Individual difference measures for work and career: Utilizing the advantages offered by latent variable methods to address measurement and construct validity issues

Der Fakultät Wirtschaftswissenschaften der Leuphana Universität Lüneburg
zur Erlangung des Grades

- Dr. phil. -

vorgelegte Dissertation von Anne Herrmann
geb. in: Teterow, Deutschland

Betreuer und Gutachter:	Prof. Dr. Hans-Rüdiger Pfister, Leuphana Universität Lüneburg
Gutachter:	Prof. Dr. Alexander Freund, Leuphana Universität Lüneburg
Gutachter:	Prof. Dr. Cornelius König, Universität des Saarlandes

September 2013

Acknowledgments

Completing my dissertation has been a challenging endeavor. I feel very grateful for having received support and encouragement from many individuals. I would like to express my gratitude to my first supervisor, Prof. Dr. Hans-Rüdiger Pfister, for encouraging me to pursue my academic research while giving me great freedom to work independently. He also provided me with the privilege to spend more than two years at the Institute of Experimental Industrial Psychology at Leuphana University where I conducted the main part of my PhD research. I would like to thank my dissertation committee, Prof. Dr. Alexander Freund and Prof. Dr. Cornelius König, for supporting me. Special thanks to Prof. Dr. Andreas Hirschi for our collaborative work. He challenged me to do the best I can. Friends, family and colleagues have supported me in various and countless ways during my work on the dissertation. I cannot see how I could possibly list all of them and, more importantly, how I could thank them appropriately by means of mentioning them here. Hence, let me just say: You all know who you are and how much I appreciate how you have helped me.

Table of Content

Detailed Table of Content

Synopsis

Introduction

Research in work and organizational psychology frequently conducts studies where psychological constructs are investigated that are not directly observable. Instead these constructs have to be assessed indirectly. One commonly used approach to measure the degree to which a construct is exhibited by an individual is using self-report questionnaires. These questionnaires contain a set of items that are considered to be indicators of the construct of interest. In order to conduct research in the field of work and organizational psychology and reach meaningful conclusions, reliable and valid measures of these constructs are needed. Evidence of the reliability and validity of these measures has to be provided based on thorough research.

Existing questionnaires were typically developed using classical test theory (Allen & Yen, 2002). The internal structures of the instruments were usually examined using exploratory factor analysis (EFA). To quantify the strength of a relationship between two constructs, a correlation coefficient based on manifest scores is often used in published research. However, more recently, latent variable approaches have been introduced. Compared to previously applied methods, they offer more in-depth analyses than previous methods into the functioning of self-report measures as well as into the relationships among constructs. Thus, they provide researchers in the field of work and organizational psychology with new opportunities to examine the instruments and determine if they are suitable to obtain meaningful results. They also offer new approaches to investigate the relationships between constructs, particularly when assessed over time.

The research conducted in this PhD thesis and reported in three papers aims at utilizing these opportunities to examine the measurement properties of a selection of self-report questionnaires. It focusses on investigating the reliability and validity of questionnaires whereby some have already been used frequently while others were developed only recently. All are relevant for research studies that address psychological questions relevant for the world of work. The aim is to provide researchers in this research area with a thorough assessment of the measurement properties of these self-report questionnaires. The findings are intended to demonstrate their suitability for future research studies conducted in work and organizational psychology as well as for practical applications.

Advantages offered by latent variable approaches to address measurement issues

Applying different methods based on latent variables offers several advantages over previous approaches when addressing measurement issues. The methods provide a better understanding of how each item functions and how it represents the constructs. Furthermore, the measurement invariance of an instrument can be assessed using structural equation modeling (SEM), one latent variable approach. This means it can be determined if a scale measures the same constructs for different groups. While some other methods to assess measurement invariance exist, e.g. based on logistic regression (Swaminathan & Rogers, 1990), these are very laborious and hence not applied frequently. Assessing measurement invariance using SEM is far more straightforward. Increasingly strict constraints can be applied to test for different levels of invariance (Lance, Vandenberg, & Self, 2000). Once measurement invariance of a scale has been established, substantial conclusions on differences between groups on the particular psychological construct are warranted and comparisons between groups become meaningful. Thus, evidence of measurement invariance is particularly relevant as it makes the application of psychometric instruments in high-stake situations such as personnel selection defensible (Horn & McArdle, 1992).

Increasingly, researchers are urged to collect data on more than one measurement point. One reason for advocating longitudinal studies is that it allows investigating and hence better understanding developmental patterns of psychological constructs. What is more, longitudinal studies are useful to shed light on causal mechanisms, which are not possible in cross-sectional survey research. In these instances, when longitudinal data has to be analyzed, the advantages of latent variable approaches become apparent and the benefits they offer can be fully utilized.

First, based on SEM, it is possible to determine a scale's measurement invariance across time. This is a prerequisite for subsequent analyses based on longitudinal assessments (Horn & McArdle, 1992).

Second, the differential stability across time can be assessed as the standardized covariance between the latent variables at two measurement points. This approach overcomes some issues that occur when assessing the test-retest reliability as a Pearson correlation coefficient between manifest scores at both measurement points. The Pearson correlation coefficient is attenuated due to measurement error for the construct at each measurement point. At the same time, it is inflated due to the specific variance shared between like-items across time. Assessing temporal stability based on a latent variable model controls for these effects by (a) allowing error terms of like-items across time to correlate, and (b) controlling for measurement error in the latent variable (Marsh et al., 2010).

Third, latent growth modeling (LGM), another latent variable approach, can be applied to explore developmental patterns of psychological constructs in longitudinal research. It also allows for investigating if and how developmental patterns of constructs may be related.

Finally, cross-lagged analysis (CLA) can be used when longitudinal data is available. In CLA, the potential causal mechanisms between the constructs of interest can be examined. This is possible because in a CLA the internal stability over time and the concurrent relationships of two constructs are controlled for. Hence this method offers a complementary perspective to LGM by focusing on how the variables are related to each other from one point in time to the next.

Utilizing the advantages offered by latent variable approaches to investigate the reliability and validity of individual difference measures applied in organizational research

Self-report questionnaires applied in work and organizational psychology are required to be reliable and valid measures of the constructs under examination. As outlined above, methods based on the latent variable approach offer new and superior opportunities to assess the measurement properties of these self-report questionnaires. The methods hence recommend themselves to be applied to provide cumulative evidence of the scales' suitability to address substantial research questions. The three papers presented as part of this PhD thesis utilize the benefits of different latent variable approaches to provide new insights into the functioning of well-established as well as newly developed self-report questionnaires. Thus, the findings support the application of the scales in future research studies. In addition, conceptual questions are addressed informed by the research gap identified for the scales under examination.

The studies presented across the three papers constituting the PhD thesis share an overarching theme by having these four aspects in common: All three papers (1) focus on the insights gained from self-reports of individuals in non-experimental settings, (2) are concerned with questionnaires that assess psychological constructs related to and relevant for the context of work and career, (3) thoroughly examine the measurement properties of these questionnaires, in particular with regard to their reliability and validity, and (4) apply statistical methods based on latent variable approaches to assess their measurement properties.

The first paper conducts research based on two established and well-researched personality instruments. It investigates the frequently debated issue of the internal structure of these two multi-scale personality instruments, thus addressing a concern about construct validity. The second paper includes several single-scale career attitude measures previously used in published research. The paper examines the differential construct validity and aspects of predictive validity of these measures in a longitudinal study. Finally, the third paper investigates the measurement properties of a scale which has only shortly before been developed in English and deployed in one published study. For the research presented here, the scale was first translated into German. The paper then addresses questions with regard to the scale's internal structure, development over time and its predictive validity.

The research questions, the analytical approaches, the results and the contributions of each paper are described in the following sections.

I. Paper "Simple Measures and Complex Structures: Is it worth employing a more complex model of personality in Big Five inventories?"

The first paper, published in the *Journal of Research in Personality*, addresses a question arising from the poor performance of five-factor personality inventories in confirmatory factor analyses (CFA). Repeatedly and across several personality instruments, inacceptable model fit was obtained when applying CFA. The supposedly simple structure of the instruments could not be confirmed. This has prompted some researchers to question the construct validity of these personality measures. Others have doubted that the strict assumptions imposed onto the data in CFA are appropriate. They suggested applying Exploratory Structural Equation Modeling (ESEM) instead (Marsh et al., 2010).

Rather than joining the debate on the suitability or superiority of any of these methods to describe personality, the research question was focusing on the implications of applying CFA and ESEM by asking: What is the impact that the application of either method has on the construct validity of personality inventories? Thus, the contribution of the paper is to allow for a more nuanced discussion that takes the implications on scores and hence on the construct validity of using either CFA or ESEM into account.

For this purpose, CFA and ESEM were applied to construct better-fitting – though more complex – models. The investigation is based on data from a diverse sample of 620 respondents who completed two established and well-researched personality questionnaires (NEO PI-R and 16PF questionnaire). The congruence between the original version of each instrument and the modified versions obtained from applying CFA or ESEM was examined by computing correlations between scores obtained from either version. In addition, changes in construct validity of the original and the modified versions of both questionnaires were examined using the multitrait-multimethod approach first introduced by Campbell and Fiske (1959). With this method, the convergent validity of two instruments is determined by assessing the agreement in measuring the same constructs, i.e. the five matching factors across both questionnaires. Furthermore, the discriminant validity is assessed across and within both instruments. The discriminant validity is supported when the correlations between factors that are conceptually unrelated are low.

With regard to the congruence between the original and the modified model scores, the study found that scores derived from either method do not differ substantially, as indicated by high correlations between the original scores and the scores based on CFA and ESEM. When applying ESEM, the convergent validity declines but the discriminant validity improves in comparison to the results obtained from the original model versions of both personality inventories. When applying CFA, the convergent as well as the discriminant validity decrease in both instruments.

The research reported in this paper contributes to the ongoing debate on the internal structure of personality that has arisen from the application of CFA. It provides researchers and practitioners alike with important information by quantifying the relative score changes that one would expect if one were to implement the more complex model of personality suggested by ESEM and CFA. The decrease in construct validity observed when applying the more complex structure of personality proves that retaining the simple structure of the current questionnaires is not only a defensible option, but may even be favorable. Its findings hence provide support for the continued use of current personality inventories in research and practical applications. Furthermore, the results also refute potential concerns regarding the validity and applicability of previous research based on current personality instruments that has been raised when the inventories failed to be supported by CFA.

From a conceptual viewpoint, the research reported in the first paper shows that it may be ill-advised to reject personality theory based on CFA results. The two Big Five models used to describe personality may not necessarily account for every relationship between subscales. However, this study shows that increasing the complexity has in fact a negative impact on the construct validity of the measures, thus providing a less useful assessment of an individual's personality. The domains in the more complex models contained such a large number of subscales that made it impossible to interpret the results. In addition, the conceptual overlap of domains due to multiple assignments of subscales onto domains further decreased the insights one could gain from assessing personality applying the more complex model. This shows that it is necessary to consider that theories are designed to explain phenomena and need to simplify the more complex relationships observed between constructs in the real world. Meehl (1990) argues that models can be useful even if they simplify reality. The results of this study suggest that models *need* to simplify reality so that they can be useful.

II. Paper "Calling and Career Preparation: Investigating Developmental Patterns and Temporal Precedence"

The second paper, published in the *Journal of Vocational Behavior*, demonstrates how latent variable methods can be fruitfully applied to assess the relationships and developmental patterns of several related career attitudes based on longitudinal data. Just like personality traits, career attitudes are typically assessed using self-report questionnaires. However, career attitudes tend to undergo more changes than personality, particularly during time at university.

Thus, to fully understand the relationships between career attitudes, it is essential to assess them over a period of time. Conducting longitudinal rather than cross-sectional research allows examining how these constructs may influence each other over time. It hence provides a comprehensive understanding of the relationships between career attitudes.

This paper addresses questions on the relationships and developmental patterns of calling and three dimensions of career preparation (i.e., career planning, career decidedness, and career self-efficacy beliefs) based on longitudinal data and utilizing several latent variable methods. Dobrow and Tosti-Kharas (2011, p. 1003) define calling as a consuming, meaningful passion people experience toward a career domain. Calling and career preparation are assumed to be closely related. The study aimed at exploring the nature of and reason for this relationship as this issue has not been satisfyingly investigated in other studies so far. We assessed calling as well as the three dimensions of career preparation in a one-year longitudinal three-wave study based on a diverse sample of German university students (N = 846). The constructs were assessed using single-scale instruments previously applied in career-related research. To explore the development of the investigated constructs as well as the relationships between calling and career preparation in a more encompassing way, we applied LGM and CLA.

Before addressing these substantial questions, the measurement properties of the scales applied in this study were examined thoroughly. Using CFA, we show that calling is a distinct construct that captures something different than the career preparation scales. Furthermore, we provide evidence of the internal consistency within each scale. In addition, we established the longitudinal measurement invariance of the scales deployed in our study, thus providing support for their appropriateness for the subsequent analyses based on longitudinal data (Horn & McArdle, 1992). We hypothesized the existence of reciprocal effects between calling and the three dimensions of career preparation. To investigate relationships in the developmental patterns of the constructs, we applied LGM, thus utilizing the benefits this method offers for the analysis of longitudinal data. We found positive correlations between the intercepts of calling and the intercepts of all three career preparation measures. With regard to similarities in the developmental patterns, we found the slope of calling to be positively related to those of career decidedness and self-efficacy belief but not to career planning.

Finally, the application of CLA showed that calling predicts a subsequent increase in self-efficacy beliefs. Reversely, career decidedness predicts an increase in the presence of calling. For calling and career planning, a reciprocal relationship was found, i.e. calling and career planning influence each other over time.

These findings advance the literature by showing that having a sense of control over one's vocational development, clarity about personal preferences and career goals (i.e., career decidedness) as well as envisioning future career stages and possible selves (i.e., career planning) can strengthen and confirm a sense of calling among university students. The findings of the study advance the theoretical understanding of how a calling develops as well as how and why it is related to other prominent career development constructs, specifically, three dimensions of career preparation. In sum, our results suggest that showing higher career preparedness in terms of career decidedness and career planning can help people to develop and/or confirm a sense of calling in their careers. In turn, experiencing a calling appears to be a motivating force for engaging in career preparation. Having a calling might thus help to navigate a complex career terrain and address career development tasks (Hall & Chandler, 2005).

Gaining a more nuanced understanding of the relationships between calling and the three dimensions of career preparation can also inform the work of career counselors. Our results imply that helping clients to find or develop a calling can be beneficial because having a calling is related to an increased engagement in career preparation. This may have positive effects on the general ability of an individual to cope with vocational demands. Furthermore, encouraging university students to engage in the various aspects of career preparation might also be important in order to help them in developing a calling.

In sum, the paper’s contribution is two-fold. First, it demonstrates the construct validity of four important career scales and proves their suitability to be applied in longitudinal research. Second, it increases our understanding of how and why the presence of a calling is related to career preparation and in doing so also enhances our knowledge of how callings emerge and develop over time.

III. Paper “The Protean Career Orientation: Investigating Gender Differences, Temporal Stability, and Predictive Utility”

The third paper, under review in the *Journal of Applied Psychology* at the time of this writing, investigates a protean career orientation which is considered as one of the important career concepts that characterize the contemporary ‘new careers’ (Enache, Sallan, Simo, & Fernandez, 2011; Sullivan & Baruch, 2009). Individuals with a protean career orientation feel responsible for managing their own career and are self-directed in their career development. They are guided by their own values in pursuing their career goals. In comparison, individuals with a traditional career orientation rely more on their organization to take responsibility for their career progression. Having a traditional career orientation is also characterized by placing a higher importance on objective career success such as salary and position (Baruch, 2008). Considering the increasingly dynamic work environment and the decline in traditional career progression within organizations, individuals with higher levels of a protean career orientation may be more capable to navigate their way through their career. Thus, it becomes ever more important to understand a protean career orientation. To investigate this career orientation, it is hence desirable (a) to be able to assess individuals with regard to their level of a protean career orientation, and (b) to conduct research into the protean career orientation to understand the mechanisms related to this fairly new construct, such as its temporal stability and its incremental predictive utility.

A protean career orientation is frequently contrasted with a traditional career orientation (Hall, 1996, 2004). However, existing empirical research did not assess similarities and differences between the two career orientations. In the research reported here, we therefore used a scale assessing a protean career orientation which also measures a traditional career orientation based on a separate set of items. We conducted several studies based on independent samples as part of this research that applied this scale measuring both career orientations. Advancing extant research, the measure we applied allows to empirically examine the relation of the two career orientations regarding several criterion variables and to compare the respective effects. Where applicable, we used this advantage in the reported studies by contrasting results obtained from both career orientations.

Based on a one-dimensional measure for a protean career orientation, we present validation and model testing studies with three independent samples. First, we provide evidence of the scale’s construct validity by establishing its unidimensionality. We further show that the protean career orientation is distinct from a traditional career orientation. However, the two orientations are positively correlated rather than just being at opposite end of the career orientation spectrum. We demonstrate the scale's measurement invariance across gender separately for a sample of university students (*N*=1,224) and working professionals (*N*=526). This proves that the scale assesses the same construct for both genders and can be applied to examine gender differences in a protean career orientation. No differences were found between genders in either sample. We also provide support of its measurement invariance between university students and employees. This demonstrates that the scale is suitable to assess this career orientation at different stages in the career and that meaningful comparison between both groups can be made. We found no significant differences in their levels of protean career orientation. The findings from both group comparisons indicate that males and females as well as university students and working professionals are equally well-equipped with regard to their career orientation to function in a work environment that requires them to rely less on an organization to take care of their career progression but instead to be self-directed in managing their professional development.

Second, we examined the measurement invariance across six months among university students (*N* = 419) and working professionals (*N* = 156). We found support for the measurement invariance of the protean career orientation scale across time. Thus, this new scale assessing an increasingly important career orientation is suitable to be used in future studies using a longitudinal research design. This is particularly relevant because sophisticated career career-related research usually needs data collected at several points in time. We also established the scale's differential stability over the course of six months. We found the protean career orientation to be more stable among professionals than among students. This finding is plausible considering that the latter are still at university exploring career options. This makes changes in their career orientation more likely in comparison to individuals who have gained work experience and can be expected to have developed a more stable attitude towards what matters to them in their careers.

Third, and moving on to address conceptual questions related to career attitudes, we could show that a protean career orientation partially mediates the relation between a proactive personality and proactive career behaviors among university students and working professionals. Fourth, we could demonstrate that a protean career orientation partially mediates the relation between core self-evaluations and career satisfaction among working professionals. The last two findings provide some insights into the importance of a protean career attitude in explaining career outcomes. They provide an understanding of how this career attitude contributes to promoting career satisfaction and displaying proactive career behaviors. Finally, based on a cross-lagged study, we could show that career satisfaction predicts a protean career orientation. However, a protean career orientation does not predict an increase in career satisfaction.

Overall, the paper contributes to the field of career research in two ways. First, it thoroughly examines the measurement properties of a scale developed to assess the protean career orientation. The research conducted here therefore provides researchers with a well-examined scale that can be applied in future studies to address further questions into an increasingly important career orientation. Demonstrating measurement invariance across time and gender also supports the scale's application in longitudinal studies and when examining career attitudes in relation to gender. Second, the paper examines several relevant conceptual questions that help to better understand how a protean career orientation acts as a mediator between personality characteristics and other career-related constructs. Investigating these conceptual questions, we answered calls for advancing career studies investigating the so-called 'new careers' through rigorous empirical evaluations of newly emerging constructs (Arthur, 2008; Inkson, Gunz, Ganesh, & Roper, 2012; Sullivan & Baruch, 2009).

Summary and Discussion

Research in work and organizational psychology frequently utilizes self-report questionnaires to gain insights in psychological constructs relevant in the work context. In order to conduct such research while ensuring that meaningful conclusions can be drawn, it is essential to have reliable and valid measures of the constructs under investigation. For this purpose, sophisticated data-analytical methods are needed that allow for a thorough investigation of the instruments to understand the functioning of the items, the scales as well as the relationships between scales.

A new set of statistical methods based on latent variables has been made available to researchers recently. These methods offer numerous opportunities for a more thorough investigation of the measurement properties of self-report instruments. The research in this PhD thesis utilized these opportunities. It thoroughly examined the measurement properties of several self-report questionnaires based on a set of latent variable methods. The results offer researchers a more detailed account of the suitability of these measures to be applied in cross-sectional and longitudinal research. The findings also provide practitioners with confidence that the instruments have acceptable measurement properties and that their application in organizational settings is defensible. Supporting the instruments' measurement invariance for gender and across time is particularly relevant for researchers as well as practitioners. It provides researchers with the required information to employ the measures in studies looking at gender differences and investigating the development of constructs over time. It is also relevant for practitioners who need to ensure that measures used in high-stake situations offer fair assessments for candidates from different groups. Furthermore, the research on measurement invariance reported here demonstrates how latent variable methods can be used to address methodological and substantial questions simultaneously.

To conclude, the cumulative findings of the herein conducted PhD research advance the field of work and organizational psychology by providing important new insights regarding the validity of different self-report questionnaires in research and practice. At the same time, they have demonstrated the utility of latent variable methods in gaining a more in-depth understanding of individual difference measures.

References

Allen, M., & Yen, W. (2002). Introduction to Measurement Theory. Long Grove, IL: Waveland Press.

Arthur, M. B. (2008). Examining contemporary careers: A call for interdisciplinary inquiry. *Human Relations, 61*(2), 163-186. doi: 10.1177/0018726707087783

Baruch, Y. (2008). *Development and validation of a measure for protean career*. Paper presented at the 68th Annual Meeting of the Academy of Management, Anaheim, CA.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin, 56*(2), 81. doi: 10.1037/h0046016

Dobrow, S. R., & Tosti-Kharas, J. (2011). Calling: The development of a scale measure. *Personnel Psychology, 64*(4), 1001-1049. doi: 10.1111/j.1744-6570.2011.01234.x

Enache, M., Sallan, J. M., Simo, P., & Fernandez, V. (2011). Examining the impact of protean and boundaryless career attitudes upon subjective career success. *Journal of Management & Organization, 17*(4), 459-473. doi: 10.5172/jmo.2011.17.4.459

Hall, D. T. (1996). Protean careers of the 21st century. *The Academy of Management Executive, 10*(4), 8-16. doi: 10.5465/AME.1996.3145315

Hall, D. T. (2004). The protean career: A quarter-century journey. *Journal of Vocational Behavior, 65*(1), 1-13. doi: 10.1016/j.jvb.2003.10.006

Hall, D. T., & Chandler, D. E. (2005). Psychological success: When the career is a calling. *Journal of Organizational Behavior, 26*(2), 155-176. doi: 10.1002/job.301

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*(3-4), 117-144. doi: 10.1080/03610739208253916

Inkson, K., Gunz, H., Ganesh, S., & Roper, J. (2012). Boundaryless careers: bringing back boundaries. *Organization Studies, 33*(3), 323-340. doi: 10.1177/0170840611435600

Lance, C. E., Vandenberg, R. J., & Self, R. M. (2000). Latent Growth Models of Individual Change: The Case of Newcomer Adjustment. *Organizational*

Behavior and Human Decision Processes, 83(1), 107-140. doi: 10.1006/obhd.2000.2904

Marsh, H. W., Lüdtke, O., Muthen, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22*(3), 471-491. doi: 10.1037/a0019227

Meehl, P. E. (1990). Appraising and Amending Theories: The Strategy of Lakatosian Defense and Two Principles that Warrant It. *Psychological Inquiry, 1*(2), 108-141. doi: 10.1207/s15327965pli0102_1

Sullivan, S. E., & Baruch, Y. (2009). Advances in career theory and research: a critical review and agenda for future exploration. *Journal of Management, 35*(6), 1542-1571. doi: 10.1177/0149206309350082

Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27*(4), 361-370. doi: 10.1111/j.1745-3984.1990.tb00754.x

I. Paper "Simple Measures and Complex Structures: Is it worth employing a more complex model of personality in Big Five inventories?"

