., & Schüz, B. (submitted). *Stages of Health Behavior Change and Mindsets: A Latent Class Appraoch*, Manuscript submitted for publication.

In press

Richert, J., Lippke, S., & Ziegelmann, J. P. (in press). Intervention-engagement and its role in the effectiveness of stage-matched interventions promoting physical exercise. *Research in Sports Medicine: An International Journal*

2010

Richert, J., Lippke, S., & Schwarzer, R. (2010). Comparison of Individual Criteria and Externally Imposed Criteria for Stage Allocation: Findings from an Internet Study Addressing Physical Activity, *Measurement in Physical Education and Exercise Science*, 14: 1–16.

Richert, J., Reuter, T., Wiedemann, A., Lippke, S., Ziegelmann, J.P., & Schwarzer, R. (2010). Differential Effects of Planning and Self-Efficacy on Fruit and Vegetable Consumption, *Appetite, 54(3),* 611-614.

Schwarzer, R., Richert, J., Kreausukon, P., Remme, L., Wiedemann, A.U., & Reuter, T. (2010). Translation of Intentions into Dietary Behaviors via Planning Requires Self-Efficacy*., International Journal of Psychology, 45(4),* 260-268.

2009

Schwarzer, R., & Richert, J. (2009). Health psychology at the International Congress of Psychology in Berlin, Germany. *Zeitschrift für Gesundheitspsychology, 17, 40-42.*

2008

Richert, J., & Lippke, S. (2008). Theoretical approaches (Health knowledge, attitudes, and practices - theoretical approaches for evidence-based practice. In R. W. Pierce & R. I. Schwartz (Eds.), *New Perspectives on Knowledge, Attitudes and Practices in Health* (pp. 133-156). Hauppauge, NY: Nova Science Publishers.

2007

Richert, J., Lippke, S., Ziegelmann, J. P., & Petermann, F. (2007). Compliancesteigerung durch stadienpassende Interventionen [Increasing compliance with stage-matched interventions]. In F. Petermann & I. Ehlebracht-König (Eds.), *Motivierung, Krankheitsbewältigung und Compliance* (pp. 135-149). Regensburg, Germany: Roderer Verlag.

Published Abstracts

2010

Koring, M., Richert, J.M., Reuter, T., Wiedemann, A. U., & Schwarzer, R. (2010). *Synergistic effects of planning and self-efficacy on physical activity.* Paper presented at the 24th conference of the European Health Psychology Society, Cluj, Romania, 1st-4th September 2010.

Lange, D., & Richert, J. (2010). *Gender differences in the relationships between intentions, planning, and health behavior: A moderated mediation analysis.* Paper presented at the 24th conference of the European Health Psychology Society, Cluj, Romania, 1st-4th September 2010.

Lippke, S., Schwarzer, R., Richert, J., & Ziegelmann, J.P. (2010). *Selbstwirksamkeitserwartung als vermittelnder Faktor bei der Wirkung von Gesundheitsförderungsprogrammen [The role of self-efficacy in the effectiveness of health promotion programs].* Poster Session presented at the 10. Kongress für Gesundheitspsychologie, Bremen, Germany, 27th – 30th Septemper, 2010.

Parschau, L., Richert, J., Reuter, T., & Wiedemann, A.U. (2010). *An internet-based health promotion program: Who will return to follow-up assessments?* Poster session presented at the 24th conference of the European Health Psychology Society, Cluj, Romania, 1st-4th September 2010.

Parschau, L., Richert, J., & Wiedemann, A.U. (2010). *Ein Online-Programm zur Gesundheitsförderung: Welche soziodemografischen, sozial-kognitiven und verhaltensbezogenen Faktoren erhöhen die Wahrscheinlichkeit eines Drop-outs? [An online health promotion program: which socio-demographic, social-cognitive and behavior related factors increase the likelihood of drop-out?]* Paper presented at the 10. Kongress für Gesundheitspsychologie, Bremen, Germany, 27th – 30th Septemper, 2010.

Richert, J., & Lippke, S. (2010). *The validation of a scale measuring engagement in the domain of health promotion interventions.* Poster session presented at the 24th conference of the European Health Psychology Society, Cluj, Romania, 1st-4th

September 2009.