Abstract

The poor performance of five-factor personality inventories in confirmatory factor analyses (CFA) prompted some to question their construct validity. Others doubted the CFA's suitability and suggested applying Exploratory Structural Equation Modeling (ESEM). The question arises as to what impact the application of either method has on the construct validity of personality inventories. We addressed this question by applying ESEM and CFA to construct better-fitting, though more complex models based on data from two questionnaires (NEO PI-R and 16PF). Generally, scores derived from either method did not differ substantially. When applying ESEM, convergent validity declined but discriminant validity improved. When applying CFA, convergent and discriminant validity decreased. We conclude that using current personality questionnaires that utilize a simple structure is appropriate.

Complete reference: Herrmann, A. & Pfister, H.-R. (2013). Simple measures and complex structures: Is it worth employing a more complex model of personality in Big Five inventories? Journal of Research in Personality, 47(5), 599–608. doi: 10.1016/j.jrp.2013.05.004

Introduction

Researchers who investigate normal adult personality have reached a consensus on five broad factors, often called the 'Big Five' (Goldberg, 1990), and on their conceptual definitions (Digman, 1990; McCrae & Costa, 1999; Norman, 1963). These factors are known as Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness, although other terms are used as well. This general consensus has allowed for cumulative research and meta-analyses of important aspects of the construct, including the development of personality over an individual's lifespan (Judge, Higgins, Thoresen, & Barrick, 1999; Terracciano, McCrae, & Costa, 2010), differences between groups (Goldberg, Sweeney, Merenda, & Hughes, 1998; Schmitt, Realo, Voracek, & Allik, 2008), the existence of a general factor of personality (Musek, 2007; van der Linden, te Nijenhuis, & Bakker, 2010), a prediction of external criteria (Grucza & Goldberg, 2007; Hurtz & Donovan, 2000), and many more. In research and practice, personality is predominantly assessed using self-report questionnaires. Many of these questionnaires contain items that contribute to one of many first-order scales that are combined to represent the Big Five factors.

The internal structure of personality, i.e., the assignment of subscales to the five factors, has commonly been examined using an exploratory factor analysis (Aluja, Rossier, Garcia, & Verardi, 2005; Cattell & Cattell, 1995; Costa & McCrae, 1992b). This assignment is extremely important because it forms the basis for obtaining scores for the higher-order personality factors. In general, a simple structure (Thurstone, 1947) where each first-order scale is uniquely assigned to only one of the Big Five factors is assumed to be appropriate.

As in many other research areas in which constructs are assessed using self-report questionnaires, CFAs were eventually applied to personality data. The results of these studies were largely discouraging. The CFA model fit indices frequently exceeded proposed cut-off values for acceptable model fits and, based on CFA standards, did not confirm the simple structure (Church & Burke, 1994; Hopwood & Donnellan, 2010; McCrae, Zonderman, Costa, Bond, & Paunonen, 1996; Vassend & Skrondal, 2011). Several cross loadings (i.e., links between first-order scales and factors other than the originally postulated higher-order personality factors) usually needed to be included in the model to achieve an acceptable fit. The more complex models, however, were difficult to interpret and often displayed less of a good fit in cross-validation samples (e.g.,Church & Burke, 1994; Hopwood & Donnellan, 2010).

This has raised concerns if the currently proposed composition of the broad factors provides an adequate assessment of an individual's personality. These higher-order scores are commonly used in research studies and in practical applications of personality instruments. Thus, confidence is required regarding the suitability of the Big Five factors as a 'common language' for describing personality. Adding additional cross loadings as suggested by CFA also changes the meaning of the observed scores. Subsequently, one must question how the construct validity of personality instruments is affected when subscales contribute to more than one broad factor.

In the present study we address these concerns in two ways: First, we determine the 'change of scores' which – in this examination – refers to a difference in the relative position of an individual within a sample on the trait continuum measured as the correlation between the original scores and scores obtained after incorporating the CFA cross loadings. Second, we examine the impact on the instruments' construct validity resulting from the modified models.

To complement our investigation and consider more recent trends in factor analysis, we also apply Exploratory Structural Equation Modeling (ESEM, Asparouhov & Muthen, 2009), a method that integrates CFA and EFA. ESEM is less restrictive than CFA as it does not constrain the non-target loadings to be zero. In difference to CFA, in ESEM a model can be specified only with regard to the number of factors. Further restrictions can be added and tested using chi-square difference tests. In difference to EFA, ESEM provides typical CFA parameters, such as standard errors and goodness of fit statistics as well as the possibility to test for measurement invariance between groups and across time (Asparouhov & Muthen, 2009). Due to these possibilities and advantages of ESEM, it has been promoted to be applied in the psychometric evaluation of psychological instruments (Marsh, Liem, Martin, Morin, & Nagengast, 2011).

We applied a CFA and ESEM to data from 620 respondents who completed two established personality questionnaires (the NEO PI-R and the 16PF questionnaire). Using two different sets of modification criteria to determine cross loadings when conducting the CFA, we generated two more complex models for each instrument. We computed scores based on these modified CFA models using two different approaches: (a) we applied the scoring rules for the instrument provided in the respective test manual but added the additional subscales, as identified in the CFA, and (b) we used the factor scores obtained from the respective modified CFA model. The first approach mirrors current usage in research, in which manifest, rather than latent, Big Five scores are employed (Barrick & Mount, 1996; Grucza & Goldberg, 2007; Hurtz & Donovan, 2000; Salgado, 2003). The second approach uses scores that correspond more directly with the CFA models. With regard to the application of ESEM, we used the factor scores obtained from applying the method from both instruments.

To assess the relative score changes, we computed correlations between scores from the original model and the scores obtained from the CFA and ESEM models. The results of this analysis support a more nuanced discussion of the discrepancy between current personality theories and the more complex model of personality, as suggested by the CFA. Applying ESEM offers further insight into how Big Five scores based on a more recent factor-analytical method.

To determine the impact on the questionnaires' construct validity, we applied the multitrait-multimethod (MTMM) approach, which was developed by Campbell and Fiske (1959), to the original model as well as the models proposed by CFA and ESEM. A comparison of the MTMM results across the models showed the extent to which the relationships within and between the five factors of both instruments changed as one moved from a simple to a more complex structure, thus determining changes in the convergent and discriminant validity.

Previous studies have focused mainly on investigating the congruence between results obtained from the EFA and CFA of an instrument without examining the impact of the observed discrepancies on scale scores and construct validity beyond the internal structure (e.g., Aluja, Blanch, & Garcia, 2005; Borkenau & Ostendorf, 1990; McCrae, et al., 1996). In other studies, CFAs were applied to several instruments, but it was not determined how the relationships between the constructs were affected by changes in the model proposed by the CFAs (e.g., Church & Burke, 1994; Hopwood & Donnellan, 2010). In our study, we address those gaps by determining how the scores of and the relationships between personality scales change when the internal structure is more complex, as suggested by CFA. As a result, we extend the examination of construct validity beyond the internal structure to focus on changes in the convergent and discriminant validity within and across the two instruments. The study thus follows a suggestion made, among others, by Hopwood and Donnellan (2010) that "there is a need to document that misspecifications have practical or substantive consequences beyond simply contributing to model misfit" (p. 343).

Considering the complexities and difficulties in identifying the correct model in CFA based on modification indices and other model assessment criteria (Fan & Sivo, 2007; MacCallum, Roznowski, & Necowitz, 1992), we do not aim at determining the "true" model of personality. Instead, we provide an empirical illustration, i.e., to demonstrate by way of example the impact that this added complexity would have on scores and construct validity. By also applying ESEM to both instruments, we shed light on how this more recent but increasingly used method may affect the resulting factor scores and subsequently the instruments' construct validity.

Method

Measures

The data from two hierarchical self-report personality instruments were used in this study:

1) Cattell's 16 Personality Factor Questionnaire, 5th Edition (16PF, Conn & Rieke, 1994) consists of 185 items with a three-choice response format that measures 16 primary factors. The 15 non-cognitive factors are then combined into five factors, commonly called 'global factors'.

2) The Revised NEO Personality Inventory (NEO-PI-R, Costa & McCrae, 1992b) comprises 240 items with a five-point Likert response format. It assesses 30 facets of personality that are used to compute five higher-level domain scores.

The 16PF and the NEO-PI-R differ in that the first-order level of personality is described with 15 and 30 scales, respectively. An alignment exists, however, between the second-order level, where there is a NEO domain counterpart for each 16PF global factor. The counterparts for both instruments are 16PF-Extraversion and NEO-Extraversion, 16PF-Anxiety and NEO-

Neuroticism, 16PF-Self-Control and NEO-Conscientiousness, 16PF-Independence and NEO-Agreeableness and, finally, 16PF-Tough-Mindedness and NEO-Openness (Cattell & Mead, 2008). The last two pairs are defined in the opposite direction.

Different views exist on when to consider a psychometric questionnaire a "Big Five Instrument". We follow a definition by McCrae and John (1992): “The five-factor model of personality is a hierarchical organization of personality traits in terms of five basic dimensions” (p. 175) which applies to the NEO PI-R as well as the 16PF. These two Big Five instruments were included in this study because they differ profoundly in their development and in the approach to computing the second-order factors. This method safeguards against drawing conclusions about personality constructs that are actually a result of characteristics of a particular instrument. The 16PF questionnaire was developed based on empirical analyses. An EFA of the item parcels was carried out to identify the primary personality traits. These primary factors were subjected to a second-order EFA to extract five global factors (Cattell & Cattell, 1995). Based on the size of the EFA factor loadings, a set of contributing primary factors and their respective weightings were selected in the computation of each global factor score. The global factor computation is thus data-driven with regard to the assignment of primary factors and their relative importance. This approach also resulted in multiple assignments; six of the 15 first-order factors contributed to two global factors. The authors of the NEO PI-R instrument reached a consensus on five factors and used those factors as a starting point to develop a hierarchical model of personality (Costa & McCrae, 1992b). Based on psychological literature and conceptual considerations, six facets were then determined for each factor that reflected relevant and diverse aspects of the respective higher-order construct. The five second-order factor scores are computed using unit-weighting. In other words, a simple sum score was obtained by adding up the scores of the contributing six facets. Unlike the 16PF, each facet contributed to one domain only.

Sample

The sample used in this study included 620 respondents and was a subset of the Eugene Springfield Community Sample (ESCS; see Grucza & Goldberg, 2007 for information on data collection procedures and further sample details). The original ESCS sample was slightly larger (N = 857 for the NEO PI-R and N = 680 for the 16PF). In our study, only the 620 participants who completed both instruments were included in the sample. Of these, 97% were Caucasian, 57% were female, and roughly half the sample had achieved at least a college degree. The age of the respondents ranged from 18 to 85 years old (M = 52, SD = 13).

Because both questionnaires were completed by the same sample, unknown characteristics of different samples can be ruled out as an explanation for any observed differences between instruments. This set-up also allows us to examine the construct validity across both instruments.

Analyses

CFAs and ESEM

CFAs were conducted using the software R (2012) and the package 'lavaan' (Rosseel, 2011), which have been shown to generate the same results as other software packages (Narayanan, 2012), to examine the second-order structure of the 16PF and the NEO PI-R. Because the data were non-normally distributed, we used a robust maximum likelihood estimation method that provided robust standard errors and Satorra-Bentler scaled test statistics (Satorra & Bentler, 2001). The original models were specified as follows: for each model, all loadings of the manifest variables on the five factors were assumed to be zero, except for the latent factors to which the manifest variable was assigned in the original model as specified in the test manual. The covariances between the latent variables were freely estimated in both models. Cattell used an oblique factor rotation when conducting the EFA of the 16PF during questionnaire development because it reflected his idea of interrelated personality factors (Cattell & Cattell, 1995). While the theoretical NEO PI-R model proposed orthogonal domains, the five domains displayed considerable intercorrelations which have been attributed to conceptual overlap of facets that may relate to more than one broad factor (Costa & McCrae, 1992a).

We used four fit indices that cover different aspects of model fits that were identified as particularly suitable for personality data, in which comparatively low target loadings and several secondary loadings were expected (Beauducel & Wittmann, 2005): (a) the Comparative Fit Index (CFI), an incremental fit index; (b) the Standardized Root Mean Square Residual

(SRMR), an absolute fit; (c) the Root Mean Square Error of Approximation (RMSEA), an index that favors simple over more complex models; and (d) the Satorra-Bentler corrected χ^2 (SB-χ^2) test, a significance test used when data are distributed non-normally, as was the case in our study.

The purpose of our study was to gauge the impact that the cross loadings suggested by CFA had on Big Five scores. Hence, we decided to apply two different approaches to model modification in the analyses, resulting in two alternative CFA models per personality instrument. Both approaches reflected different ideas about what should guide modifications and what type of modifications were justifiable. The modification process for Model 1, the first alternative CFA model, was guided by the modification index (MI), which provides the researcher with a direct measure of the change in the model fit chi-square if the parameter was freed. Starting from the original model for both instruments, we computed MIs for successive models, each time freeing the factor loading or residual correlation between subscales with the highest MI until an acceptable model fit was obtained. There is a lively debate on the appropriateness of using general cut-off values for goodness-of-fit statistics (Marsh, Hau, & Wen, 2004). In the absence of hard-and-fast rules, we opted for the frequently applied cut-off values suggested by Hu and Bentler (1999): a CFI greater than .90, an SRMR less than .08, and an RMSEA less than .06.

For the second alternative CFA model, Model 2, the approach to model modification was guided by the intention to control for Type I as well as Type II errors (Saris, Satorra, & van der Veld, 2009). A Type I error is present if a parameter that is fixed to zero in the original model is classified as a misspecification and is therefore estimated in the revised model even though its population value is zero. A Type II error occurs when a parameter fixed to zero is not classified as a misspecification even though its population value is not zero (Hu & Bentler, 1998). We used the MI to identify paths to be freed, this time only releasing paths where the MI was greater than 10, thus applying a chi-square test with a significance level of .001 (df = 1). A large sample size, however, increases the likelihood of Type I errors for this value (Saris, et al., 2009). Thus, we combined information provided by the MI with the Expected Parameter Change (EPC). This value indicates the size of a currently fixed parameter if it were to be freely estimated in a revised model. It is a standardized value that can be viewed as an effect size. There are no rules as to what minimum the EPC should take on to justify freeing the respective parameter. All loadings were included that indicated that at least 10% of the variance in the manifest variable was explained by the respective latent factor. Thus, we opted for a conservative cut-off value of .316 (absolute value). This value also lies between suggested values found in the literature, such as .30 (Kline, 1994) and .40 (Saris, et al., 2009). We started from the original model, this time freeing the parameter with the highest MI and an EPC > .316 for each successive model until no further indices complied with the criteria outlined above. These rather conservative criteria were applied to avoid obtaining an over-fitted model that (a) is not replicable when fitted to another sample and (b) is not a parsimonious description of the relationships between variables (MacCallum, et al., 1992). For Model 2, the model modifications were restricted to releasing paths between indicators and latent variables. Error terms between manifest variables were not included because their conceptual meaning had been questioned (Gerbing & Anderson, 1984).

ESEM was conducted using the software Mplus 6.11 (Muthén & Muthén, 2012) and applying an oblique geomin rotation, thus allowing the factors to covary. The same model fit indices as for the CFA were computed.

Correlations between scores from the original model and the CFA and ESEM models

For the first set of scores based on the modified CFA models (M1m and M2m), we applied the scoring rules for the instrument provided in the respective test manual but added the additional subscales, as identified in the CFA. Thus, the modified NEO PI-R domain scores were obtained as a unit-weighted sum of the raw scores of the six original facets and the additional facets. The modified 16PF scores were computed as a weighted sum of the original and additional primary factors identified in the CFA. We applied the average weighting of the original primary factors for each global factor to the additional subscales, thus neither downplaying nor overestimating their impact. Research has shown that weighting only produces minor relative changes in scores compared to unit weighting, especially under conditions where the number of components is high, where these components are correlated and where their weights vary only slightly (Bobko, Roth, & Buster, 2007). All three conditions apply to the components contributing to the global factors of the three 16PF models. Therefore, we could rule out that the weighting unduly affects correlations between scores. For the second set of scores based directly on the CFA results (M1c and M2c), the factor scores of the respective modified CFA model for both instruments were calculated. In lavaan, these scores are estimated based on a regression method referred to as 'modal posterior estimator'. Correlations between the original model with the ESEM model are based on the ESEM factor scores (EM) of the respective instrument.

Two sets of correlations between scores from the original and the modified models for the factors of both instruments were computed: (a) Pearson correlation coefficients to measure the strength of the linear relationship between scores from the original and the modified models and (b) Spearman correlation coefficients to quantify the change in rank order, thus determining the concordance of the ordering of individuals on each broad domain between the original and the more complex models.

MTMM

We used the multitrait-multimethod matrix (MTMM) developed by Campbell and Fiske (1959) as a framework to compare the level of convergent and discriminant validity of both instruments across the original scores and the scores obtained from ESEM as well as from both CFA modified models based on the two different approaches to score computation. Thus, six MTMM matrices were computed. Convergent validity is confirmed when high correlations are observed for corresponding scales, i.e., for scales that measure the same constructs across both instruments (monotrait-heteromethod, MTHM). Discriminant validity reflects the idea that traits that are not conceptually related should display considerably lower correlations than the ones between corresponding traits. Discriminant validity is supported when the non-diagonal intercorrelation coefficients within one method (heterotrait-monomethod, HTMM) are low and the non-diagonal intercorrelation coefficients between the traits of the two methods (heterotrait-heteromethod, HTHM) are even lower (Campbell & Fiske, 1959). Further support of construct validity is provided when the pattern of correlations between traits is similar for both methods. To compare the evidence of construct validity of the original model with the modified models, we calculated the means for each set of coefficients constituting different aspects of convergent and discriminant validity for each model separately. For this purpose, we used Fisher's transformation because it has been shown to be the preferable procedure when averaging correlations (Silver & Dunlap, 1987).

Results

CFAs and ESEM

The original simple structure models underlying both instruments exhibited an unacceptable model fit when conducting CFA (NEO PI-R: SB-χ^2 = 3493.44, df = 395, p < .001, SRMR = .13, RMSEA = .11, CFI = .61; 16PF: SB-χ^2 = 669.94, df = 74, p < .001, SRMR = .08, RMSEA = .11, CFI = .76). Altogether, 42 modifications (29 released paths, 13 residual covariances) to the NEO PI-R were required to obtain an acceptable model fit for Model 1 (SB-χ^2 = 1116.85, df = 353, p < .001, SRMR = .06, RMSEA = .06, CFI = .90). For the 16PF, fewer modifications (six released paths, six residual covariances) needed to be included until an acceptable model fit was achieved for Model 1 (SB-χ^2 = 197.98, df =62, p < .001, SRMR = .04, RMSEA = .06, CFI = .95). When applying the more conservative criteria (MI > 10 and EPC > .316) to derive Model 2, 12 and five paths were added to the NEO PI-R and the 16PF, respectively, until no fixed parameter fulfilled the a priori criteria for being freely estimated. Neither of the final two models achieved an acceptable model fit (NEO PI-R: SB-χ^2 = 2167.78, df = 383, p < .001, SRMR = .09, RMSEA = .09, CFI = .77; 16PF: SB-χ^2 = 391.29, df = 69, p < .001, SRMR = .05, RMSEA = .09, CFI = .87).

In Table 1 and Table 2, we provide an overview of the contributing subscales for the original and the two CFA modified models of both instruments. Many, but not all, of the subscales added to the domains in Models 1 and 2 were logical. For example, it is plausible to assign Dominance to the 16PF factor of Extraversion. It is less intuitive, however, to know how Abstractedness is related to the 16PF factor of Anxiety. Similarly, it seems reasonable to add Warmth and Positive Emotion to the NEO domain of Agreeableness. The negative link between Aesthetics and the NEO domain of Extraversion, however, is hard to explain conceptually.

Table 1

Overview of contributing primary factors for the 16PF models

16PF global factor	**Contributing 16PF primary factors in the original model**	**Additional primary factors in Model 1**	**Additional primary factors in Model 2**
Anxiety	Vigilance (L), Apprehension (O), Tension (Q4), Emotional Stability (-C)	–	M, E
Extraversion	Warmth (A), Liveliness (F), Social Boldness (H), Privateness (-N), Self-Reliance (-Q2)	E	E
Tough-Mindedness	Warmth (-A), Sensitivity (-I), Abstractedness (-M), Openness to Change (-Q1)	Q4, F, L	C
Independence	Dominance (E), Vigilance (L), Social Boldness (H), Openness to Change (Q1)	-O	-O
Self-Control	Rule-Consciousness (G), Abstractedness (-M), Perfectionism (Q3), Liveliness (-F)	N	–

Note. '-' indicates a reversed loading of the primary factor onto the global factor.

Table 2

Overview of contributing facets for the NEO PI-R models

NEO PI-R Domain	Contributing NEO PI-R facets in the original model	Additional facets in Model 1	Additional facets in Model 2
Neuroticism	Anxiety (n1); Angry Hostility (n2); Depression (n3); Self Consciousness (n4); Impulsiveness (n5); Vulnerability (n6)	-c1, a5, -a1, -c5, -c6, o3	-c1, a5, -e3
Extraversion	Warmth (e1); Gregariousness (e2); Assertiveness (e3); Activity (e4); Excitement Seeking (e5); Positive Emotions (e6)	-a2, n5, c4, a3, a1, -o2, -c6, o3	a3, a1,c4
Openness	Fantasy (o1); Aesthetics (o2); Feelings (o3); Actions (o4); Ideas (o5); Values (o6)	a6	–
Agreeableness	Trust (a1); Straightforwardness (a2); Altruism (a3); Compliance (a4); Modesty (a5); Tender Mindedness (a6)	e1, e6, -n2, e2, c3, -n3, o3	e1, -n2, e6, e2
Conscientiousness	Competence (c1); Order (c2); Dutifulness (c3); Achievement Striving (c4); Self-Discipline (c5); Deliberation (c6)	e4, e3, -o1, -a1, -n5, a3, -n6	-n5, e4

Note. '-' indicates a reversed loading of the facet onto the domain

The application of ESEM provided better model fit due to the less restrictive assumptions (NEO PI-R: SB-χ^2 = 1231.49, df = 295, $p < .001$, SRMR = .03, RMSEA = .07, CFI = .90; 16PF: SB-χ^2 = 197.61, df = 40, $p < .001$, SRMR = .03, RMSEA = .08, CFI = .94). For the NEO PI-R, the facets displayed substantial loadings on their respective domain. Only very few higher non-target loadings were observed. For the 16PF, the subscales displayed substantial loadings on their respective factor. Only for Independence a less clear pattern emerged and some differences with regard to the assigned subscales according to the test manual were found.

Correlations between scores from the original model and the CFA and ESEM models

Pearson correlation coefficients between counterparts of the Big Five scores based on the original model and the CFA- and ESEM-based scores for the NEO PI-R and the 16PF are shown in Table 3. The Spearman coefficients were almost identical to the Pearson coefficients (maximum difference .03). As such high similarity was found, and in order to save journal space, the Spearman coefficients are not reported.

The Pearson coefficients of the original model scores with the modified CFA scores which were computed as instructed by the respective test manual but with the additional subscales suggested by CFA (M1m and M2m) were fairly high. They ranged from .82 to .99 for the NEO PI- R and from .85 to .97 for the 16PF. Only three coefficients were below .90. The Pearson coefficients of the original model scores with the CFA factors scores obtained from the modified models (M1c and M2c) were also fairly high for the NEO PI-R (.78 to .98). However, a reduced agreement was found for the 16PF (.52 to .96), with particularly low coefficients for Tough-Mindedness and Independence. A similar pattern emerged for the ESEM factor scores (EM): Fairly high Pearson coefficients with the original model scores were obtained for the NEO PI-R (.87 to .98). The agreement for the 16PF was in general lower (.62 to .97). The lowest coefficient was obtained for Independence, the factor which also

Table 3
Pearson correlation coefficients of the 16PF and NEO PI-R factors across models

NEO PI-R[a]	OM - M1m		OM - M2m		OM - RV	OM - M1c	OM - M2c	OM - EM
	r	r_r	r	r_r	r_{rv}	r	r	r
Neu.	.95	.90	.95	.94	.68	.95	.97	.98
Ext.	.93	.84	.95	.92	.63	.80	.78	.87
Open.	.99	.98	n.a.	n.a.	.67	.96	.98	.98
Agree.	.82	.80	.88	.88	.61	.92	.92	.90
Cons.	.93	.83	.97	.94	.60	.97	.98	.97
16PF[a]	**OM - M1m**		**OM - M2m**		**OM - RV**	**OM - M1c**	**OM - M2c**	**OM - EM**
	r	r_r	r	r_r	r_{rv}	r	r	r
Anx.	n.a.	n.a.	.90	.87	.59	-.93[b]	-.89[b]	-.93[b]
Ext.	.97	.97	.97	.97	.73	.90	.96	.97
T-M.	.85	.85	.95	.95	.65	-.70[b]	-.72[b]	-.79[b]
Ind.	.95	.94	.95	.93	.60	.60	.52	.62
S-C.	.94	.94	n.a.	n.a.	.65	.79	.89	.82

Note. N = 620; Neu. = Neuroticism, Ext. = Extraversion, Open. = Openness to Experience, Agree. = Agreeableness, Conc. = Conscientiousness, Anx. = Anxiety, T.-M. = Tough-Mindedness, Ind. = Independence, S-C. = Self-Control, OM = Original model, M1m/M2m = Model 1/Model 2 scores computed based on scoring rules from respective test manual and additional scales included as suggested by CFA, M1c/M2c = Model 1/Model 2 CFA factor scores, EM = Exploratory Structural Equation Modeling factor scores, RV = Random Variable Model, r = Pearson correlation coefficient, r_r = Pearson correlation coefficient with random variables added to M1 and M2, r_{rv} = Pearson correlation coefficient with maximum number of random variables added to the model, n.a. = not applicable as scores were the same for the two models.
[a] All correlations $p < .001$.
[b] Factor scores are reversed.

displayed the least clear pattern of subscale loadings in the ESEM solution.

When comparing the agreement of the original scores with both sets of CFA-based modified scores (M1c/M2c versus M1m/M2m), the M1m/M2m scores displayed a higher agreement with the original Big Five scores. The M1m and M2m scores are computed following the instructions in the manual, albeit with some subscales added as suggested by CFA. Thus, the score obtained from the respective modified model M1m and M2m still contains the four to six subscales and applies the same weighting to these subscales as in the original model. The correlation coefficient between scores from the original and the modified CFA model M1m and M2m is therefore always in large part a correlation with itself. The M1c and M2c scores of the modified models also share the four to six subscales with the original model. However, the weighting of these subscales in the score computation was based on the CFA factor loadings and hence may differ from what was applied in the original model scores, thus offering one explanation for the slightly lower agreement.

To gauge the impact of this shared variance of scores between the original and the modified CFA models M1m and M2m, we generated a set of random subscales with scores for each respondent. These new variables were specified to have means and standard deviations similar to the subscales of the two instruments and zero-correlations with each other and with the original subscales. Using scores from these random variables, we computed two matching sets of alternative broad factor scores for each individual based on the two alternative models for both questionnaires.

For the first set of alternative scores, we added the same number of random variables to the computation of each broad factor score as was added to obtain scores for the two modified models of each questionnaire. For example, based on the CFA, eight additional facets were assigned to the factor of Extraversion in the NEO PI-R Model 1. Thus, we added eight random variables when computing the NEO PI-R Extraversion scores for the random-variable Model 1. The correlation coefficients, r_r, between the original model scores and the scores for random-variable Model 1 and random-variable Model 2 were only marginally smaller or sometimes equal to the coefficients obtained when adding scales based on a CFA of the original data and applying the scoring rules in the respective test manual (see Table 3). Therefore, adding the same number of zero-correlated random variables to the original model creates just as much relative change in the original scores as does the addition of scales identified by the CFA to the broad construct.

For the second set of alternative scores, we added the maximum number of random variables to the computation of each broad factor score, considering the overall number of narrow scales in each instrument. Thus, 24 random variables were added when computing each of the five NEO scores, and 10 or 11 random variables were added when computing each of the five 16PF scores for these random-variable models. We then computed Pearson correlation coefficients, r_{rv}, between the original scores and the scores obtained from this random-variable procedure for both models across both questionnaires. These correlation coefficients, shown in Table 3, were considerably smaller, ranging from .60 to .68 for the NEO PI-R and from .59 to .73 for the 16PF.

MTMM

The results of the MTMM-analyses are shown in Tables 4 to 6. Overall, convergent validity was supported for the original model (see lower-left triangle in Table 4): four of the five MTHM coefficients are considerable larger than all heterotrait coefficients. Only the relationship between NEO PI-R Agreeableness and 16PF Independence is smaller than two of the HTMM coefficients. Furthermore, the discriminant validity of the instruments is supported because the HTMM coefficients for both instruments are generally smaller than the MTHM coefficients and larger than the HTHM coefficients.

The pattern of the four MTMM matrices obtained for the two modified CFA models across both approaches to score computation is less clear (see Table 5 and 6). Evidence for convergent validity is less convincing because the correlation coefficients between the 16PF and NEO PI-R counterparts are consistently lower. Furthermore, evidence for the discriminant validity for these four MTMM matrices is weak as indicated by high correlations between conceptually unrelated factors across all four MTMM matrices based on the modified CFA model scores. Out of 80 HTMM and HTHM coefficients, 19 and 18 coefficients for Model 1 and Model 2, respectively, exceed an absolute value of .40.

Table 4

Multitrait–multimethod correlation matrix of the original model and for the ESEM model

		NEO PI-R					16PF				
		Neu.	Ext.	Open.	Agree.	Cons.	Anx.	Ext.	T-M.	Ind.	S-C.
NEO PI-R	Neu.	(.94/.95)	**-.28**	**-.16**	**.04**	**-.32**	*.56*	*.00*	*-.14*	*-.52*	*-.07*
	Ext.	**-.30**	(.91/.94)	**.45**	**.00**	**.17**	*-.10*	*.59*	*.61*	*.33*	*.02*
	Open.	**-.05**	**.33**	(.92/.93)	**-.17**	**.00**	*.02*	*.05*	*.40*	*.22*	*-.49*
	Agree.	**-.21**	**.05**	**.04**	(.90/.92)	**-.03**	*-.32*	*.03*	*.08*	*-.53*	*.23*
	Cons.	**-.44**	**.20**	**-.13**	**.14**	(.91/.93)	*-.04*	*-.05*	*.01*	*.26*	*.50*
16PF	Anx.	.68	*-.31*	*-.14*	*-.24*	*-.21*	(.87/.87)	**-.05**	**-.12**	**-.13**	**-.15**
	Ext.	*-.09*	.66	*.22*	*.17*	*-.05*	**-.29**	(.91/.85)	**.57**	**.16**	**.15**
	T-M.	*-.05*	*-.23*	-.66	*-.10*	*.23*	**.04**	**-.41**	(.85/.90)	**.22**	**-.11**
	Ind.	*-.14*	*.47*	*.34*	-.35	*.12*	**-.08**	**.38**	**-.38**	(.84/.89)	**-.03**
	S-C.	*-.11*	*-.12*	*-.44*	*.24*	.57	**-.04**	**-.20**	**.49**	**-.22**	(.86/.88)

Note. N = 620; Neu. = Neuroticism, Ext. = Extraversion, Open. = Openness to Experience, Agree. = Agreeableness, Conc. = Conscientiousness, Anx. = Anxiety, T.-M. = Tough-Mindedness, Ind. = Independence, S-C. = Self-Control. Reliability coefficients are in parentheses (OM/EM); the monotrait-heteromethod correlations are underscored; the triangular heterotrait-monomethod matrices are in boldface; the square heterotrait-heteromethod matrices are in italics. Coefficients displayed in the lower-left triangle are based on the original model with scores computed based on scoring rules from the respective test manual. Coefficients displayed in the upper-right triangle are based on ESEM factor scores; Tough-Mindedness is reversed.
All correlations (absolute values) >.12 are $p < .001$, .09 - .11 are $p < .01$, .06 - .08 are $p < .05$, < .06 are *n.s.*

Table 5

Multitrait–multimethod correlation matrix of Model 1

		NEO PI-R					16PF				
		Neu.	Ext.	Open.	Agree.	Cons.	Anx.	Ext.	T-M.	Ind.	S-C.
NEO PI-R	Neu.	(.95/.95)	**-.43**	**-.20**	**.03**	**-.28**	-.71	*-.20*	*.18*	*-.33*	*-.19*
	Ext.	**-.17**	(.90/.95)	**.54**	**-.62**	**-.06**	*.26*	.34	*-.01*	*.35*	*-.24*
	Open.	**.04**	**.27**	(.92/.93)	**-.13**	**-.13**	*.09*	*.22*	.36	*.18*	*-.53*
	Agree.	**-.48**	**.45**	**.28**	(.94/.95)	**.20**	*.09*	*.15*	*.28*	-.47	*.25*
	Cons.	**-.66**	**.24**	**-.16**	**.31**	(.93/.93)	*.29*	*.05*	*-.22*	*.09*	.43
16PF	Anx.	.64	*-.21*	*-.14*	*-.46*	*-.29*	(.87/.88)	**.42**	**-.12**	**.19**	**.45**
	Ext.	*-.07*	.65	*.24*	*.36*	*.10*	**-.28**	(.91/.92)	**.56**	**-.45**	**.08**
	T-M.	*.07*	*-.10*	-.60	*-.33*	*.12*	**.38**	**-.27**	(.88/.84)	**-.67**	**-.45**
	Ind.	*-.28*	*.42*	*.28*	-.01	*.27*	**-.31**	**.53**	**-.17**	(.86/.85)	**-.18**
	S-C.	*-.20*	*-.27*	*-.45*	*.04*	.45	**.04**	**-.43**	**.41**	**-.27**	(.87/.86)

Note. *N* = 620; Neu. = Neuroticism, Ext. = Extraversion, Open. = Openness to Experience, Agree. = Agreeableness, Conc. = Conscientiousness, Anx. = Anxiety, T.-M. = Tough-Mindedness, Ind. = Independence, S-C. = Self-Control. Reliability coefficients are in parentheses (M1m/M1c); the monotrait-heteromethod correlations are underscored; the triangular heterotrait-monomethod matrices are in boldface; the square heterotrait-heteromethod matrices are in italics. Coefficients displayed in the lower-left triangle are based on scores computed based on scoring rules from the respective test manual and additional scales included as suggested by CFA. Coefficients displayed in the upper-right triangle are based on CFA factor scores; Anxiety and Tough-Mindedness are reversed.

All correlations (absolute values) >.12 are p < .001, .09 - .11 are p < .01, .06 - .08 are p < .05, < .06 are n.s.

Table 6

Multitrait–multimethod correlation matrix of Model 2

		NEO PI-R					16PF				
		Neu.	Ext.	Open.	Agree.	Cons.	Anx.	Ext.	T-M.	Ind.	S-C.
NEO PI-R	Neu.	(.94/.94)	**-.18**	**-.11**	**-.18**	**-.49**	*-.53*	*-.17*	*.31*	*-.42*	*.03*
	Ext.	**-.51**	(.92/.93)	**.44**	**-.69**	**-.07**	*.05*	*.40*	*-.03*	*.27*	*-.29*
	Open.	**-.11**	**.31**	(.92/.93)	**-.07**	**-.20**	*.22*	*.31*	*.41*	*-.04*	*-.57*
	Agree.	**-.29**	**.59**	**.16**	(.93/.94)	**.25**	*.39*	*.11*	*.28*	*-.42*	*.25*
	Cons.	**-.57**	**.43**	**-.09**	**.20**	(.92/.92)	*.15*	*.00*	*-.31*	*.18*	*.44*
16PF	Anx.	.52	*-.24*	*.07*	*-.48*	*-.28*	(.88/.88)	**.61**	**.24**	**-.11**	**-.08**
	Ext.	*-.19*	.63	*.23*	*.35*	*.04*	**-.10**	(.91/.91)	**.56**	**-.39**	**-.07**
	T-M.	*-.22*	*-.08*	-.60	*-.08*	*.30*	**-.38**	**-.30**	(.85/.81)	**-.84**	**-.31**
	Ind.	*-.45*	*.41*	*.32*	-.11	*.20*	**.04**	**.53**	**-.20**	(.86/.84)	**-.18**
	S-C.	*-.09*	*.00*	*-.44*	*.14*	.51	**-.27**	**-.20**	**.51**	**-.21**	(.86/.86)

Note. *N* = 620; Neu. = Neuroticism, Ext. = Extraversion, Open. = Openness to Experience, Agree. = Agreeableness, Conc. = Conscientiousness, Anx. = Anxiety, T.-M. = Tough-Mindedness, Ind. = Independence, S-C. = Self-Control. Reliability coefficients are in parentheses (M2m/M2c); the monotrait-heteromethod correlations are underscored; the triangular heterotrait-monomethod matrices are in boldface; the square heterotrait-heteromethod matrices are in italics. Coefficients displayed in the lower-left triangle are based on scores computed based on scoring rules from the respective test manual and additional scales included as suggested by CFA. Coefficients displayed in the upper-right triangle are based on CFA factor scores; Anxiety and Tough-Mindedness are reversed.

All correlations (absolute values) >.12 are p < .001, .09 - .11 are p < .01, .06 - .08 are p < .05, < .06 are n.s.

An examination of the MTMM matrix based on the ESEM (see upper-right triangle in Table 4) showed that the convergent validity of the ESEM model was supported because the MTHM coefficients were of consistently high magnitude. More noteworthy however was the discriminant validity of the instruments assessed by the HTMM and the HTHM coefficients: Considerably lower correlations between conceptually unrelated Big Five factors based on the ESEM scores were obtained than based on the original model, particularly for the 16PF.

No absolute rules are available as to what can be considered sufficient evidence of construct validity based on MTMM results (Bagozzi & Yi, 1991). Instead, the pattern of correlation coefficients should be judged to assess the instrument's construct validity. To judge whether the pattern of one model provides a stronger support of construct validity than another, we computed mean values for the five MTMM matrices separately using Fisher's transformation.

We found consistently weaker support of construct validity in the four MTMM matrices based on the modified CFA models in comparison to the original model. First, the mean values of the MTHM matrices are considerably smaller in the modified CFA models (.50, .49, .49, and .44 for M1m, M1c, M2m, and M2c respectively, compared to .59 in the original model), indicating a decline in convergent validity for the modified models. Second, a mean increase in the HTMM matrices of the modified CFA models was observed for the 16PF (.31, .37, .28, and .38 for M1m, M1c, M2m, and M2c respectively, compared to .26 in the original model), and more pronounced for the NEO PI-R (.32, .27, .34, and .28 for M1m, M1c, M2m, and M2c respectively, compared to .19 in the original model). The reduced differentiation between non-matching traits is caused by several subscales that now contribute to more than one factor, creating not only a conceptual overlap but also shared variance that leads to increased correlations among broad domains. Finally, there is also a mean increase for the HTHM matrices, albeit only marginal (.24, .21, .23, and .22 for M1m, M1c, M2m, and M2c respectively, compared to .20 for the original model). Overall, a considerable decline in convergent and discriminant validity compared to the original was obtained for all modified CFA models.

The results based on ESEM display a less consistent pattern and are therefore discussed separately. Compared to the original model, the convergent validity of the ESEM model was slightly reduced as indicated by a mean value of .52 for the MTHM matrix. In fact, four of the five coefficients based on ESEM factor scores were considerably smaller than in the original model. Interestingly, the Big Five factor Agreeableness/Independence, which typically displays the least agreement across both instruments, was found to be more similar when using ESEM factor scores (MTHM correlation coefficient of -.53, compared to -.35 in the original model). Particularly remarkable however is the improved discriminant validity of the instruments when using ESEM factor scores. A mean decrease in the HTMM matrices of the ESEM model was observed for the NEO PI-R (.17, compared to .19 in the original model), and more pronounced for the 16PF (.18, compared to .26 in the original model). A slight mean decrease was also observed for the HTHM matrix (.19, compared to .20 in the original model).

Discussion

CFA and ESEM were applied to two personality instruments based on the Big Five framework to determine the impact the factor structure suggested by these factor-analytical methods had on relative scores as well as on the construct validity of the NEO PI-R and the 16PF. MTMM analyses based on the Big Five scores of the CFA models revealed a considerable decrease in the convergent and discriminant validity of the questionnaires. Results based on ESEM were more promising in that the discriminant validity was improved in comparison to the original model. However, with the exception of Agreeableness/Independence, a considerable decrease in the convergent validity was observed.

The results – particularly those based on CFA models – highlight some important issues. Several additional links between subscales and factors were suggested by CFA, indicating that the imposed simple internal structure may not be an adequate description of the construct personality. Introducing these additional links may indeed result in a model that better reflects the internal structure of personality. It has been argued before that the five factors are not as distinct as often suggested. In fact, even Costa and McCrae (1992b) acknowledge that some secondary loadings are “appropriate and meaningful”

(p. 45), such as a high negative loading of the Neuroticism facet Angry Hostility on the domain Agreeableness. In the interest of simplicity and interpretability a decision was made to assign each subscale to one factor only and to exclude any additional relationships with other factors.

However, retaining the simple structure seems to be not only advisable in order to ensure the measures' interpretability. The present study shows that deciding against a more complex structure also avoided a negative impact on their convergent and discriminant validity. While introducing additional links in the models has led to an increase of internal validity by better reflecting the complex relationships between subscales and higher-order factors, this improvement was achieved at the expense of the instruments' convergent and discriminant validity which is not desirable. First, a decline in convergent validity resulted in a decreased consensus on the five broad personality factors. This impedes comparisons of research findings on personality conducted using different measures and will make it more difficult to combine them in meta-analyses. Thus, the more complex structure jeopardizes the benefit of a five-factor framework. Second, a decline in discriminant validity as indicated by higher intercorrelations showed that the broad factors are conceptually less differentiated and hence might be less useful in applied settings. This tradeoff between the instrument's capability to adequately represent the complex internal structure of personality while preserving its convergent and discriminant validity cannot easily be resolved.

Furthermore, while there are good theoretical reasons to question the proposed simple structure, CFA should not be the method of choice to determine a more appropriate representation of the internal structure of personality. First, different assignments of scales to factors were obtained depending on the modification criteria and cut-off criteria applied in CFA, especially for the 16PF. Second, some of the additional links identified in CFA may not reflect conceptual relationships but are method artifacts, due to response styles such as social desirability (Ziegler & Buehner, 2009) or particular item content, such as negatively phrased items (Biderman, Nguyen, Cunningham, & Ghorbani, 2011). While these effects were not examined in the present study, it is important to remember that they may provide an explanation for some of the relationships found between subscales.

A high agreement between the original scores and the modified scores computed following the respective test manual (M1m and M2m) were obtained. Including an additional path in a CFA model when the respective subscale displays a high loading on a factor results in a small relative change in an individual's score. This is because very little additional variance is added. At the same time, the conceptual benefit is questionable because the constructs reflected by these modified composite scores become increasingly complex and less distinct with respect to their conceptual meaning.

Furthermore, by adding the same number of random scales as had been performed in the modified models, the magnitude of relative change was approximately the same as what had been obtained when adding scales suggested by the CFA, with the exception of Extraversion and Conscientiousness of the NEO PI-R. As the additional subscales were specified to be unrelated in this simulation, they quantify the maximum relative change that may occur in such an instance, regardless of which subscales may be assigned. This is particularly informative because the assignment of subscales to factors based on the CFA has been shown to depend on the decision criteria applied during the process of model modification. In the second simulation, the maximum number of subscales was added to each broad domain. It shows the maximum relative score change if one were to add all remaining subscales to each factor. While this may not present a realistic scenario, it offers a benchmark against which the observed differences between the original and the modified model can be judged.

A reduced agreement between the original scores with the CFA factor scores based on the modified CFA models (M1c and M2c) and ESEM scores were obtained, particularly for the 16PF. The fact that the original Big Five scores based on conventional scoring yield different results from applying CFA and ESEM factor scores has important implications for research and practice using personality questionnaires. First, a research study may yield different results depending on how the Big Five scores were obtained. Second, scores based on the modified CFA models and ESEM models are conceptually different constructs because their conceptual meaning is determined by the specific combination of contributing subscales. As such, potentially different findings between studies are not only likely but also plausible as analyses will be based on personality factors that do not share the same conceptual meaning. Third, regarding the applicability of research findings based on CFA and EFA factor scores, caution needs to be exercised as these results may not be directly transferable to practical applications where conventional scoring is used.