Wendt, S., Richert, J., Mohr, C., & Dill, J. (2010). *A Comparison of a 3 Stage & 4 Stage Model of Walking Behavior.* Poster session presented at the 5th Annual Northwest Ecological Community Conference, Portland, Oregon, 15th October 2010.

Wiedemann, A.U., Richert, J., & Koring, M. (2010). *Action plans: Effects of experimentally manipulated number of plans and plan recall.* Paper presented at the 24th conference of the European Health Psychology Society, Cluj, Romania, 1st-4th September 2010.

2009

Koring, M., Wiedemann, A., & Richert, J. (2009). *Who benefits from planning? Moderators of intervention effectiveness.* 23rd Annual Conference of the European Health Psychology Society. Poster session presented at the 23rd conference of the European Health Psychology Society, Pisa, Italy, 24th September 2009.*

Richert, J., & Lippke, S. (2009). *Diagnostik von Stadien aufgrund eines externen vs. selbstgewählten Kriteriums – Vergleich und Vorteile der verschiedenen Ansätze [Diagnostic of stages based on external and self-set criteria – comparison and advantages of different approaches]* Paper presented at the 41. Jahrestagung der Arbeitsgruppe für Sportpsychologie, Leipzig, Germany, 21st May 2009.

Richert, J., & Lippke, S. (2009). *Stage-Allocation Criteria in the domain of physical activity.* Poster session presented at the 23rd conference of the European Health Psychology Society, Pisa, Italy, 24th September 2009.

Richert, J., Lippke, S., & Plotnikoff, R.C. (2009). *The Protection Motivation Theory within the Stages of the Health Action Process Approach: A test in the exercise domain.* Paper presented at the 23rd conference of the European Health Psychology Society, Pisa, Italy, 24th September 2009.

Richert, J., Lippke, L., & Schwarzer, R. (2009). *Macht es Sinn, aktiv Sporttreibende danach zu unterscheiden, ob sie etwas verändern wollen oder nicht? – Eine Weiterentwicklung des HAPA um ein viertes Stadium [Does it make sense to differentiate physically active individuals based on their intention to change? – Expanding the Health Action Process Approach by a fourth stage]*. Paper presented at the 9. Kongress für Gesundheitspsychologie, Zürich, Swizerland, 25th August 2009.

Schwarzer, R., Richert, J., Mallach, N., Cao, D., Reuter, T.(2009). *Selbstwirksamkeit moderiert die Intentions-Planungs-Verhaltens-Kette [Self-efficacy moderates the intention-planning-behavior chain]*. Paper presented at the 9. Kongress für Gesundheitspsychologie, Zürich, Swizerland, 25th August 2009.

2008

Richert, J., Lippke, S. &, Ziegelmann, J. P. (2008). *Intervention-Engagement and its role in the effectiveness of stage-matched interventions: Evidence from an online study on physical activity.* Poster session presented at the XXIX International Congress of Psychology 2008 (ICP), Berlin, Germany, 23.07.2008.

Richert, J., Lippke, S., Ziegelmann, J. P. &, Schwarzer, R. (2008). *Effects of engagement in interventions on their outcomes: Evidence from a randomized controlled trial.* Paper presented at the DHP/EHPS Conference, Bath, England, 9th-12th September 2008.

* won the 2009 EHPS Young Researcher Award

2007

Richert, J., Lippke, S., Ziegelmann, J. P., & Schwarzer, R. (2007). *Interventions-Engagement und dessen Bedeutung für die Wirksamkeit stadienpassender Interventionen: Eine Online-Studie zu körperlicher Aktivität [Intervention-Engagement and its role in the effectiveness of stage-matched interventions. An online study on physical activity].* Paper presented at the 11. Tagung der DGPs-Fachgruppe Pädagogische Psychologie, Berlin, Germany, 19th September 2007.

.

Erklärung

Hiermit versichere ich, dass ich die vorgelegte Arbeit selbständig verfasst habe. Andere als die angegebenen Hilfsmittel habe ich nicht verwendet. Die Arbeit ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Jana Richert

Berlin, März 2011