Personality questionnaires have repeatedly exhibited good criterion-related validity (e.g., Grucza & Goldberg, 2007; Hurtz & Donovan, 2000). The poor support of their internal structure has raised the question of how these measures can predict external criteria. However, the simple structure may in fact be beneficial for the measures' predictive capabilities. Several studies have demonstrated that broader domains reduce the predictive power of personality (Dudley, Orvis, Lebiecki, & Cortina, 2006; Tett, Steele, & Beauregard, 2003). In addition, from a conceptual and practical viewpoint, using these more complex structures seems to be less useful because it is harder to interpret relationships between broader domains and external criteria. This study provides reasoning for the continued use of current personality instruments that have demonstrated criterion-related validity despite CFA findings that suggest a more complex structure.

The results also refute potential concerns regarding the validity and applicability of previous research based on current personality instruments that has been raised when the inventories failed to be supported by CFA. More importantly, the decrease in construct validity when applying the more complex structure of personality proves that retaining the simple structure of the current questionnaires is not only a defensible option, but may even be favorable.

Limitations

In our study, models were specified that reflect the proposed structure according to the respective test manuals and the current typical applications of CFA. Other modeling approaches have been suggested, such as circumplex models (Fabrigar, Visser, & Browne, 1997), and bifactor models that incorporate either method factors (Biderman, et al., 2011) or a general factor (Chen, Hayes, Carver, Laurenceau, & Zhang, 2012). These may overcome some issues related to the application of more conventional CFA models to personality data. Their application should be encouraged as they may also provide different views on the internal structure of personality.

The model modifications were based on a data-driven approach. Adding only conceptually sound links between facets and domains may have led to different models that are easier to interpret. It is questionable, however, whether such an arbitrary approach to utilizing CFA results can be defended and is superior to an exclusively conceptual approach to theory development. In any case, it would have led to even fewer additional subscales per broad domain, thus resulting in even smaller relative score changes.

The study did not examine the impact the structures proposed by CFA and ESEM have on criterion-related validity. However, given that the instruments have shown to be less construct-valid, an examination of their criterion-related validity seems not indicated as construct validity should be a requisite before proceeding to this next question.

Recommendations and Conclusions

Considering the limitations and ambiguities regarding the results obtained from the CFA, one should not dismiss current measures of personality and question their construct validity merely based on the poor fit based on this analytical method. Furthermore, it may be ill-advised to reject personality theory based on CFA results. Theories are designed to explain phenomena and need to simplify the more complex relationships observed between constructs in the real world. Meehl (1990) argues that models can be useful even if they simplify reality. One may add that models *need* to simplify reality so that they can be useful. While current personality measures are not without flaws and do not fulfill the model fit criteria proposed for CFA applications, their continued use seems justified as they have demonstrated good criterion-related validity. This study also shows that increasing the measures' complexity to comply with CFA standards and improved their internal validity led to a reduced convergent and discriminant validity, suggesting that there is a trade-off between these two aspects of construct validity.

Our results based on ESEM were more promising with regard to the findings on the instruments' construct validity, particularly regarding their discriminant validity. ESEM also offers multi-group analyses and longitudinal analyses, both with tests for measurement invariance (Asparouhov & Muthen, 2009). It hence enables the application of sophisticated methods typically associated with the CFA/structural equation modeling framework but without requiring the instrument to fulfill the more stringent CFA criteria. We believe it to be a useful tool in developing and evaluating self-report questionnaires assessing personality and encourage its application.

Acknowledgments. We would like to thank Lew Goldberg and Maureen Barckley for making the Eugene Springfield Community Sample data available for this study.

References

Aluja, A., Blanch, A., & Garcia, L. F. (2005). Reanalyzing the 16PF-5 second order structure: Exploratory versus confirmatory factorial analysis. *European Journal of Psychology of Education, 20*(4), 343-353. doi: 10.1007/bf03173561

Aluja, A., Rossier, J., Garcia, L. F., & Verardi, S. (2005). The 16PF5 and the NEO-PI-R in Spanish and Swiss Samples: A Cross-Cultural Comparison. *Journal of Individual Differences, 26*(2), 53-62. doi: 10.1027/1614-0001.26.2.53

Asparouhov, T., & Muthen, B. (2009). Exploratory Structural Equation Modeling. *Structural Equation Modeling, 16*(3), 397-438. doi: 10.1080/10705510903008204

Bagozzi, R. P., & Yi, Y. (1991). Multitrait-Multimethod Matrices in Consumer Research. *Journal of Consumer Research, 17*(4), 426-439.

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81*(3), 261-272. doi: 10.1037/0021-9010.81.3.261

Beauducel, A., & Wittmann, W. W. (2005). Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. *Structural Equation Modeling, 12*(1), 41-75. doi: 10.1207/s15328007sem1201_3

Biderman, M. D., Nguyen, N. T., Cunningham, C. J. L., & Ghorbani, N. (2011). The ubiquity of common method variance: The case of the Big Five. *Journal of Research in Personality, 45*(5), 417-429. doi: 10.1016/j.jrp.2011.05.001

Bobko, P., Roth, P. L., & Buster, M. A. (2007). The Usefulness of Unit Weights in Creating Composite Scores. *Organizational Research Methods, 10*(4), 689-709. doi: 10.1177/1094428106294734

Borkenau, P., & Ostendorf, F. (1990). Comparing exploratory and confirmatory factor analysis: A study on the 5-factor model of personality. *Personality and Individual Differences, 11*(5), 515-524. doi: 10.1016/0191-8869(90)90065-y

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin, 56*(2), 81. doi: 10.1037/h0046016

Cattell, H. E. P., & Mead, A. D. (2008). The Sixteen Personality Factor Questionnaire (16PF®). In G. J. Boyle, G. Matthews & D. H. Saklofske

(Eds.), *The Sage handbook of personality theory and assessment* (pp. 135-159). Los Angeles Calif. u.a.: SAGE.

Cattell, R. B., & Cattell, H. E. P. (1995). Personality Structure and the New Fifth Edition of the 16PF. *Educational and Psychological Measurement, 55*(6), 926-937. doi: 10.1177/0013164495055006002

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P., & Zhang, Z. (2012). Modeling General and Specific Variance in Multifaceted Constructs: A Comparison of the Bifactor Model to Other Approaches. *Journal of Personality, 80*(1), 219-251. doi: 10.1111/j.1467-6494.2011.00739.x

Church, A. T., & Burke, P. J. (1994). Exploratory and confirmatory tests of the Big Five and Tellegen's three- and four-dimensional models. *Journal of Personality and Social Psychology, 66*(1), 93-114. doi: 10.1037/0022-3514.66.1.93

Conn, S. R., & Rieke, M. L. (1994). *The 16PF fifth edition technical manual*: Institute for Personality and Ability Testing Champaign, IL.

Costa, P. T., & McCrae, R. R. (1992a). 'Four ways five factors are not basic': Reply. *Personality and Individual Differences, 13*(8), 861-865. doi: 10.1016/0191-8869(92)90002-7

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

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417-440. doi: 10.1146/annurev.ps.41.020190.002221

Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A Meta-Analytic Investigation of Conscientiousness in the Prediction of Job Performance: Examining the Intercorrelations and the Incremental Validity of Narrow Traits. *Journal of Applied Psychology, 91*(1), 40-57. doi: 10.1037/0021-9010.91.1.40

Fabrigar, L. R., Visser, P. S., & Browne, M. W. (1997). Conceptual and Methodological Issues in Testing the Circumplex Structure of Data in Personality and Social Psychology. *Personality and Social Psychology Review, 1*(3), 184-203. doi: 10.1207/s15327957pspr0103_1

Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research, 42*(3), 509-529. doi: 10.1080/00273170701382864

Gerbing, D. W., & Anderson, J. C. (1984). On the Meaning of Within-Factor Correlated Measurement Errors. *Journal of Consumer Research, 11*(1), 572-580.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216-1229. doi: 10.1037/0022-3514.59.6.1216

Goldberg, L. R., Sweeney, D., Merenda, P. F., & Hughes, J. E. (1998). Demographic variables and personality: the effects of gender, age, education, and ethnic/racial status on self-descriptions of personality attributes. *Personality and Individual Differences, 24*(3), 393-403. doi: 10.1016/s0191-8869(97)00110-4

Grucza, R. A., & Goldberg, L. R. (2007). The Comparative Validity of 11 Modern Personality Inventories: Predictions of Behavioral Acts, Informant Reports, and Clinical Indicators. *Journal of Personality Assessment, 89*(2), 167-187. doi: 10.1080/00223890701468568

Hopwood, C. J., & Donnellan, M. B. (2010). How Should the Internal Structure of Personality Inventories Be Evaluated? *Personality and Social Psychology Review, 14*(3), 332-346. doi: 10.1177/1088868310361240

Hu, L.-t., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*(4), 424-453. doi: 10.1037/1082-989x.3.4.424

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55. doi: 10.1080/10705519909540118

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology, 85*(6), 869-879. doi: 10.1037/0021-9010.85.6.869

Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology, 52*(3), 621-652. doi: 10.1111/j.1744-6570.1999.tb00174.x

Kline, P. (1994). *An easy guide to factor analysis*. London: Routledge.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111*(3), 490-504. doi: 10.1037/0033-2909.111.3.490

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. *Structural Equation Modeling, 11*(3), 320-341. doi: 10.1207/s15328007sem1103_2

Marsh, H. W., Liem, G. A. D., Martin, A. J., Morin, A. J. S., & Nagengast, B. (2011). Methodological Measurement Fruitfulness of Exploratory Structural Equation Modeling (ESEM): New Approaches to Key Substantive Issues in Motivation and Engagement. *Journal of Psychoeducational Assessment, 29*(4), 322-346. doi: 10.1177/0734282911406657

McCrae, R. R., & Costa, P. T. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 139-153). New York: Guilford Press.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*(2), 175-215. doi: 10.1111/j.1467-6494.1992.tb00970.x

McCrae, R. R., Zonderman, A. B., Costa, P. T., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the Revised NEO Personality Inventory: Confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology, 70*(3), 552-566. doi: 10.1037/0022-3514.70.3.552

Meehl, P. E. (1990). Appraising and Amending Theories: The Strategy of Lakatosian Defense and Two Principles that Warrant It. *Psychological Inquiry, 1*(2), 108-141. doi: 10.1207/s15327965pli0102_1

Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality, 41*(6), 1213-1233. doi: 10.1016/j.jrp.2007.02.003

Muthén, L., & Muthén, B. (2012). Mplus user's guide Version 6.11. Los Angeles, CA: Muthen & Muthen.

Narayanan, A. (2012). A Review of Eight Software Packages for Structural Equation Modeling. *The American Statistician, 66*(2), 129-138. doi: 10.1080/00031305.2012.708641

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology, 66*(6), 574. doi: 10.1037/h0040291

Rosseel, Y. (2011). *lavaan: an R package for structural equation modeling and more, version 0.4-11*: Available at http://lavaan.ugent.be (accessed January 9, 2012).

Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology, 76*(3), 323-346. doi: 10.1348/096317903769647201

Saris, W. E., Satorra, A., & van der Veld, W. M. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling, 16*(4), 561-582. doi: 10.1080/10705510903203433

Satorra, A., & Bentler, P. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*, 507-514. doi: 10.1007/bf02296192

Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology; Journal of Personality and Social Psychology, 94*(1), 168. doi: 10.1037/0022-3514.94.1.168

Silver, N. C., & Dunlap, W. P. (1987). Averaging correlation coefficients: Should Fisher's z transformation be used? *Journal of Applied Psychology, 72*(1), 146.

Terracciano, A., McCrae, R. R., & Costa, P. T. (2010). Intra-individual change in personality stability and age. *Journal of Research in Personality, 44*(1), 31-37. doi: 10.1016/j.jrp.2009.09.006

Tett, R. P., Steele, J. R., & Beauregard, R. S. (2003). Broad and Narrow Measures on Both Sides of the Personality-Job Performance Relationship. *Journal of Organizational Behavior, 24*(3), 335-356. doi: 10.1002/job.191

Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.

van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The General Factor of Personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality, 44*(3), 315-327. doi: 10.1016/j.jrp.2010.03.003

Vassend, O., & Skrondal, A. (2011). The NEO personality inventory revised (NEO-PI-R): Exploring the measurement structure and variants of the five-factor model. *Personality and Individual Differences, 50*(8), 1300-1304. doi: 10.1016/j.paid.2011.03.002

Ziegler, M., & Buehner, M. (2009). Modeling Socially Desirable Responding and Its Effects. *Educational and Psychological Measurement, 69*(4), 548-565. doi: 10.1177/0013164408324469

II. Paper "Calling and Career Preparation: Investigating Developmental Patterns and Temporal Precedence"

Abstract

The presence of a calling and career development are assumed to be closely related. However, the nature of and reason for this relationship has not been thoroughly investigated. We hypothesized the existence of reciprocal effects between calling and three dimensions of career preparation and assessed the change of the presence of a calling, career planning, decidedness, and self-efficacy with three waves of a diverse sample of German university students (N = 846) over one year. Latent growth analyses revealed that the intercepts of calling showed a significant positive correlation with the intercepts of all career preparation measures. The slope of calling was positively related to those of decidedness and self-efficacy but not to planning. Cross-lagged analyses showed that calling predicted a subsequent increase in planning and self-efficacy. Planning and decidedness predicted an increase in the presence of a calling. The results suggest that calling and career preparation are related due to mutual effects but that effects differ for different career preparation dimensions.

Complete reference: Hirschi, A. & Herrmann, A. (2013). Calling and career preparation: Investigating developmental patterns and temporal precedence. Journal of Vocational Behavior, 83 (1), 51-60. doi: 10.1016/j.jvb.2013.02.008

Introduction

In today's post-industrialized economies, many people seek intrinsically motivating work. Empirical research (e.g., Wrzesniewski, McCauley, Rozin, & Schwartz, 1997) shows that a considerable number of individuals in various professions are searching for or trying to implement a calling in their career. Dobrow and Tosti-Kharas (2011, p. 1003) defined calling as a consuming, meaningful passion people experience toward a career domain while Dik and Duffy (2009) described it as a transcendent summons to a meaningful career that is used to serve others. Regardless of the specific definition, the presence of a calling is often described as a psychological resource that promotes vocational development and connected to identity, confidence, resilience, and adaptability (Hall & Chandler, 2005). Empirical studies confirmed a positive relationship of calling and several career development variables such as career decidedness (Duffy & Sedlacek, 2007; Steger, Pickering, Shin, & Dik, 2010) or career self-efficacy beliefs (Dobrow & Tosti-Kharas, 2011; Duffy, Allan, & Dik, 2011; Hirschi, 2011). However, the reason for and nature of the relationship has not been clearly addressed. As a consequence, we do not know whether callings promote, hinder, precede, follow, or are reciprocally related to pivotal career development constructs. However, such knowledge is crucial to increase our understanding of how a calling emerges and how it affects career development (Dobrow, in press).

The major general contribution of the present study is that it is the first study to our knowledge to investigate the developmental intersection between calling and career development variables with a true longitudinal design encompassing three measurement points – a feature generally very rare in career and organizational research. Specifically, the present study examines the relation of change trajectories of the presence of a calling and three dimensions of career preparation (Skorikov, 2007): career planning, decision-making, and confidence among university students with three measurement points over a period of one year. Moreover, we investigate to what extent the presence of a calling precedes and/or follows the development of the career preparation dimensions with a cross-lagged study. In this way, the study makes three key contributions. First, we contribute to the calling literature by investigating how callings change over time and what factors affect such changes. Second, we extend this literature by linking calling with three dimension of career preparation among university students and show if and how calling affects those pivotal career development variables. Third, we contribute to career development research by demonstrating how career preparation affects the emergence of a presence of a calling.

Theoretical Background

Dobrow (in press) suggested that researchers must consider a calling to be a dynamic phenomenon that changes over time and addressed the need for research examining calling in conjunction with possible antecedents and outcomes. Longitudinal research investigating the relation of calling and career development variables has only begun to emerge (Dobrow, in press; Dobrow & Tosti-Kharas, 2011; Duffy, Manuel, Borges, & Bott, 2011) and generally reported positive relations. However, these studies have not tested lagged effects, which would establish whether a change in career development variables is related to a change in calling (or vice versa) and thus make a case for their mutual influence above and beyond mere concurrent relations. The present study extends existing research and attempts to increase our understanding of how calling and dimensions of career preparation are related over time.

Specifically, we investigated the intersection of calling and three dimensions of career development that represent three commitment-focused dimensions of career adaptability (Savickas, 2005) that Skorikov (2007; Stringer, Kerpelman, & Skorikov, 2011) defined as career preparation: Planning, decision-making, and confidence. Previous research showed that those dimensions are pivotal constructs of career development and related to important career outcomes such as, for example, fewer career concerns among first year university students (Creed, Fallon, & Hood, 2009) or better psychological adjustment after high school (Skorikov, 2007). In our study, we investigated the relation of those career preparation variables with calling among a large and diverse group of German university students. This allowed us to tap into a critical period in terms of career and identity development. First, engaging in career preparation is particularly pivotal for students to address the specific career task of transitioning from university to work or postgraduate degrees, which is characterized by the need for active career planning, decision-making, addressing uncertainty about future work, an active job search, and career self-management(see Abele & Spurk, 2009, for a study with German university graduates). Second, during the period of emerging adulthood, which encompasses the university years, the development and establishment of a student's identity, values, goals, and life structures are particularly prevalent (Arnett, 2000). Therefore, investigating the emergence of a presence of a calling appears timely for this population. Empirical research confirmed that the concept of a calling is relevant for a considerable number of university students in the US (Hunter, Dik, & Banning, 2010) and Germany (Hagmaier & Abele, 2012; Hirschi, 2011).

Because university students are concerned with career preparation and the notion of a calling is important for a considerable number among them, investigating the developmental intersection of career preparation and calling seems important and fruitful to increase our understanding of the nature, antecedents, and consequences of callings. A calling is frequently considered as a psychological resource that positively affects career development (Hall & Chandler, 2005) and empirical research showed positive relations to different career development variables (Dobrow & Tosti-Kharas, 2011; Duffy, Allan, et al., 2011; Duffy & Sedlacek, 2007; Hirschi, 2011; Steger, et al., 2010). It is hence reasonable to expect meaningful positive relations with dimensions of career preparation among university students. However, previous research has not investigated the developmental intersection of those career constructs.

In the next paragraphs, we review the relation of calling with the three dimensions of career preparation in terms of career planning, career decision making (i.e., career decidedness), and career confidence (i.e., career self-efficacy believes). All three dimensions can be conceptualized as positive indicators of career preparation (Skorikov, 2007) and we hence expect no fundamental differences in their relation to the presence of calling. However, empirical research (Stringer, et al., 2011) suggested that they show different antecedents, developmental patterns, and outcomes which makes it important to treat them as distinct dimensions in their own right.

Career planning and calling

People with a sense of calling should be motivated to proactively consider and plan their career because they are likely to aim to implement their calling at work, which allows them to live their calling and achieve higher job satisfaction (Duffy, Bott, Allan, Torrey, & Dik, 2012). As such, callings can represent an ideal possible (future) work self that motivates anticipatory and future-oriented career behavior, such as career planning (Strauss, Griffin, & Parker, 2012). On the other hand, it is also possible that active career planning facilitates the development and confirmation of one's calling. Career planning entails future-oriented thinking and envisioning future work states (Savickas, 1997). Thus, it allows people to envision themselves in different future work contexts and to construct a possible future self (Markus & Nurius, 1986), which is important in identity construction and finding meaning. This future-oriented identity construction can lead to the discovery or confirmation of a sense of calling. Therefore, one could expect a concurrent as well as reciprocal relationship over time between the presence of a calling and career planning,

Hypothesis 1: Career planning and the presence of a calling are positively related to each other (a) within and (b) across time; (c) more career planning will predict an increase in the presence of a calling; (d) a stronger presence of a calling will predict an increase in career planning.

Career decidedness and calling

Based on theoretical and empirical grounds, we can expect a close positive relation of career decidedness and the presence of a calling (Duffy & Sedlacek, 2007; Hirschi, 2011; Steger, et al., 2010). Theoretically, this relationship can be explained in the way in which a calling gives people a sense of direction in their career because it entails a certain vocational path toward which one feels called (Dik & Duffy, 2009). As such, the presence of a calling facilitates a career choice that implements one's self-concept into the work role (Duffy & Sedlacek, 2010). Second, calling is related to self-clarity (Duffy & Sedlacek, 2007), which is in turn an important prerequisite for career decision making and career decidedness (Super, 1990). Therefore, a calling can be assumed to enhance decidedness because it facilities career decision making by providing clarity regarding oneself and one's goals. However, having a clear perspective of one's career in terms of career decidedness could reinforce a sense of direction, control, meaning, and purpose in one's career (Savickas, 2005), which could strengthen or develop the presence of a calling. Therefore, we can assume that career decidedness and calling reinforce each other over time.

Hypothesis 2: Career decidedness and the presence of a calling are positively related to each other (a) within and (b) across time; (c) more career decidedness will predict an increase in the presence of a calling; (d) a stronger presence of a calling will predict an increase in career decidedness.

Career self-efficacy and calling

People with a sense of calling are assumed to express their strengths through their calling and as such should possess high confidence in their ability to master career-related tasks (Hall & Chandler, 2005). Similarly, individuals with a sense of calling are enacting their "true selves" in the work role, which entails expressing their core strengths (Peterson, Park, Hall, & Seligman, 2009). Therefore, although the presence of a calling may not directly translate into a specific level of a person's objective ability (Dobrow, in press), we could assume that a calling can promote a sense of career self-efficacy. Supporting this assumption, research on college students has found that those with a sense of calling reported on average more career decision making self-efficacy (Duffy, Allan, et al., 2011), that career self-efficacy was a defining component across different types of callings (Hirschi, 2011), and that a calling predicted career self-efficacy even several years later (Dobrow & Tosti-Kharas, 2011). At the same time, a sense of efficacy in mastering work- and career-related tasks could in turn facilitate the development of the presence of a calling because a sense of competence is essential to develop intrinsic motivation and self-determination in a given domain (Deci & Ryan, 2000), both important components of the presence of a calling (Dik & Duffy, 2009; Hall & Chandler, 2005). Hall and Chandler (2005) also stated that finding one's calling can lead to a success cycle where positive career experiences that emerge out of one's calling reinforce the person's self-confidence.

Hypothesis 3: Career self-efficacy and presence of a calling are positively related to each other (a) within and (b) across time; (c) more career self-efficacy will predict an increase in the presence of a calling; (d) a stronger presence of a calling will predict an increase in career self-efficacy.

Materials and Method

Participants and Procedure

We used a panel design with refreshment sample (Deng, Hillygus, Reiter, Si, & Zheng, in press) to assess two groups of students across all majors enrolled at a medium-sized German university. Specifically, we collected three waves of data, each six months apart (T1 to T3). We chose a time lag of six months between the waves because we deemed this period to be sufficient to observe any meaningful change in the assessed career variables that might occur. Previous research successfully applied the same time lag when examining change in career constructs (e.g., Strauss, et al., 2012). Group one participated in all three waves. Group two was the refreshment sample consisting of new participants recruited six months after T1 and hence participating only in the last two waves (T2 and T3). This procedure hence assesses different groups of participants with temporally overlapping measurement points (T2 and T3) in order to assess common developmental trends. Data were collected with a web-based questionnaire and participation in a lottery drawing offering two prizes of EUR 450 each were offered as an incentive at each assessment point.

The *first group* of students were recruited by sending an email invitation to all students in the second semester of their second and third years of study (approx. 3,500 students), resulting in response levels of N = 1,207 and 34% (T1). Participating students were contacted again two times, each six months apart, resulting in response rates of 45% (T2) and 24% (T3), respectively, with 206 participating in both follow-ups. The *second group* consisted of students starting their second study year at T2 (approx.1,800 students) and were also invited by email, resulting in a response rate of N = 700 and 39%. Participants were again contacted six month later (T3) with a response rate of 30%.

One advantage of using a panel design with refreshment sample over a classical longitudinal panel design is that attrition from the first group can be compensated with a new random sample of participants (group two) (Hirano, Imbens, Ridder, & Rubin, 2001). We compared participants from the first group at T2 to those from the second group at T2 on the assessed variables. The results showed no significant differences on any of the assessed constructs, indicating no group effects and hence supporting all subsequent analyses being conducted with treated participants as one group. Due to design and individual attrition, not all students participated in all three measurement waves. The impact of "missingness" on the study was assessed by examining the relationship between the number of missing time points per participant and the other study variables. The results showed that missingness was not significantly correlated with any of the assessed variables. Because we did not find any indication of a systematic bias of missingness nor significant differences between the first group and the refreshment sample, all participants participating on at least two measurement points were retained for the final sample. For participants who did not provide data on one occasion, we estimated missing data with a full information maximum likelihood estimator of missing data. This procedure was shown to yield very accurate parameter estimates and has been particularly recommended for longitudinal studies where missing data is common (Graham, 2009). In fact, is has been shown to be the preferable approach as it leads to less biased results in comparison to listwise deletion where only participants with complete data on all measurement occasions are retained (Duncan & Duncan, 1995).

The final sample consisted of 846 students, 64% were female, and the mean age was 23.73 years, *SD* = 2.40, at the first time of study participation. Participants enrolled in 33 different majors, ranging from mechanical engineering to social work, with the largest groups stemming from Management and Entrepreneurship (16 %), Business Psychology (16%), Business Administration (14 %), Environmental Science (7 %), and Business Law (5 %). As is customary in Germany, race was not assessed.

Measures

Unless stated otherwise, all measures used a five-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*)*.* Cronbach's alpha estimates, means, standard deviations, and correlations between measures are reported in Table 1.

Presence of calling

The German language version (Hirschi, 2011) of the presence subscale of the brief calling scale (BCS; Dik, Eldridge, Steger, & Duffy, 2012) was applied. It consisted of two statements (''I have a calling to a particular kind of work'', and ''I have a good understanding of my calling as it applies to my career''). This measure has the advantage of not imposing a specific notion of calling on the study participants. A recent validation study (Dik, et al., 2012) found that the BCS scores showed moderate to strong correlations with scores of other measures of calling (r = .24 to .69) and with informants' reports of participants' perceptions of their calling (r = .27 to .46). Previous research using this scale reported high correlations between the two items (r = .76 to .82) and showed moderate to high relationships with career decision making self-efficacy, intrinsic work motivation, religious commitment, and meaning in life (Duffy & Sedlacek, 2007; Steger, et al., 2010).

Table 1

Reliability, Correlations, Means and Standard Deviations for Calling and Career Preparation Scales

		1	2	3	4	*M*	*SD*
Time 1 (N = 633)	1. Planning	*.88*	.73	.45	.46	3.28	0.84
	2. Decidedness		*.88*	.45	.48	3.51	0.88
	3. Self-efficacy			*.78*	.31	2.50	0.96
	4. Calling				*.72*	3.17	1.02
Time 2 (N = 760)	1. Planning	*.87*	.72	.38	.45	3.29	0.85
	2. Decidedness		*.89*	.43	.47	3.46	0.89
	3. Self-efficacy			*.81*	.32	2.64	0.97
	4. Calling				*.72*	3.11	0.98
Time 3 (N = 505)	1. Planning	*.87*	.74	.40	.45	3.29	0.86
	2. Decidedness		*.89*	.50	.46	3.47	0.91
	3. Self-efficacy			*.81*	.41	2.83	1.00
	4. Calling				*.73*	3.12	1.01

Note. Entries in *italic* in diagonal are the Cronbach's alpha reliability coefficients for the career preparation scales and the bivariate correlations of the two calling items respectively.

All correlations > .14 are $p < .001$, .06-.14 are $p < .01$

Career planning

Planning was assessed with the German six-item (e.g., “I have a strategy for reaching my career goals”) career planning scale proposed by Abele and Wiese (2008), adopted from respective scales from Gould (1979) and Wayne, Liden, Kraimer and Graf (1999). Abele and Wiese (2008) reported a reliability of α = .86 and support for the construct validity of the scale among a large group of university-educated German professionals in terms of medium relationships with subjective and objective career success.

Career decidedness

We applied the German-language adaptation of the vocational identity scale (Holland, Daiger, & Power, 1980; Jörin, Stoll, Bergmann, & Eder, 2004) using seven items (e.g., “I’m not sure yet which occupations I could perform successfully”). Research with the German-language version reported scale reliabilities between α = .81 and .89 and showed that the scale correlated highly with other measures of career decidedness, moderately with career planning, and low with career exploration among adolescents and college students (Hirschi, Niles, & Akos, 2011; Jörin Fux, 2006).

Career self-efficacy

We used the six-item (e.g., “Whatever comes my way in my job, I can usually handle it”) German short version of the occupational self-efficacy scale developed and validated by Rigotti, Schyns, and Mohr (2008) with a six-point Likert scale from 1 (*not at all true*) to 6 (*completely true*). Rigotti et al. (2008) reported a scale reliability of α = .84 and evidence for construct validity among a large group of German employees with moderate relationships to job satisfaction, organizational commitment, job performance, and job insecurity.

Analytical Approach

In order to test our hypotheses that calling and career preparation are related within and across time, Hypotheses 1 to 3 (a) and (b), we first applied Latent Growth Modeling (LGM), a statistical analysis that estimates growth trajectories of intraindividual change over time (for an introduction, see Martens & Haase, 2006). Specifically, we assessed whether, over the assessed three time points, the intercept (initial levels) and slope (intraindividual change trajectory) of calling were related to the intercepts and slopes of the career preparation measures.

In order to assess our hypotheses which suggest that calling and career preparation predict change in each other over time, Hypotheses 1 to 3 (c) and (d), we next applied cross-lagged analyses (CLA, see Martens & Haase, 2006, for a basic introduction). This type of analysis is particularity useful to estimate whether a variable temporally precedes and/or follows another variable. While LGM is concerned with intercepts and slopes over the entire assessed time span, CLA focuses on how the variables are related to each other from one point in time to the next. Hence, the latter provides a complementary perspective to LGM. In all analyses, calling and the different career preparation dimensions were assessed as latent constructs with their respective items as indicators. All analyses were conducted using M*plus* (Version 6.1; Muthén & Muthén, 2010) with the robust maximum likelihood estimation MLR.

To assess model fit, the Satorra-Bentler corrected (SB-χ^2) significance test (2001) was used which is suitable for nonnormally distributed data as is the case in our study. It is an absolute fit index that indicates how well the model fits the sample data. A significant test result (i.e. $p < 0.05$) suggests that the data differs significantly from the proposed model. However, because the test is very sensitive to sample size, it was supplemented with the comparative fit index (CFI) and the root mean square error of approximation (RMSEA). The CFI is a normed goodness-of-fit index that ranges from 0.0 to 1.0. Higher values indicate better fit relative to the independence model. The index adjusts for model parsimony and model complexity. Values close to .95 and above indicate acceptable model fit (Hu & Bentler, 1999). The RMSEA is a residual-based fit index. In addition to the noncentrality parameter, the sample size and degrees of freedom are included in its computation. A perfectly fitting model will obtain an RMSEA of zero. The index increases as the model misspecification becomes more severe. Values of .06 or less are considered acceptable (Hu & Bentler, 1999). Model comparisons were based on the Satorra-Bentler scaled χ^2-difference test where the degrees of freedom are specified as the difference in degrees of freedoms between both models (Satorra & Bentler, 2001).

Results

Preliminary Analyses

Before testing the hypotheses, it is necessary to prove that calling is a distinct construct that captures something different than the career preparation scales. We thus compared the model fit of a single-factor model with a model distinguishing calling from planning, decidedness, and self-efficacy among all students who participated at T1 (N = 1,207). Poor model fit was obtained for the one-factor model (SB-$\chi 2$ = 3413.47, df = 189, p < .001; CFI = .70; $RMSEA$ = .12). The fit of the four-factor model achieved good model fit (SB-$\chi 2$ = 1106.16, df = 183, p < .001; CFI = .92; $RMSEA$ = .06) and provided significantly better fit than the one-factor model (SB-scaled $\Delta\chi^2$ = 2307.31, df = 3; p < .001). Moreover, we established that a three-factor model distinguishing the three career preparation dimensions provided a significant better fit than a model where the three dimensions are treated as indicators of a single career preparation factor (one-factor model: SB-$\chi 2$ = 2726.91, df = 152, p < .001; CFI = .73; $RMSEA$ = .12; three-factor model: SB-$\chi 2$ = 1001.45, df = 149, p < .001; CFI = .91; $RMSEA$ = .07; model comparison: SB-scaled $\Delta\chi^2$ = 1725.45, df = 3; p < .001). This confirmed our approach to analyze the relationships between calling and each of the three career preparation dimensions separately. Confirming the scales' construct validity, all standardized factor loadings of the scale items on their respective career preparation construct were of considerable size (.54 to .92) and highly significant (all p < .001).

Prior to assessing change over time, it is further necessary to provide evidence of measurement invariance across time points (Horn & McArdle, 1992). Measurement invariance assures that the measures assess the same construct at different points in time regarding factor structure and item functioning (for more details on the procedure see Lance, Vandenberg, & Self, 2000). To proceed with LGM analyses, it was necessary to demonstrate at least scalar invariance. Scalar invariance is confirmed when equivalent factors structures and equal factor loadings are observed across time points (Horn & McArdle, 1992). All scales either fulfilled or exceeded this minimum requirement and the suitability of the scales for the subsequent LGM was confirmed.

Test of Hypotheses

Latent growth model

First, we assessed linear and non-linear univariate LGM of calling and the three career preparation scales to establish which growth curve best describes the change of each construct over time. Non-linear growth was modeled by freely estimating the slope factor at T3 (Curran & Hussong, 2003). For career planning and decidedness linear growth was confirmed. For self-efficacy and calling a non-linear model provided significantly better fit than a linear model and thus was used in subsequent analyses. The results support models of linear and non-linear growth despite the fact that the manifest means reported in Table 1 did not change substantially over time. This is possible because, firstly, the constructs are modeled as latent variables in LGM and estimates of growth are hence not based on the assumption that the constructs are equally well represented by each item. Second, the data can imply growth even if this growth is statistically nonsignificant, as indicated by nonsignificant slope means for our constructs. The means and variances of the intercepts of all constructs were significant, with the latter indicating that there were differences between individuals with regard to their initial level in the assessed constructs. Moreover, the slope variance of all career preparation measures was significant, suggesting meaningful differences between individuals with regard to their rates of change. This insight is provided by LGM because this method does not simply assess to what extent the sample mean changes across time. Instead the increase is modeled by a latent slope factor. Thus, while the assumption of growth over the course of the study may not be supported by our data in general, it is still possible that some individuals in the sample increased considerably while others remained stable or even experienced a decline in their in level of the construct in question. Such difference in developmental patterns, also referred to as variability of change, is assessed by the slope variance. We hence proceeded to investigate bivariate LGM in order to detect to what extent the observed interindividual variability of initial level and subsequent change in calling and the career preparation measures are related.

We specifically examined the hypotheses that calling would be significantly related to career preparation (a) within and (b) across time, and specified bivariate latent growth models estimating the correlations of the slopes and intercepts of

Table 2

Model Fit Indices, Parameters Estimates and Correlations between Intercepts and Slopes for Bivariate Latent Growth Curve Models

	Model fit			Correlations of Intercepts and Slopes	
	SB-χ2 *(df)*	*CFI*	*RMSEA*	Intercept calling with intercept career preparation measure	Slope calling with slope career preparation measure
Planning	611.96 (257)	.95	.04	.48***	-.09
Decidedness	530.97 (337)	.98	.03	.62***	.47*
Self-efficacy	401.09 (260)	.98	.03	.50***	.29*

Note. *N* = 846; *RMSEA* = Root Mean Square Error of Approximation; *CFI* = Comparative Fit Index.

$^{**}p < .01$.

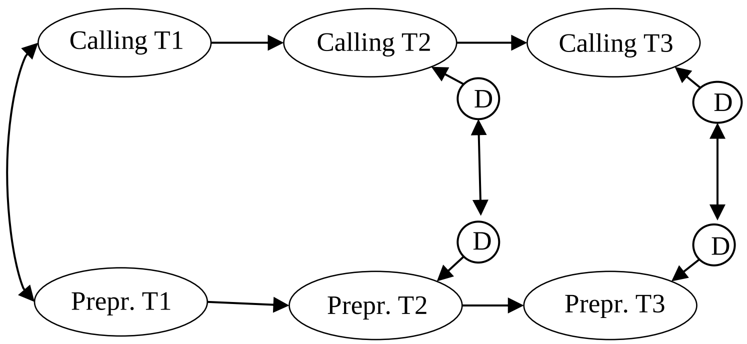
$^{***}p < .001$.

calling and one of the career preparation scales, respectively. Table 2 displays the model fit indices and correlations between slopes and intercepts of the variables (full results of the LGM analyses can be obtained from the authors upon request). Confirming the bivariate correlations among the observed measures reported in Table 1, the results in Table 2 showed significant correlations between the intercept of calling and the intercepts of each of the three career preparation scales, ranging from .48 to .62. This confirms significant relations of calling and career preparation within time, supporting H1to H3 (a). Second, a positive relationship between the slopes of calling and the slopes of decidedness and self-efficacy was obtained, suggesting that an increase in one of the constructs was associated with an increase in the other construct and confirming H2 and H3 (b), respectively. For planning, no significant relationship between its slope with the slope of calling was observed, thus rejecting H1 (b).

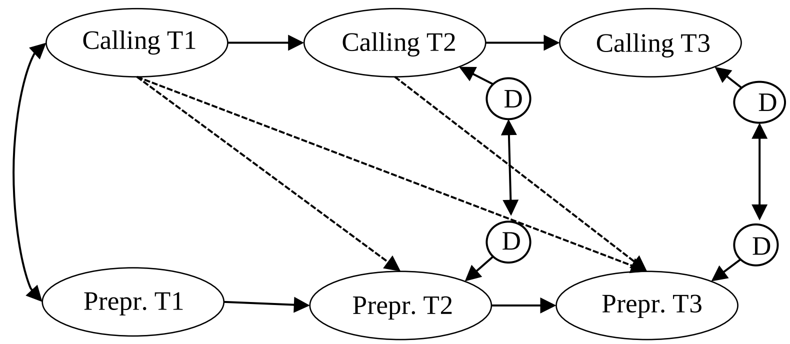
Cross-lagged analysis

Prior to examining the cross-lagged models, we tested measurement models which allowed the latent constructs assessed at the tree time points to correlate freely (Bagozzi & Edwards, 1998). Each measurement model displayed acceptable to good fit with *RMSEA* = .03 to .05 and *CFI* = .93 to .97. To assess the longitudinal associations between calling and each of the three career preparation scales, we next conducted comparisons between a series of nested cross-lagged models (see Figure 1). The starting point was the autoregressive model (M1) which estimates the stability of the constructs over time (Burkholder & Harlow, 2003). In the second model (M2), cross-lagged pathways were added from calling assessed at previous waves to the career preparation measures assessed at later waves. For the third model (M3), the relationships were reversed, and paths leading from the career preparation measures to calling were specified. The final model (M4) contained both cross-lagged effects, thus testing reciprocal effects.

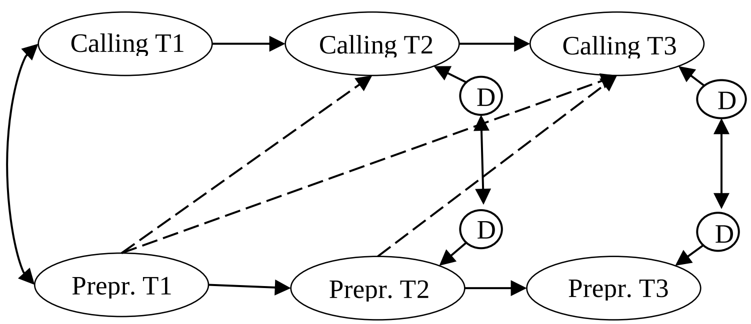
First, we tested the autoregressive model (M1) and found acceptable fit for all calling- career preparation models, with fit indices ranging from .04 to .05 for the *RMSEA* and from .93 to .96 for the *CFI*. We then tested whether either or both of the cross-lagged models (M2 or M3) provided a significantly better fit to the data than the more parsimonious autoregressive model. If significant, the better-fitting model of these two models was compared with the fully cross-lagged model (M4) to



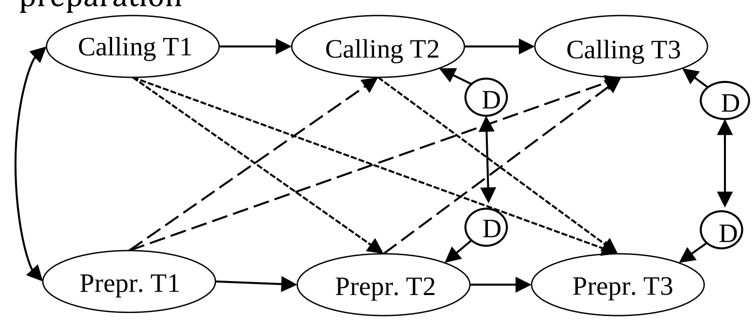
Model 1. Autoregressive Model



Model 2. Cross-lagged model of calling onto career preparation



Model 3. Cross-lagged model of career preparation onto calling



Model 4. Fully cross-lagged model of career preparation and calling

Figure 1. Models 1 to 4 of the cross-lagged analysis of calling and the respective career preparation measure over three time points. Prepr. = career preparation measure, D = disturbance terms associated with the latent variables at T2 and T3. For clarity, only the structural model is shown. All latent constructs were measured by their respective items. For clarity, items and paths representing residual covariances between like-items of the three measurement points of calling and career preparation respectively are omitted.

Table 3

Cross-lagged Standardized Regression Paths and Autoregressive Paths

	Cross-lagged paths (Standard regression estimate)						Autoregressive path[a]			
	Career preparation → Calling			Calling → Career preparation			Career preparation		Calling	
	T1→T2	T1→T3	T2→T3	T1→T2	T1→T3	T2→T3	T1→T2	T2→T3	T1→T2	T2→T3
Planning	.05	$.18^{*}$	-.04	-.06	$.11^{**}$	-.03	.60	.68	.59	.68
Decidedness	$.14^{**}$	$.17^{*}$	.02	.06	.03	.08	.73	.71	.58	.67
Self-efficacy	-.02	.04	-.01	$.21^{***}$	.02	$.19^{***}$	.36	.42	.65	.75

Note. $N = 846$.

[a] All autoregressive paths $p < .001$.

* $p < .05$.

$^{**}p < .01$.

$^{***}p < .001$.

determine the most appropriate model (Martens & Haase, 2006). Table 3 shows the autoregressive paths linking the same constructs across time points and the cross-lagged standardized regression paths between calling and the career preparation scales of the fully cross-lagged models.

For career planning and calling, the fully cross-lagged model was found to be most appropriate, confirming a mutual relation of calling and career planning as stated in H1 (c) and (d). With regard to decidedness and calling, M3, which specified temporal precedence of decidedness over calling, was identified as best-fitting, confirming H2 (c) but rejecting H2 (d) stating that calling would also precede decidedness. For self-efficacy the best-fitting model was M2, which specified that calling precedes the career preparation construct, confirming H3 (d). However, the fully cross-lagged model did not improve the model fit in either case, refuting H3 (c), the assumption that self-efficacy would also precede presence of calling.

In sum, our CLA analyses showed that calling temporarily preceded career planning and self-efficacy, but not decidedness. On the other hand, career planning and decidedness, but not self-efficacy, temporally preceded presence of calling.

Discussion

Previous theoretical and empirical work has suggested that calling and dimensions of career preparation are significantly and positively related. However, the nature of their relationship has not been thoroughly examined. Our study increases our understanding of how and why the presence of a calling is related to career preparation and in doing so also enhances our knowledge of how callings emerge and develop over time – a question also largely unaddressed in the empirical literature. First, we found that the level of the presence of a calling related positively and moderately to career planning and high to decidedness and self-efficacy. This finding supports theoretical assumptions that people with a calling would also possess more career metacompetencies (Hall & Chandler, 2005) as well as empirical findings showing positive correlations between calling and career decidedness and self-efficacy (Dobrow & Tosti-Kharas, 2011; Duffy & Sedlacek, 2007; Hirschi, 2011; Steger, et al., 2010). We further showed that the presence of a calling is empirically distinct from the assessed dimensions of career preparation and thus add to the existing literature suggesting the empirical distinctness of calling from constructs such as, for example, work engagement or career commitment (Dobrow & Tosti-Kharas, 2011).

Second, moving beyond establishing mere concurrent relations, the present study advances the literature by providing an in-depth analysis of how such relations can be explained by using a longitudinal design and applying LGM and CLA, both statistical methods particularly suited to investigate developmental change among multiple variables. Using bivariate LGM, we first investigated whether changes in one construct were related to changes in the other. As expected, change in the presence of a calling showed a moderate positive relation with changes in decidedness and self-efficacy, indicating that the constructs develop in parallel over time and that students who changed in the degree of the presence of a calling also changed similarly in their level of decidedness and self-efficacy. However, no relations between the slopes of calling and planning were observed, indicating that a change in calling was not related to corresponding changes in career planning. Therefore, whereas we could confirm that calling is significantly related to career preparation within time, our study advances the existing literature by suggesting that different change processes across time might be at work.

To further examine the relationship between the presence of a calling and career preparation over time, we then conducted CLA. Whereas other longitudinal studies have established that calling is related to the degree of vocational development (Duffy, Manuel, et al., 2011) or career self-efficacy (Dobrow & Tosti-Kharas, 2011) even across several years, our results provide a more sophisticated analysis of their relation by investigating cross-lagged effects. Going beyond extant research, CLA allowed us to tap more closely into the temporal precedence linking the presence of a calling and career preparation by controlling for internal stability over time and the concurrent relationships of the constructs. On a general level, the results supported a model of reciprocal effects between the presence of a calling and dimensions of career preparation. However, the direction of temporal precedence differed between career preparation dimensions, and we did not find consistent support for full reciprocal effects.

Specifically, the presence of a calling preceded increases in career planning and self-efficacy but not career decidedness. It thus appears that the presence of a calling motivates students to envision their vocational future and make plans for their careers, possibly to find ways of actualizing their callings in the work role. Calling also seems to enhance confidence in mastering challenges at work. Conversely, possibly because students with a sense of calling already possess high career decidedness (Duffy & Sedlacek, 2007; Hirschi, 2011), the presence of a calling did not substantially further enhance decidedness over time.

In turn, the presence of a calling was also preceded by aspects of career preparation, specifically career decidedness and planning. These findings support the notion that callings are dynamic (Dobrow, in press) and advance the literature by showing that having a sense of control over one's vocational development, clarity about personal preferences and career goals (i.e., career decidedness) as well as envisioning future career stages and possible selves (i.e., career planning) can strengthen and confirm a sense of calling among students. One could imagine that experiencing certainty about one's future career and making corresponding plans can contribute to the emergence of a calling because it might help students to discover their passion towards a particular career. Consistent with other studies reporting no relationship between ability and calling (Dobrow, 2012), we could not confirm that self-efficacy promotes a sense of calling.

To summarize, our results enrich our understanding of how a calling develops and how and why it is related to career preparation among university students. We can confirm that the presence of a calling is meaningfully related to career preparation within time. This relation can in turn be explained by callings preceding changes in certain aspects of career preparation (i.e., planning, self-efficacy) and certain aspects of career preparation (i.e., decidedness, planning) preceding changes in a calling.

Limitations

Some limitations must be considered when interpreting our results. First, we assessed a relatively brief time span of only one year. Although we tapped into a developmentally critical period and other research has observed meaningful change in career development variables within similar time lags, this approach may nonetheless limit our ability to observe developmental patterns among the assessed variables that may become apparent over the course of several years. This limitation is especially important to note because our results show that the presence of a calling and career preparation were relatively stable constructs in our sample over the assessed time-frame. Future studies are encouraged to assess developmental relationships over longer periods of time. Second, our sample was restricted to university students, and future research must examine the generalizability of our results among working samples as well. Third, we relied on self-report measures. Although the longitudinal design does diminish method effects, common method bias may be an issue and could be avoided by future research applying multi-source measures. Fourth, we applied a brief calling scale that allows participants to use their own notion of calling. Although the scale has received empirical support in several other studies, one limitation is that it is not clear what participants mean when indicating a "calling". A recent study by Hagmaier and Abele (2012) among German university students suggests that our measure taps mostly into the notion of a "transcendent guiding force" as a defining component of calling. Future studies are encouraged to assess developmental relations of calling and other career constructs with other measures of calling to enrich our understanding of how different aspects of calling are related to career development. Moreover, it is possible that the two items of our applied calling measure refer to different concepts. The first item addresses the presence of a calling more generally while the second refers to whether somebody knows how to apply a calling to her or his career development. A post-hoc analysis revealed that the second item showed consistently larger correlations with the applied career preparation measures than the first item. It might thus be important to distinguish between having a calling and knowing how to implement it into a career and we encourage further research to address this distinction. Fifth, we measured confidence in terms of occupational self-efficacy beliefs which might imply certain validity constrains among a student sample. Future research could investigate efficacy believes regarding other career relevant domains such as decision making. Sixth, although our research design allowed cross-lagged analyses that are particularly useful for investigating potential causal mechanisms among constructs in field research, one must be careful when making causal claims. Inferences about causality may be wrong because the assessed variables have not reached equilibrium or because variables that might alter the influences are missing from the model (Shadish, Cook, & Campbell, 2002). Therefore, causal effects must be further studied in more rigorously controlled experiments to be certain about the true causal influences between calling and career preparation. Finally, as common in longitudinal research attrition was an issue in our sample. Attrition might have been somewhat larger than occurred in other studies with university students (Duffy, Manuel, et al., 2011) because we did not sample students attending a specific class or study field but all students attending the university which made it harder to track them for follow-up surveys. However, we believe that this setback is compensated by the increased external validity of our sample compared to investigating considerably narrower selection of students attending a particular subject class. Moreover, we did not find systematic effects of missingness in our data. Furthermore, utilizing sophisticated estimation procedures that provide accurate estimates of missing data as done in the present study is beneficial for two reasons. First, we were able to use the data of a large number of participants and hence increase the power of our analyses. Second, we avoided the potential bias of listwise deletion of participants with incomplete data (see Graham, 2009, for more details on how to treat missing data).

Conclusions and Implications

Our study advances the theoretical understanding of how callings develop as well as how and why they are related to other prominent career development constructs, specifically, dimensions of career preparation. In sum, our results suggest that showing higher career preparedness in terms of career decidedness and planning can help people to develop and/or confirm a sense of calling in their careers. In turn, experiencing a calling appears to be a motivating force for engaging in career preparation and might thus help to navigate a complex career terrain and address career development tasks (Hall & Chandler, 2005).

With regard to counseling practice, addressing callings might be important for a considerable number of clients (e.g., Hunter, Dik, & Banning, 2010; Wrzesniewski, McCauley, Rozin, & Schwartz, 1997). Our results imply that helping clients to find or develop a calling can be beneficial because callings may have positive effects on the general ability to cope with vocational demands by increasing subsequent engagement in career preparation. Dik and Duffy (2009) suggested that introspection might be important in order to *find* a calling stemming from an external source. Offering a complementary perspective, our results imply that increasing the degree of career preparation might also be important in order to *develop* (Dobrow, in press) a calling among university students. For example, clarifying personal preferences and career goals and envisioning possible future states and selves (Markus & Nurius, 1986) may be useful tools in this regard. Career counselors could enhance their regular practice by linking such activities more explicitly to questions of meaning and purpose in work and how clients might develop a sense of calling in their career.

References

Abele, A. E., & Spurk, D. (2009). The longitudinal impact of self-efficacy and career goals on objective and subjective career success. *Journal of Vocational Behavior, 74*(1), 53-62. doi: 10.1016/j.jvb.2008.10.005

Abele, A. E., & Wiese, B. S. (2008). The nomological network of self-management strategies and career success. *Journal of Occupational and Organizational Psychology, 81*(4), 733-749. doi: 10.1348/096317907X256726

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist, 55*(5), 469-480. doi: 10.1037/0003-066X.55.5.469

Bagozzi, R. P., & Edwards, J. R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods, 1*(1), 45-87.

Burkholder, G. J., & Harlow, L. L. (2003). An illustration of a longitudinal cross-lagged design for larger structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 10*(3), 465-486.

Creed, P. A., Fallon, T., & Hood, M. (2009). The relationship between career adaptability, person and situation variables, and career concerns in young adults. *Journal of Vocational Behavior, 74*(2), 219-229. doi: 10.1016/j.jvb.2008.12.004

Curran, P. J., & Hussong, A. M. (2003). The use of latent trajectory models in psychopathology research. *Journal of Abnormal Psychology, 112*(4), 526-544. doi: 10.1037/0021-843x.112.4.526

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268. doi: 10.1207/S15327965pli1104_01

Deng, Y., Hillygus, D. S., Reiter, J. P., Si, Y., & Zheng, S. (in press). Handling Attrition in Longitudinal Studies: The Case for Refreshment Samples. *Statistical Science*.

Dik, B. J., & Duffy, R. D. (2009). Calling and vocation at work: Definitions and prospects for research and practice. *The Counseling Psychologist, 37*(3), 424-450. doi: 10.1177/0011000008316430

Dik, B. J., Eldridge, B. M., Steger, M. F., & Duffy, R. D. (2012). Development and validation of the Calling and Vocation Questionnaire (CVQ) and Brief Calling

Scale (BCS). *Journal of Career Assessment, 20*(3), 242-263. doi: 10.1177/1069072711434410

Dobrow, S. R. (in press). Dynamics of calling: A longitudinal study of musicians. *Journal of Organizational Behavior*. doi: 10.1002/job.1808

Dobrow, S. R., & Tosti-Kharas, J. (2011). Calling: The development of a scale measure. *Personnel Psychology, 64*(4), 1001-1049. doi: 10.1111/j.1744-6570.2011.01234.x

Duffy, R. D., Allan, B. A., & Dik, B. J. (2011). The presence of a calling and academic satisfaction: Examining potential mediators. *Journal of Vocational Behavior, 79*(1), 74-80. doi: 10.1016/j.jvb.2010.11.001

Duffy, R. D., Bott, E. M., Allan, B. A., Torrey, C. L., & Dik, B. J. (2012). Perceiving a calling, living a calling, and job satisfaction: Testing a moderated, multiple mediator model. *Journal of Counseling Psychology, 59*(1), 50-59. doi: 10.1037/a0026129

Duffy, R. D., Manuel, R. S., Borges, N. J., & Bott, E. (2011). Calling, vocational development, and well being: A longitudinal study of medical students. *Journal of Vocational Behavior, 79*(2), 361–366. doi: 10.1016/j.jvb.2011.03.023

Duffy, R. D., & Sedlacek, W. E. (2007). The presence of and search for a calling: Connections to career development. *Journal of Vocational Behavior, 70*(3), 590-601. doi: 10.1016/j.jvb.2007.03.007

Duffy, R. D., & Sedlacek, W. E. (2010). The salience of a career calling among college students: Exploring group differences and links to religiousness, life meaning, and life satisfaction. *Career Development Quarterly, 59*(1), 27-41.

Duncan, T. E., & Duncan, S. C. (1995). Modeling the processes of development via latent variable growth curve methodology. *Structural Equation Modeling: A Multidisciplinary Journal, 2*(3), 187-213. doi: 10.1080/10705519509540009

Gould, S. (1979). Characteristics of career planners in upwardly mobile occupations. *The Academy of Management Journal, 22*(3), 539-550.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*(1), 549-576. doi: 10.1146/annurev.psych.58.110405.085530

Hagmaier, T., & Abele, A. E. (2012). The multidimensionality of calling: Conceptualization, measurement and a bicultural perspective. *Journal of Vocational Behavior, 81*(1), 39-51. doi: 10.1016/j.jvb.2012.04.001

Hall, D. T., & Chandler, D. E. (2005). Psychological success: When the career is a calling. *Journal of Organizational Behavior, 26*(2), 155-176. doi: 10.1002/job.301

Hirano, K., Imbens, G. W., Ridder, G., & Rubin, D. B. (2001). Combining panel data sets with attrition and refreshment samples. *Econometrica, 69*(6), 1645-1659. doi: Doi 10.1111/1468-0262.00260

Hirschi, A. (2011). Callings in career: A typological approach to essential and optional components. *Journal of Vocational Behavior, 79*(1), 60-73. doi: 10.1016/j.jvb.2010.11.002

Hirschi, A., Niles, S. G., & Akos, P. (2011). Engagement in adolescent career preparation: Social support, personality and the development of choice decidedness and congruence. *Journal of Adolescence, 34*(1), 173-182 doi: 10.1016/j.adolescence.2009.12.009

Holland, J. L., Daiger, D. C., & Power, P. G. (1980). *My Vocational Situation*. Palo Alto, CA: Consulting Psychologists Press.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*(3), 117-144. doi: 10.1080/03610739208253916

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55. doi: 10.1080/10705519909540118

Hunter, I., Dik, B. J., & Banning, J. H. (2010). College students' perceptions of calling in work and life: A qualitative analysis. *Journal of Vocational Behavior, 76*(2), 178-186. doi: 10.1016/j.jvb.2009.10.008

Jörin Fux, S. (2006). *Persönlichkeit und Berufstätigkeit [Personality and Work]*. Göttingen: Cuviller Verlag.

Jörin, S., Stoll, F., Bergmann, C., & Eder, D. (2004). *Explorix® - das Werkzeug zur Berufswahl und Laufbahnplanung [Explorix - the tool for career choice and career planning]*. Berne, Switzerland: Hans Huber.

Lance, C. E., Vandenberg, R. J., & Self, R. M. (2000). Latent growth models of individual change: The case of newcomer adjustment. *Organizational*

Behavior and Human Decision Processes, 83(1), 107-140. doi: 10.1006/obhd.2000.2904

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist, 41*(9), 954-969.

Martens, M. P., & Haase, R. F. (2006). Advanced applications of structural equation modeling in counseling psychology research. *The Counseling Psychologist, 34*(6), 878-911.

Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Authors.

Peterson, C., Park, N., Hall, N., & Seligman, M. E. P. (2009). Zest and work. *Journal of Organizational Behavior, 30*(2), 161-172. doi: 10.1002/job.584

Rigotti, T., Schyns, B., & Mohr, G. (2008). A short version of the Occupational Self-Efficacy Scale: Structural and construct validity across five countries. *Journal of Career Assessment, 16*(2), 238-255. doi: 10.1177/1069072707305763

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514. doi: 10.1007/Bf02296192

Savickas, M. L. (1997). Career adaptability: An integrative construct for life-span, life-space theory. *Career Development Quarterly, 45*(3), 247-259.

Savickas, M. L. (2005). The theory and practice of career construction. In S. D. Brown & R. W. Lent (Eds.), *Career development and counseling* (pp. 42-70). Hoboken, NJ: Wiley & Sons Inc.

Shadish, W., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental design for generalized causal inference*. Boston, MA: Houghton-Mifflin.

Skorikov, V. B. (2007). Continuity in adolescent career preparation and its effects on adjustment. *Journal of Vocational Behavior, 70*(1), 8-24. doi: 10.1016/j.jvb.2006.04.007

Steger, M. F., Pickering, N. K., Shin, J. Y., & Dik, B. J. (2010). Calling in work: Secular or sacred? *Journal of Career Assessment, 18*(1), 82-96. doi: 10.1177/1069072709350905

Strauss, K., Griffin, M. A., & Parker, S. K. (2012). Future work selves: How salient hoped-for identities motivate proactive career behaviors. *Journal of Applied Psychology, 97*(3), 580-598. doi: 10.1037/a0026423

Stringer, K., Kerpelman, J., & Skorikov, V. (2011). Career preparation: A longitudinal, process-oriented examination. *Journal of Vocational Behavior, 79*(1), 158-169. doi: 10.1016/j.jvb.2010.12.012

Super, D. E. (1990). A life-span, life-space approach to career development. In D. Brown & L. Brooks (Eds.), *Career choice and development* (2nd ed., pp. 197-262). San Francisco, CA: Jossey-Bass.

Wayne, S. J., Liden, R. C., Kraimer, M. L., & Graf, I. K. (1999). The role of human capital, motivation and supervisor sponsorship in predicting career success. *Journal of Organizational Behavior, 20*(5), 577-595. doi: 10.1002/(sici)1099-1379(199909)20:5<577::aid-job958>3.0.co;2-0

Wrzesniewski, A., McCauley, C., Rozin, P., & Schwartz, B. (1997). Jobs, careers, and callings: People's relations to their work. *Journal of Research in Personality, 31*(1), 21-33. doi: 10.1006/jrpe.1997.2162

III. Paper "The Protean Career Orientation: Investigating Gender Differences, Temporal Stability, and Predictive Utility"

Abstract

In an increasingly dynamic work environment, a protean career orientation gains importance. However, extant research has not sufficiently addressed the temporal stability and incremental predictive utility of a protean career orientation or issues of causality linking this orientation to career outcomes. Based on a unidimensional measure for a protean career orientation, we present a series of studies that (1) establish the scale's unidimensionality and measurement invariance across gender within separate samples of students and working professionals as well as measurement invariance between both samples; (2) demonstrate measurement invariance and differential stability over six months among students and professionals; (3) show that a protean career orientation partially mediates the relationship between personality dispositions (i.e., proactive personality, core self-evaluations) and proactive career behaviors and career satisfaction among students and employees; (4) demonstrate that a protean career orientation possesses incremental predictive validity regarding proactive career behaviors and career satisfaction beyond personality dispositions among students and employees; and (5) based on a cross-lagged study among employees, address the issue of causality linking a protean career orientation to career satisfaction by showing that career satisfaction predicts a protean career orientation but that a protean career orientation does not predict career satisfaction.

Complete reference: Herrmann, A., Hirschi, A., & Baruch, Y. (under review, September 2013). *The Protean Career Orientation: Investigating Gender Differences, Temporal Stability, and Predictive Utility*. Journal of Applied Psychology.

Introduction

Changes in the workplace over the past few decades have spawned an increase in research investigating new career types, characterized by increased self-directedness, flexibility, and the aim of subjective career success. However, the field of career studies suffers from fragmentation in terms of its theoretical underpinnings and conceptual frameworks, which poses a challenge to scholars in the field (Arnold & Cohen, 2008; Arthur, Hall, & Lawrence, 1989). One of the few widely accepted theoretical concepts is the protean career, described as being flexible, self-directed, and values-driven (Hall, 1996, 2004). Together with the boundaryless career (Arthur & Rousseau, 1996), the protean career is considered to be among the most important career concepts that characterize the more contemporary ‘new careers’ (Enache, Sallan, Simo, & Fernandez, 2011; Sullivan & Baruch, 2009).

Yet, our understanding of contemporary career concepts is hampered by the lack of reliable and valid measures for evaluating emerging new constructs. Such measures would enable researchers to study individual and organizational career development and to test related conceptual frameworks. While the value of new concepts certainly depends on their theoretical innovation, this must be complemented by relevance and applicability within a global context (Cerdin & Pargneux, 2009). The need for further investigation is becoming particularly important in light of recent criticism of the developing notion of ‘new careers,’ particularly due to the lack of rigorous empirical evaluation of theoretical concepts (Inkson, Gunz, Ganesh, & Roper, 2012). As a result, validation studies are critical to ensure the continued development of new knowledge. However, these studies are difficult to conduct due to their complexity and the rigor required for a worthy validation study (Gill & Hodgkinson, 2007; Plouffe & Grégoire, 2011). Such challenges include, for example, the need to provide evidence of temporal stability in the development of new career orientations and to establish the incremental validity of new career orientations beyond the already-established predictive utility of general personality dispositions (Hinkin, 1998).

To answer this challenge and to enable an evaluation of the protean career concept within a global context, we employed a German measure of the protean career orientation (i.e., a career attitude characterized by valuing flexibility, self-directedness, and career success according to personal values) within a varied population, including professionals from various industries as well as university students. The general contributions of this paper include (1) providing an empirical evaluation of a measure of the protean career orientation and (2) offering substantive new insights into the functioning of a protean career orientation among employees and students. The specific contributions of the paper stem from (a) establishment of the factor structure and applicability of a measure for protean and traditional career orientations with respect to gender; (b) investigation of the interindividual stability of a protean career orientation; (c) examination of the incremental predictive utility of a protean career orientation, beyond that of personality dispositions, for estimating career outcomes; and (d) consideration of the temporal precedence that a protean career orientation has and its relation to subjective career success.

Overview of the Studies

We present four studies that investigate a protean career orientation among university students and employees. Investigating the same research questions and measures across different samples is highly relevant for making inferences about the generalizability of the research findings. However, most existing research on the protean career orientation has relied on only one type of sample. To address this issue, we conducted Study 2 with university students while Studies 3 and 4 examined working professionals. Study 1 included both students and professionals.

In Study 1, we introduce a new German language adaptation of a measure to assess a protean and a traditional career orientation. An individual with a protean career orientation values self-directedness and defines career success according to their personal values. An individual with a traditional career orientation places importance on objective career success such as salary and position (Baruch, 2008). We first aim to confirm the applicability and factor structure of the translated measure among both university students and employees and to then investigate possible gender differences in a protean career orientation. We also assess the measurement invariance across students and employees. Using a longitudinal study design in Study 2, we examine measurement invariance and differential stability of a protean career orientation over six months. We further provide evidence of concurrent, discriminant, predictive, and incremental validity of a protean career orientation by investigating how such an orientation is related to students' proactive dispositions and career management behaviors. Study 3 aims to replicate and advance findings from Study 2 among employees by assessing whether a protean career orientation mediates the relationship between personality (i.e., core self-evaluations and proactivity) and career outcomes (i.e., career satisfaction and proactive career behaviors). Study 3 also provides further evidence for the concurrent, discriminant, and incremental validity of our measure. Finally, Study 4 elaborates on the findings from Study 3 and uses a cross-lagged design to address the issue of temporal precedence in the relationship between a protean career orientation and career satisfaction. It also provides further evidence of measurement invariance and differential stability of a protean career orientation across time and among employees.

Study 1: Examination of Factor Structure, Construct Discrimination, and Measurement Invariance across Gender among University Students and Employees

We measured a protean and traditional career orientation using a German translation of an existing reliable and valid measure (Baruch, 2008). This measure assesses a protean career orientation as a one-dimensional construct – in contrast to the frequently applied measure from Briscoe, Hall, and De Muth (2006) which was developed to assess a protean career orientation with two dimensions: (1) a self-directed career management attitude, and (2) a values-driven career attitude. However, subsequent research has not always confirmed the proposed two-factor structure of the measure because the values-driven scale emerged as problematic in non-US samples (Chan et al., 2012). As a consequence, most empirical research (Briscoe, Henagan, Burton, & Murphy, 2012; De Vos & Soens, 2008; Park, 2009; Verbruggen & Sels, 2008) has assessed only the first dimension (i.e., self-directed career management attitude) of the protean career orientation. This severely limits the content validity of such a measurement approach to assessing the broader construct of a protean career orientation. The measure applied herein avoids this pitfall, assessing the protean career orientation as a one-dimensional construct, as originally presented in the literature by Hall and colleagues (Hall & Mirvis, 1996; Hall & Moss, 1998). Moreover, the measure also enables the assessment of a traditional career orientation. Although traditional and protean career orientations are frequently contrasted, (Hall, 1996, 2004), existing empirical research has not yet assessed the similarities and differences between the two. Advancing extant research, our applied measure enables one to empirically examine the relationship between the two career orientations and to compare their respective effects along several criterion variables. Where applicable, we used this advantage in the reported studies by contrasting the results obtained from both career orientations.

The first two authors, both of whom are native German speakers fluent in English, psychologists and intimately familiar with the construct assessed by the scales, independently translated the protean and traditional career orientation scales that had originally been developed and validated in English (Baruch, 2008; Baruch, Bell, & Gray, 2005; Baruch & Quick, 2007). Content was translated from the scales' source versions into the German target versions. After initial translations, these authors convened a reconciliation meeting during which any differences in translations were discussed and after which a final translated version of each item was agreed upon. This procedure was chosen because it is particularity useful in ensuring authenticity, connotation and comprehensibility, which are frequently compromised when incorporating a back-translation approach (van de Vijver & Leung, 1997). The applied measure uses seven items to assess a protean career orientation and three items to assess a traditional career orientation (see Table 1 and 2), with each item being rated along a seven-point Likert scale.

First in Study 1, we established the factor structure of the German version of the protean career scale within two samples, one with university students and one with employees. Use of multiple independent samples has been encouraged in an effort to develop and validate a new scale to support its factor structure (Hinkin, 1995; MacKenzie, Podsakoff, & Podsakoff, 2011). We applied exploratory factor analysis (EFA) to the protean career items to establish the scale's unidimensionality. This was followed by a confirmatory factor analysis (CFA) which enabled a more precise evaluation of the measurement model (Hinkin, 1998). Second, we tested the discriminant validity

Table 1

Item Content and Item Statistics of the Protean Career Orientation Scale for Study 1

	Item	Students (N = 1,224)			Employees (N = 526)		
		M	*SD*	CITC	*M*	*SD*	CITC
PC 1	For me, career success means how I am doing compared to my goals and values / Für mich bedeutet Karriereerfolg, wie ich im Vergleich zu meinen Zielen und Werten dastehe	5.50	1.26	.36	5.53	1.19	.48
PC 2	If I have to find a new job outside the organization, it would be easy / Wenn ich eine neue Stelle außerhalb des Unternehmens finden muss, wäre das einfach.	4.33	1.19	.32	4.55	1.50	.41
PC 3	I am in charge of my own career / Ich übernehme die Verantwortung für meine eigene Karriere	5.92	0.94	.50	5.90	1.03	.63
PC 4	I navigate my own career, according to my plans / Ich steuere meine eigene Karriere meinen Plänen entsprechend.	5.20	1.15	.44	5.12	1.29	.59
PC 5	I take responsibility for my own development / Ich übernehme Verantwortung für meine eigene Entwicklung	6.01	0.85	.48	6.02	0.96	.60
PC 6	For me, career success means having a high level of freedom and autonomy / Für mich bedeutet Karriereerfolg, ein hohes Maß an Freiheit und Autonomie zu haben.	5.35	1.29	.34	5.59	1.23	.45
PC 7	For me, career success means being flexible / Für mich bedeutet Karriereerfolg flexibel zu sein	5.00	1.35	.40	5.06	1.38	.41

Note. PC = Protean career orientation, *M* = mean, *SD* = standard deviation, *CITC* = corrected item-total correlation.

Table 2

Item Content and Item Statistics of the Traditional Career Orientation Scale for Study 1

	Item	Students (*N* = 1,224)			Employees (*N* = 526)		
		M	*SD*	CITC	*M*	*SD*	CITC
TC 1	For me, career success means having a high income / Für mich bedeutet Karriereerfolg ein hohes Einkommen zu erzielen	5.19	1.42	.64	5.24	1.32	.62
TC 2	For me, career success means high status / Für mich bedeutet Karriereerfolg hoher Status	4.75	1.50	.66	4.62	1.49	.62
TC 3	For me, career success means reaching a senior level position / Für mich bedeutet Karriereerfolg eine hohe Position zu erreichen	4.97	1.45	.72	4.93	1.45	.70

Note. TC = Traditional career orientation, *M* = mean, *SD* = standard deviation, *CITC* = corrected item-total correlation.

of the protean career scale by applying both types of factor analyses to a combined set of the protean career items with the three items that assess traditional career orientation. Because these represent two distinct constructs, we expected to obtain a two-factor solution in the EFA and to confirm this solution in the subsequent CFA. Third, we investigated the measurement invariance of the protean career scale across gender. Demonstrating equivalence between males and females would show that the scale assesses the same construct for each (Vandenberg & Lance, 2000). Such a finding is also necessary to allow for meaningful group comparisons based on the scale (Sass, 2011), as has been conducted for the protean career orientation in some existing studies (Briscoe, et al., 2006).

Investigations of gender differences are important because research suggests that, on average, men and women differ significantly in certain career attitudes. For example, Mainiero and Sullivan (2005) explained that women usually follow less traditional career paths compared to men. Other research has shown that men typically prefer jobs that provide high salaries, power, prestige, and career opportunities while women typically prefer jobs that allow for helping others and spending time with family, but which also develop their knowledge and skills (Konrad, Ritchie, Lieb, & Corrigall, 2000). Hence, men may appear more oriented towards traditional careers than women, and women may be seen as valuing intrinsic aspects of career success more than men. Providing some support for this notion, Gerber, Wittekind, Grote, and Staffelbach (2009) found that men reported a traditional/promotion career orientation more often than women. However, a majority of the existing research has either not addressed gender differences in a protean career orientation or yielded inconsistent results. For example, Briscoe et al. (2006) found no gender differences in varying samples of undergraduate and part-time working MBA students. Conversely, Segers, Inceoglu, Vloeberghs, Bartram, and Henderickx (2008) studied a large international sample of working professionals and found that women scored higher than men on the values-driven and self-directed dimensions of a protean career orientation scale. Study 1 advances existing research by providing evidence of measurement invariance across gender for our new measure before investigating potential differences between groups.

Finally, we also examined the measurement invariance of the protean career orientation measure between students and employees because this is a necessary prerequisite to comparing these sets of individuals who are in different career stages.

Materials and Method

Participants and procedure

For the student sample, students in their second and third years of study at a German university (approx. N = 3,500) were invited via email to participate in a study on career development. We received a response rate of approximately 35%, which is well within the typical range found in behavioral science (Baruch & Holtom, 2008). As an incentive for completing the questionnaire, participants were told they could enter a lottery drawing for two prizes of 450€ each. The resulting sample of N = 1,224 students was 63.2% female, with a mean age of M = 23.91 years (SD = 2.75). The mean current study semester was M = 4.12 (SD = 2.36). Participants were cumulatively enrolled in 34 different majors, with the largest groups studying management and entrepreneurship (19%), business administration (14%), business psychology (12%), cultural studies (9%), and environmental science (7%).

For the sample of employees, university alumni from three German universities were contacted via email (N = 927) and invited to complete the online questionnaire. Those who had not yet responded received two reminder emails, each one week apart. This strategy resulted in a final response rate of 57% (N = 526 participants). Participation in a lottery drawing with several prizes ranging from 25€ to 380€ and a total value of 880€ was offered as an incentive. The sample was 58.9% female, with a mean age of M = 28.74 (SD = 5.15). The majority of the sample had received a Master's degree or equivalent (59%) and about a third (31%) had obtained a Bachelor's degree. Participants were employed in many different industry sectors, with the largest groups working in business administration (20%), engineering (16%), education (12%), marketing and advertising (8%), and information technologies (7%). As is customary in Germany, ethnic background was not assessed.

Measures

Protean career orientation. A seven-item measure assessing the extent to which a subject exhibited a protean career orientation was administered. Each item was rated along a seven-point Likert scale. Cronbach's alpha was .69 (.68/.69 for males/females, respectively) for the student sample and .77 (.82/.73 for males/females, respectively) for the employee sample.

Traditional career orientation. A three-item measure assessing the extent to which a subject exhibited a traditional career orientation was administered. Each item was rated along a seven-point Likert scale. Cronbach's alpha for the scale was .82 for the student sample and .80 for the employee sample.

Results and Discussion

Factor Structure – EFA

Principal axis factoring (PAF) with Promax rotation was applied to assess the unidimensionality of the protean career orientation scale. We used multiple criteria to determine the appropriate number of factors, including parallel analysis (PA; Horn, 1965), the scree test (Cattell, 1966), and factor interpretability and meaningfulness (Kahn, 2006). For both the student and employee samples, results from PA pointed to a two-factor solution while the scree plot suggested a one-factor solution. The one-factor solution displayed moderate-to-high factor loadings for all items. The extraction of two factors yielded a solution with a poorly defined second factor and non-negligible cross-loadings for all items. Thus, based on interpretability – and supported by the scree plot – the one-factor solution was found to be more appropriate for both samples.

Next, PAF was conducted using the combined set of items from the protean and traditional career orientation measures to demonstrate discriminant validity of the protean career version. For both samples, PA suggested the extraction of three factors, while the scree plot indicated only a two-factor solution. The two-factor solution contained two clearly defined factors, one representing the protean construct and the other representing the traditional career construct with moderate-to-high factor loadings of their respective items and negligible cross-loadings. Extracting three factors resulted in one distinct traditional factor and two protean career factors. In this solution, all protean items displayed considerable secondary loadings on the respective other protean factor, suggesting a ‘factor splitting’ due to overextraction (Wood, Tataryn, & Gorsuch, 1996). Thus, based on interpretability – and supported by the scree plot – the two-factor solution was found to be more appropriate and suggested good discriminant validity for the protean career construct.

Factor Structure – CFA

CFA was conducted using a robust maximum likelihood estimation method (MLR; Satorra & Bentler, 2001). Model fit was assessed using multiple criteria, including (1) the Satorra-Bentler corrected χ^2 (SB-χ^2) test, (2) the comparative fit index (CFI), (3) the root mean square error of approximation (RMSEA), and (4) the standardized root mean square residual (SRMR). We used frequently applied cut-off values suggested by Hu and Bentler (1999). These values were a CFI greater than .90, an SRMR less than .08, and an RMSEA of close to .06.

The model fit obtained for the original measurement model was unsatisfactory for both samples (Students: SB-χ^2 = 244.73, *df* = 14, *p* < .01; *CFI* = .81; *RMSEA* = .12, 90% CI = .10-.13; *SRMR* = .07; Employees: SB-χ^2 = 112.54, *df* = 14, *p* < .01; *CFI* = .87; *RMSEA* = .12, 90% CI = .10-.14; *SRMR* = .07). As suggested by the modification indices, we next freed the correlated error term between items 6 and 7, which led to considerable improvement in the model fit (Students: SB-χ^2 = 116.52, *df* = 13, *p* < .01; *CFI* = .92; *RMSEA* = .08, 90% CI = .07-.09; SRMR = .05; Employees: *SB*-χ^2 = 53.39, *df* = 13, *p* < .01; *CFI* = .95; *RMSEA* = .08, 90% CI = .06-.10; *SRMR* = .05). Upon additionally freeing the error terms between items 3 and 5, an acceptable model fit was achieved (Students: *SB*-χ^2 = 77.35, *df* = 12, *p* < .01; *CFI* = .95; *RMSEA* = .07, 90% CI = .05-.08; *SRMR* = .04; Employees: *SB*-χ^2 = 34.09, *df* = 12, *p* < .01; *CFI* = .97; *RMSEA* = .06, 90% CI = .04-.08; *SRMR* = .04).

Some authors have argued that the presence of correlated error terms suggests the concurrent existence of extraneous factors (Gerbing & Anderson, 1984). However, the error terms could also be a methodological artifact resulting from the item format, in which case it is justifiable to release them in an effort to obtain a better-fitting model (Byrne, Shavelson, & Muthen, 1989). In the present case, both correlated error terms could be explained by the respective item pair sharing the same item stem, thus suggesting a method effect. Therefore, freeing both covariances was justified. Obtaining a good fit for this modified model confirmed the scale's unidimensionality. The model with two correlated error terms was thus used as a baseline model for the subsequent investigation of measurement invariance.

To demonstrate that protean and traditional career orientations are distinct rather than simply opposite constructs, we compared the model fit of a single-factor model with a two-factor model, one representing a protean career orientation and the other representing a traditional one. The one-factor model displayed a very poor fit (Students: SB-χ^2 = 687.76, *df* = 33, $p < .001$; *CFI* = .72; *RMSEA* = .13, 90% CI = .12-.14; *SRMR* = .13; Employees: SB-χ^2 = 517.30, *df* = 33, $p < .001$; *CFI* = .60; *RMSEA* = .17, 90% CI = .16-.18; *SRMR* = .16). In contrast, the two-factor model achieved a very good and significantly better model fit in both samples (Students: SB-χ^2 = 151.38, *df* = 32, $p < .001$; *CFI* = .95; *RMSEA* = .06, 90% CI = .05-.06; *SRMR* = .05; SB-corrected $\Delta\chi^2$ = 540.97, *df* = 1; $p < .001$; Employees: SB-χ^2 = 66.35, *df* = 32, $p < .001$; *CFI* = .97; *RMSEA* = .05, 90% CI = .03-.06; *SRMR* = .04; SB-corrected $\Delta\chi^2$ = 429.03, *df* = 1; $p < .001$).

Interestingly, a positive correlation between protean and traditional career orientations was observed in the two-factor model (students: $r = .18$; employees: $r = .27$, both $p < .001$). Theoretical accounts often present a protean career orientation as being opposite to a traditional career orientation (e.g., Hall, 1996). However, our results suggest that protean and traditional career orientations share some variance that might represent a general career orientation, or at least indicates that they do not necessarily oppose each other as found by Baruch and Quick (2007). Moreover, the relationship between the two orientations is stronger for individuals with work experience than for those who are students.

Measurement invariance for gender

Next, the measurement invariance of the protean career scale for gender was investigated for both samples by testing a sequence of nested CFA models with increasing invariance restrictions (cf. Vandenberg & Lance, 2000). Model comparisons were based on the Chisquare-difference test with Satorra-Bentler correction (Satorra & Bentler, 2001) and differences in CFI with a ΔCFI ≤ -.01, indicating non-invariance (Cheung & Rensvold, 2002). These comparisons were supplemented by cut-offs for ΔRMSEA and ΔSRMR. The test of invariance of factor loadings was supported when ΔRMSEA ≤ .015 or ΔSRMR ≤ .03 while the invariance of item intercepts and residual invariances, respectively, was confirmed when ΔRMSEA ≤ .015 or ΔSRMR ≤ .01 (Chen, 2007).

In the baseline model (Model 1), a one-factor solution with two correlated error terms was specified, but no further restrictions were applied. In Model 2, the factor loadings were restricted to being equal across males and females to test for metric invariance which was supported in both samples. The intercepts of corresponding items were restricted to being invariant in Model 3. The intercept of item 2 needed to be estimated freely in both groups (Model 3b), thus supporting partial scalar invariance. In Model 4, the assumption of equal item reliability could be upheld for the student sample but not for the sample of employees for which three residual variances had to be released (item 2, 4 and 7) because they differed between groups. In Model 5, the factor variances were constrained to being equal for both groups. The final model (Model 6) was the most restrictive because it imposed the additional constraint of equal factor means. Comparing the means is considered meaningful even when only partial scalar invariance is established, as was the case for both samples (Steenkamp & Baumgartner, 1998). All assumptions imposed by these subsequent model comparisons were supported. In addition, an acceptable model fit was achieved for all models. The results, displayed in detail in Tables 3 and 4, confirm that the protean career scale is an appropriate measure for assessing a protean career orientation for in males and females. Furthermore, our results substantively address the contradictory findings in extant research regarding gender differences in protean career orientations since we found no such differences in either the student or employee samples.

Measurement invariance for students versus employees

Measurement invariance of the protean career orientation scale for students and employees was assessed using the same procedure as the one used for gender, testing a series of nested CFA models with increasing invariance restrictions (cf. Vandenberg & Lance, 2000). The invariance of the factor structure and the factor loadings across both groups was supported, providing evidence of the scale's invariant construct validity across both groups. Scalar invariance was partially supported, as the intercepts of items 2 and 6 had to be

Table 3

Fit Indices for Measurement Invariance Model Comparisons across Gender for Student Sample in Study 1 (N = 1,224; 774 women / 450 men)

Model	Model equalities	*df*	SB-Chi²	Compare with model	ΔSBc-Chi²[a]	CFI	RMSEA [90% CI]	SRMR
1	Number of factors	24	90.47	-	-	.95	.07 [.05-.08]	.04
2	NF; factor loadings	30	100.10	1	12.66	.94	.06 [.05-.08]	.06
3a	NF; FL; intercepts	36	140.53	2	46.74*	.92	.07 [.06-.08]	.07
3b	NF; FL; partial intercepts	35	110.81	2	9.31	.94	.06 [.05-.07]	.06
4	NF; FL; PIC; residual variances	42	120.72	3b	16.02	.94	.06 [.04-.07]	.10
5	NF; FL; PIC; RV; factor variance	43	121.42	4	2.52	.94	.06 [.04-.07]	.10
6	NF; FL; PIC; RV; FV; factor mean	44	123.54	5	1.58	.94	.05 [.04-.07]	.10

Note. * $p < .001$; NF = Number of factors; FL = factor loadings; PIC = partial intercepts; RV = residual variances; FV = factor variance.

[a] The Satorra-Bentler corrected Chi^2-difference test was computed using the ML chi-square, the degrees of freedom and the scaling correction factor of the two nested models (Satorra & Bentler, 2001)

Table 4

Fit Indices for Measurement Invariance Model Comparisons across Gender for Working Sample in Study 1 (N = 526; 310 women / 216 men)

Model	Model equalities	*df*	SB-Chi²	Compare with model	ΔSBc-Chi²[a]	CFI	RMSEA [90% CI]	SRMR
1	Number of factors	24	42.47	-	-	.98	.05 [.03-.08]	.04
2	NF; factor loadings	30	51.99	1	13.93	.97	.05 [.03-.08]	.06
3a	NF; FL; intercepts	36	104.17	2	75.60*	.91	.09 [.07-.10]	.10
3b	NF; FL; partial intercepts	35	64.93	2	14.46	.96	.05 [.04-.08]	.07
4a	NF; FL; PIC; residual variances	42	71.78	3b	11.95	.96	.05 [.03-.07]	.17[b]
4b	NF; FL; PIC; partial residual variances	39	63.42	3b	2.83	.97	.05 [.03-.07]	.08
5	NF; FL; PIC; PRV; factor variance	40	65.05	4b	3.21	.97	.05 [.03-.07]	.11
6	NF; FL; PIC; PRV; non-invariant FV; factor mean	41	68.95	4b	1.36	.97	.05 [.03-.07]	.11

Note. * $p < .001$; NF = Number of factors; FL = factor loadings; PIC = partial intercepts; RV = residual variances; FV = factor variance.

[a] The Satorra-Bentler corrected Chi²-difference test was computed using the ML chi-square, the degrees of freedom and the scaling correction factor of the two nested models (Satorra & Bentler, 2001)

[b] Increase in SRMR indicated that invariance of residual variances was not supported. Residual variances of three items had to be released in Model 4b.

released. Similarly, the assumption of invariant item uniqueness was partially supported, as the residual variances of item 2 had to be released to obtain a non-significant change in model fit. While the variance in the protean career orientation was higher for employees than for students, the final model comparison showed that there was no systematic difference between students and employees in their mean level of a protean career orientation as evidenced by their self-reports.

Study 2: Protean Career Orientation and Proactivity in Career Development: Establishing Measurement Invariance across Time and Incremental Predictive Validity

Career research often involves an investigation of developmental patterns. As such, the use of an assessment in longitudinal studies requires evidence of its measurement invariance across time (Horn & McArdle, 1992). Such evidence has not been demonstrated in previous research on protean career orientation; therefore, this longitudinal six-month study helps close this gap in the research literature. Evidence of measurement invariance is also needed to assess the differential stability of a protean career orientation because an examination of rank-order stability as a measure of relative reliability becomes meaningful only once one has demonstrated that the same construct is indeed being measured over time.

Study 2 also investigates the relationship between a protean career orientation and engagement in proactive career behaviors. Due to the dynamic nature of contemporary careers, proactive career behaviors (e.g., networking, planning, exploration) have gained increased attention in the career success literature (Fuller & Marler, 2009). Following Hall (1996), we expect that a protean career orientation is positively related to active engagement in proactive career behaviors because people with a protean orientation are more motivated to direct their careers according to their own values. For example, positive relationships have been found between a protean career orientation and career planning and career exploration among students and employees alike (Creed, Macpherson, & Hood, 2011; De Vos & Soens, 2008) as well as between a protean career orientation and a general disposition to be proactive (i.e., proactivity) (Creed, et al., 2011) among a sample of university students. In sum, existing theory and empirical research together imply a positive relationship between a protean career orientation and proactivity in career development. Existing research on this topic, however, is sparse and has relied on cross-sectional data. Moreover, extant research has not established whether an adoption of a protean career orientation is incrementally predictive of proactive career behaviors beyond a proactive personality disposition. Finally, based on the assumption that career-specific attitudes mediate the effects of general dispositions on career outcomes, it is possible that a protean career orientation mediates the effects of a proactive disposition on one's tendency to exhibit proactive career behaviors. However, extant research does not contain investigations of this hypothesis in particular.

In light of the previous discussion, we propose the following hypotheses:

Hypotheses 1: There is a positive correlation between a protean career orientation and (a) a proactive disposition, and (b) the engagement in proactive career behaviors.

Hypothesis 2: A protean career orientation partially mediates the effects of a proactive disposition on proactive career behaviors.

Hypothesis 3: A protean career orientation is predictive of proactive career behaviors beyond a proactive disposition.

To provide further evidence of the distinctness between a protean and traditional career orientation, the same mediation model as proposed by Hypothesis 2 was also tested with traditional career orientation as the mediating variable.

Materials and Method

Participants and procedure

The 1,224 participating students from Study 1 were invited to take part in a follow-up study and to provide their email to the study investigators for this purpose. The 887 students who agreed to participate were contacted again six months later, a time lag often used in career research (e.g., Kossek, Roberts, Fisher, & Demarr, 1998; Strauss, Griffin, & Parker, 2012). We achieved a response rate of 47% in this second wave of data collection, which is above the norm in behavioral sciences (Baruch & Holtom, 2008). The assessment of measurement invariance and differential stability over time of the protean career orientation was based on data from T1 and T2. To test the mediation model, proactivity and traditional career orientation were assessed at T1, and proactive career behaviors (i.e., career engagement) were evaluated at T2.

Of the 419 students who participated, 63.5% were female, with age M = 23.63 (SD = 2.75) and study semester M = 3.78 (SD = 2.05) at T1. Respondents were enrolled in 27 different majors overall, with almost two-thirds of them studying management and entrepreneurship (19%), business psychology (15%), business administration (14%), cultural studies (8%), or environmental science (7%). As suggested by Baruch and Holton (2008), the 419 respondents who completed both waves of data collection were compared to the 805 respondents who participated only in the first wave. No significant differences were found for any of the assessed variables.

Measures

Protean and traditional career orientation. The same measures used in Study 1 were used in Study 2. Cronbach's alpha was .67 at T1 and .71 at T2 for the protean career orientation measure and .79 at T1 and .81 at T2 for the traditional career orientation measure.

Proactivity. We measured participants' self-reported proactive disposition using the seven-item personal initiative questionnaire developed by Frese, Fay, Hilburger, and Leng (1997) (e.g., "I actively attack problems."). Cronbach's alpha for this scale was .79 at T1 in our sample.

Career engagement. The degree of engagement in proactive career behaviors was assessed using the career engagement scale (Hirschi, Freund, & Herrmann, in press). This tool contains nine items that measure the general degree to which someone has engaged in different career management behaviors (e.g., career planning, career exploration, networking, positioning behavior, voluntary training) within the last six months. Cronbach's alpha of the scale was .89 at T1 and .86 at T2.

Results and Discussion

Measurement invariance over time and differential stability

To determine the measurement invariance over time, a procedure analogous to that applied in Study 1 was carried out. Measurement invariance for longitudinal data were tested by applying equality constraints to model parameters across the two time points. The baseline model (Model 1) established in Study 1 was used as a starting point, albeit with two factors, each factor representing one point in time. All equality constraints applied to the respective models yielded non-significant results when compared to the previous, less-constrained model ($p > .05$). This finding confirms the measurement invariance of the protean career scale across time and indicates that the scale measures the same basic construct across time.

We assessed the differential stability of the protean career orientation across six months as the standardized covariance between the latent variables at both measurement points, controlling for measurement error of the scale. The obtained coefficient of .63 confirms moderate stability in the measure, which is in line with findings from other studies assessing the variability of work values during emerging adulthood (Jin & Rounds, 2012).

Protean career orientation and proactivity in career management

To test Hypotheses 1, 2, and 3, we used participants' responses on proactive disposition and protean and traditional career orientation at T1 while using their proactive career behaviors at T2. Such temporal separation can reduce the potential inflation caused by common method bias (MacKenzie & Podsakoff, 2012). To provide evidence that the independent variable, the mediator, and the dependent variable are not representative of the same latent construct (Fiedler, Schott, & Meiser, 2011), we applied CFA. The three-factor model displayed a significantly better model fit than that of the one-factor model (SB-corrected $\Delta\chi2 = 671.00$, $df = 3$; $p < .001$), supporting the assumption of distinct constructs.

As recommended by previous researchers (Preacher, Rucker, & Hayes, 2007; Shrout & Bolger, 2002), we applied a bootstrapping technique with 5,000 bootstrapping samples using an Mplus syntax for mediation provided by Preacher and colleagues (Preacher, Zyphur, & Zhang, 2010). Using this procedure, the indirect effect, its 95% confidence intervals, and the standard errors were computed. A path was significant if zero was not included in the confidence interval (Shrout & Bolger, 2002). The results revealed support for all three hypotheses (see Table 5). The bivariate correlations between the three variables were positive and highly significant ($p < .001$), confirming Hypothesis 1. Additionally, a significant indirect effect existed in the mediation model and a protean career orientation predicted proactive career behaviors beyond a

Table 5

Correlations and Mediation Bootstrap Analyses for Study 2 (Student Sample: N = 419) and Study 3 (Employee Sample: N = 526)

	Correlations[a]		IV	MV	DV	Indirect effect in mediation: Point estimate[a] (SE)	Percentile 95% CI: Lower	Percentile 95% CI: Upper	β[a]
Study 2: Student sample									
	PC-PAD: .45	PC-ENG: .32	PAD	PC	ENG	.086 (.026)	.035	.137	.193
Study 3: Employee sample									
	PC-PAD: .49	PC-ENG: .35	PAD	PC	ENG	.122 (.027)	.069	.175	.247
	PC-CSE: .44	PC-CSat: .45	CSE	PC	CSat	.130 (.029)	.074	.187	.294

Note: IV = independent variable, MV = mediating variable, DV = dependent variable, PAD = proactive disposition, PC = protean career orientation, ENG = engagement in proactive career behaviors, CSE = core-self evaluations, CSat = career satisfaction, SE = standard error, CI = confidence interval, β = effect of MV on DV while controlling for IV.

[a] All correlations / point estimates / β coefficients $p < .001$

proactive disposition, thus confirming Hypotheses 2 and 3, respectively. In contrast, no significant indirect effect for traditional career orientation was found when testing if such an orientation mediates the relationship between proactivity and proactive career behaviors.

The results confirm that a protean career orientation positively predicts engagement in proactive career behaviors among university students. Advancing extant research, our study provides support for the assumption that a protean career orientation partially mediates the effects of more basic personal dispositions on career outcomes, specifically between proactivity and proactive career behaviors. This result enriches the literature by clarifying how and why a protean career orientation is related to career outcomes. We have also shown that a protean career orientation possesses incremental validity in predicting proactive career management behaviors beyond the general disposition for proactivity. Finally, we could establish the different functioning of a protean career orientation in contrast to a traditional orientation by showing that only a protean but not a traditional orientation acts as a mediator between proactivity and career engagement.

Study 3: Incremental Utility and Mediating Effects on Proactive Career Behaviors and Career Satisfaction among Employees

Existing research on the protean career orientation has documented its relationship to diverse work and career outcomes among employees, such as organizational commitment (Briscoe & Finkelstein, 2009; Cakmak-Otluoglu, 2012), career management behaviors, or career satisfaction (De Vos & Soens, 2008). However, given the importance of personal dispositions regarding diverse career and work outcomes (e.g., Ilies, Fulmer, Spitzmuller, & Johnson, 2009; Judge, Heller, & Mount, 2002), it is important to investigate to what extent positive career outcomes of a protean career orientation are dependent upon a mutual relationship to more basic personal dispositions. This would establish (a) whether a protean career orientation can incrementally predict important career outcomes beyond that which can be predicted by more general personal dispositions, and (b) whether a protean career orientation would partially mediate the effects of personal dispositions on career outcomes.

In Study 3, we used a sample of employees to replicate the finding regarding proactivity and career behaviors obtained in Study 2 with a sample of students (Hypotheses 1 to 3). Second, we investigated the relationships among a protean career orientation, core self-evaluations (CSE), and career satisfaction. Because a protean career orientation implies guiding one's career according to one's own values to achieve subjective career success, it is generally assumed that a protean career orientation is positively related to career satisfaction (Hall, 2004; Hall & Mirvis, 1996). Supporting this assumption, empirical studies have repeatedly found a positive relationship between a protean orientation and subjective evaluations of career success (Briscoe, et al., 2012; De Vos & Soens, 2008; Park, 2009; Verbruggen & Sels, 2008). However, extant research does not contain investigations into whether a protean career orientation is related to career satisfaction beyond the effects of important personality dispositions.

Specifically, we consider CSE to be the "basic, fundamental appraisal of one's worthiness, effectiveness, and capability as a person" (Judge, Erez, Bono, & Thoresen, 2003, p. 304). As such, CSE represent the dispositional core of job satisfaction, and research has repeatedly confirmed their significant relationships with a range of job attitudes and career success (Judge & Kammeyer-Mueller, 2011). People with high CSE are assumed to be more ambitious and confident in their career and more actively engaged in self-initiated career planning as well as exploration and job searching (Judge & Kammeyer-Mueller, 2011). Thus, it is likely that CSE are positively related to a protean career orientation. However, previous research has not investigated this relationship, and it is important to establish the incremental validity of a protean career orientation above CSE when explaining career satisfaction. Finally, we investigate to what extent a protean career orientation mediates the relationship between CSE and career satisfaction to increase our knowledge of whether and how career orientations mediate the effects of personality dispositions on career outcomes, as previous studies have not examined such mediating effects. To summarize, we assumed that:

Hypothesis 4: There is a positive correlation between a protean career orientation and (a) CSE, and (b) career satisfaction.

Hypothesis 5: A protean career orientation partially mediates the relation between CSE and career satisfaction.

Hypothesis 6: A protean career orientation predicts career satisfaction beyond the effects of CSE.

As above, we also tested if a traditional career orientation mediates the relationship between CSE and career satisfaction.

Materials and Method

Participants and procedure

The same sample of working professionals (N = 526) as described in Study 1 was used.

Measures

Protean and traditional career orientation. The same measures as described in Study 1 were used in Study 3. Cronbach's alpha for the two scales was .77 and .80, respectively.

Proactivity. The seven-item measure of proactivity as described in Study 2 was used. Cronbach's alpha for the scale was .84 in our sample.

Career engagement. The nine-item measure of proactive career behavior described in Study 2 was used. Cronbach's alpha for the scale was .91.

Core self-evaluations (CSE). CSE were assessed with the 12-item German-language version of the CSE scale by Judge et al. (2003; Stumpp, Hülsheger, Muck, & Maier, 2009). Cronbach's alpha for the scale was .85.

Career satisfaction. We assessed career satisfaction with a German translation (Abele & Spurk, 2009) of the well-established career satisfaction scale (Greenhaus, Parasuraman, & Wormley, 1990). Cronbach's alpha for the scale in the present sample was .88.

Results and Discussion

Protean career orientation, proactivity, and proactive career behaviors

Before testing the mediation, we applied CFA to establish that proactive disposition, protean career orientation, and proactive career behaviors are three distinct constructs (Fiedler, et al., 2011). This assumption was supported as the three-factor model displayed a significantly better model fit than the one-factor model (SB-corrected $\Delta\chi 2$ = 843.94, df = 3; p < .001). To test the mediation, we applied the same approach as described in Study 2. Hypotheses 1 to 3 were confirmed (see Table 5). The results replicate the findings obtained in Study 2 and show that also among working professionals, a protean career orientation is positively related to proactive career behaviors and proactivity (Hypotheses 1a and 1b). Moreover, we confirmed the results from Study 2 among employees that a protean orientation partially mediates the effects of proactivity on proactive career behaviors (Hypothesis 2) and that the effects of a protean orientation on proactive career behaviors are beyond the mere effects of having a disposition for proactivity (Hypothesis 3). A considerable smaller indirect effect was found for traditional career orientation as a mediator of the relationship between proactivity and proactive career behaviors. This confirms the distinct functioning of protean and traditional orientations regarding career behaviors.

Protean career orientation, CSE, and career satisfaction

The same procedures were applied to test Hypotheses 4 to 6 regarding the relationships among CSE, protean career orientation, and career satisfaction. A comparison of the model fit of a one-factor model with that of a three-factor model demonstrated that the three constructs are distinct (SB-corrected $\Delta\chi^2 = 653.18$, $df = 3$; $p < .001$). All three hypotheses were confirmed (see Table 5). First, the three constructs were positively and highly correlated (all $p < .001$) (H4). Second, a protean career orientation mediated the relationship between CSE and career satisfaction (H5). Third, a protean career orientation predicted career satisfaction beyond CSE (H6). Finally, no significant indirect effect was found for traditional career orientation as a mediator of the relationship between CSE and career satisfaction.

These results advance existing research by showing that CSE are positively related to a protean career orientation. This finding confirms the notion that CSE are important in the current career environment because they promote a self-directed and values-driven orientation to work. Moreover, the incremental validity of a protean career orientation for predicting career satisfaction above CSE significantly advances current research on career outcomes. This study contributes to the literature by showing that a protean career orientation partially mediates the relationship between CSE and career satisfaction, and thus, provides new evidence for how the relationship between CSE and career outcomes can be explained.

Study 4: Establishing Measurement Invariance across Time and Addressing Issues of Temporal Precedence Linking Career Satisfaction and Protean Career Orientation

First, we aimed to replicate Study 2 regarding measurement invariance across time with a sample of employees. Compared to students, individuals with more work experience can be expected to have developed more stable career attitudes. In line with results showing stabilization of work values after entering work (Jin & Rounds, 2012), we expected the differential stability of a protean career orientation to be higher among employees compared to students.

Second, we wanted to elaborate on the findings of Study 3 and investigate more closely how a protean career orientation and career satisfaction are related. As explained in Study 3, research has consistently found positive correlations between a protean orientation and subjective career evaluations. Usually, this relationship is explained in the way that people with a protean orientation are more active in, and successful at guiding their career according to their own values, and hence, are more likely to achieve subjective career success (Hall & Mirvis, 1996). Indeed, a protean career orientation is commonly assumed to temporally precede subjective career success.

However, we postulated that subjective career success might equally well precede a protean career orientation because it enhances a sense of competence and confidence in one's career management capabilities. According to self-determination theory (Deci & Ryan, 1985), a sense of competence is essential to developing an intrinsic motivation in a given domain. Having a strong intrinsic motivation for career development appears pivotal in the protean career orientation because it implies an intrinsic motivation to manage one's career according to one's own values. Moreover, we can assume that a sense of confidence and competence for career management is essential to develop a protean career orientation because it relies on oneself and not the organization to guide one's career (Hall, 2004). In sum, we suggest that the sense of competence that results from the subjective evaluation of career success and the increased likelihood of reaching one's intrinsic career goals that emerges from a protean career orientation can lead to reciprocal effects between a protean career orientation and career satisfaction. However, extant research has not sufficiently acknowledged this possibility, and researchers have not yet tested this assumption due to a persistent reliance on cross-sectional data. In the present study, therefore, we used longitudinal data and cross-lagged analysis to test the following assumptions:

Hypothesis 7: A protean career orientation predicts an interindividual increase in career satisfaction.

Hypothesis 8: Career satisfaction predicts an interindividual increase in a protean career orientation.

Hypothesis 9: A reciprocal effect exists between a protean career orientation and career satisfaction.

As in the previous studies, we replicated results for a traditional career orientation. That is, we examined its relationship with career satisfaction to provide a comparison.

Materials and Method

Participants and procedure

The employee sample was recruited by inviting via email a unique sample of university alumni from two German universities (N = 703) to complete an online questionnaire, followed by two reminder emails each one week apart. This strategy yielded a final response rate of 50% (N = 352) at T1. These employees were contacted again after six months and asked to participate in a follow-up study, resulting in 156 responses at T2 (response rate 44%). Of these participants, 61.5% were female, with mean age M = 30.88 (SD = 6.75) at T1. Most of them had received either a Bachelor's degree (28%) or a Master's degree or equivalent (70%). Participants were employed in many different industry sectors, with the largest groups working in education (17%), business administration (14%), engineering (12%), information technologies (10%), and marketing and advertising (8%). Participation in a lottery drawing offering two prizes of 450€ each was offered as an incentive at each assessment point. No significant differences for any of the assessed variables were found between the respondents and the non-respondents from the first wave.

Measure

Protean and traditional career orientation. The same measures as described in Study 1 were used. Cronbach's alphas for the two scales in this study were .70 and .76 at T1, and .72 and .80 at T2 for the protean and traditional scales, respectively.

Career satisfaction. The five-item measure of career satisfaction described in Study 3 was used. Cronbach's alpha for this sample was .85 at T1 and .80 at T2.

Results and Discussion

Measurement invariance over time and differential stability

Measurement invariance over time for the sample of working professionals was examined by fitting a series of nested CFA models and comparing the model fit to the respective previous, less constrained model. Equality constraints to model parameters across the two time points were applied to test for invariance across time (for a more detailed description of the procedure and applied cut-off criteria see Study 2). Measurement invariance was confirmed at each step, supporting the measurement invariance of the protean career scale across time. Similar to our finding for the student sample, a comparison of the latent means showed that the mean level of protean career orientation was stable over a time period of six months. As hypothesized, the differential stability was higher for the working sample than for the student sample, with a standardized covariance of .75 between the two latent variables of protean career orientation at T1 and T2, compared to .63 for the student sample. Considered together with the invariance of the sample mean over time, these findings suggest a high relative stability.

Cross-lagged analysis of career satisfaction and protean career orientation

Next, we applied cross-lagged analysis to examine the relationships between a protean career orientation and career satisfaction across time. A series of nested cross-lagged models was tested, starting with the autoregressive model (Model 1) in which the stability of the constructs over time is assessed (Burkholder & Harlow, 2003). In Model 2, a cross-lagged path was added from the protean career orientation at T1 to career satisfaction at T2. This was reversed in Model 3 by specifying a path from career satisfaction at T1 to protean career orientation at T2. In the full cross-lagged model (Model 4), both cross-lagged paths were included to test for the reciprocal effects of both constructs.

The autoregressive paths in Model 1 suggested that both constructs were moderately stable over a time lag of six months, with a standardized regression path of .61 for protean career orientation and .67 for career satisfaction. Furthermore, a positive and significant correlation between a protean career orientation and career satisfaction at T1 was obtained ($r = .45$, $p < .001$), confirming Study 3 as well as other research (De Vos & Soens, 2008). To test Hypotheses 7 and 8, we compared the model fit of both cross-lagged models (Model 2 and Model 3) with the autoregressive model, Model 1 (SB-$\chi^2 = 25.15$, $df = 3$, $p < .001$). The path from protean career orientation at T1 to career satisfaction at T2 in Model 2 was found to be non-significant and did not improve the model fit (SB-$\chi^2 = 21.23$, $df = 2$, $p < .001$; SB-corrected $\Delta\chi^2 = 4.36$, $df = 1$; $p > .01$), refuting Hypothesis 7 and indicating that a protean career orientation did not predict an interindividual increase in career satisfaction over six months. Model 3 displayed a significantly better fit than Model 1 (SB-$\chi^2 = 12.06$, $df = 2$, $p < .001$; SB-corrected $\Delta\chi^2 = 13.00$, $df = 1$; $p < .001$) and a significant path between career satisfaction at T1 to protean career orientation at T2 ($\beta = .237$, $p < .001$), confirming Hypothesis 8 and suggesting that career satisfaction predicts an interindividual increase in a protean career orientation over six months. Finally, the full cross-lagged model did not improve the fit over Model 3 (SB-$\chi^2 = 7.69$, $df = 1$, $p < .001$; SB-corrected $\Delta\chi^2 = 3.82$, $df = 1$; $p > .01$), suggesting no reciprocal effects over time and refuting Hypothesis 9.

When using the same procedure to investigate temporal relationships between a traditional career orientation and career satisfaction, we found no significant improvement in model fit and no significant paths between constructs across time. Furthermore, the very small and non-significant correlation between both constructs at T1 ($r = .06$, n.s.) suggests that both constructs are not related. This confirms the distinct contribution of a protean career orientation compared to a traditional career orientation relative to career satisfaction. In sum, the results significantly advance contemporary knowledge by going beyond reporting mere cross-sectional relations between a protean career orientation and subjective evaluations of career success. Specifically, our cross-lagged study provides new insights into how this relationship can be explained. We found no support for the dominant assumption that a protean career orientation leads to subjective career success. Conversely, our results rather imply that career satisfaction strengthens a protean career orientation over time. As we have argued above, this finding might be explained in the way that career satisfaction increases confidence and intrinsic motivation, both of which are key components to developing a protean career orientation.

General Discussion

Cumulatively, the four studies presented here illustrate several key theoretical contributions. First, we addressed the relationship between protean and traditional career orientations. We found that the two orientations do not simply represent opposites, but rather, two independent yet positively related career orientations. This is an important contribution to the literature, which often depicts protean and traditional as indicative of contrasting career paths and orientations (Hall, 1996). Our results imply that people can be self-directed and values-driven while simultaneously valuing traditional career outcomes such as salary and status. However, we showed that only the protean career orientation mediates the effects of personality variables on career outcomes. This finding indicates that the two career orientations may serve different functions. Apparently, the protean orientation has stronger links to personality, career attitudes and behaviors than the traditional orientation, which highlights the importance of new career orientations in the current career context (Hall, 2004). Second, we addressed contradictions in extant research regarding the potential gender differences in endorsements of a protean career orientation. After documenting measurement invariance across gender, we could then show that male and female university students and professionals did not significantly differ in the degree to which they exhibited a protean career orientation. This finding implies that while men and women might exhibit different values in their career pursuits, (e.g., salary vs. spending time with family; Konrad, et al., 2000), the preference for being self-directed, flexible, and values-driven is equally present between both genders. Third, we established that a protean career orientation is relatively stable over a significant length of time, advancing current research that had not yet addressed the important issue of temporal stability of career orientations due to previous studies' reliance on cross-sectional data. Moreover, in line with research on the stability of work values (Jin & Rounds, 2012), we showed that the protean orientation is more stable after entering the workforce than it is while being enrolled at a university, stressing the importance of the often neglected employment context in shaping career orientations (Inkson, et al., 2012). Fourth, we showed that a protean orientation partially mediates the effects of personality dispositions on proactive career behaviors and career satisfaction and that it is incrementally predictive above personality regarding these outcomes. This significantly advances extant research, which had previously assessed neither mediation nor the incremental validity of new career orientations in relation to the established predictive validity of personality characteristics (Fuller & Marler, 2009; Judge & Kammeyer-Mueller, 2011). Finally, our findings advance the research literature linking protean orientation and career success (e.g., De Vos & Soens, 2008) by showing that one should not simply think of a protean orientation as being merely an antecedent of career outcomes, but that career outcomes might also shape career orientations. Specifically, career satisfaction seems to promote a protean career orientation, and researchers should consider addressing this new finding more readily in future research.

Limitations and Suggested Future Research

While we sampled participants from educational institutions as well as industry, participants came from a young, university-educated German population. This homogeneity may have impacted the generalizability of our findings. Therefore, future studies of the protean career orientation should cover an expanded range of populations from different countries to address this potential limitation. Another limitation is that the mediations tested in Study 3 were evaluated using cross-sectional data, which prevents one from making causal inferences about the relationship between a protean career orientation and career outcomes. In fact, the results from Study 4 showed that researchers must be careful when making such causal claims based on cross-sectional data as the causal relationship inferred between constructs could actually be in the opposite direction to what is commonly assumed.

However, even though Study 4 applied a cross-lagged design to test the question of temporal precedence of a protean career orientation and career satisfaction, for several reasons one must still be careful when interpreting the results. First, we relied on a relatively small sample, which might have biased the results. Second, causal inferences might be wrong because the assessed variables had not reached equilibrium or because other influential variables might have been missing from the model. Rigorous experimental manipulation of a protean career orientation would be necessary to add certainty about the true causal influences. Related to this point, it would be interesting for future studies to investigate how a protean career orientation could be fostered through targeted career interventions. Our results suggest that a protean career orientation is somewhat less stable over time for students enrolled in college, implying that this group may be suitable for conducting intervention research. Future research should also examine more closely how the work context and work experiences might lead to the development of a protean career orientation. Furthermore, it remains to be explored how a protean career orientation may actually influence individuals' career choices and career paths. Our study provides an important starting point for such lines of inquiry.

Practical Implications

Findings from this study also have implications for the individual and organizational management of careers. Many functions of human resource management are practically conducted by line managers (Purcell & Hutchinson, 2007). Therefore, managers should be aware of their employees' attitudes, and thus, should be given the tools needed to understand, advice, and support their employees regarding career planning and management. The brevity of our applied measure makes it particularly suitable in situations in which time is limited but where insight into an employees' attitude is needed. Assessing the levels of protean career orientation within an organization's work force is of interest because it may indicate if employees tend feel responsible for developing their own careers or if they are having certain expectations for their organization to take ownership of their professional development. This would be important information in the context of career consultations for individuals, either by organizations or career consultants. An important finding for career centers at universities is that the level of protean career orientation is less stable among students. Thus, this may be an ideal time for career counselors to help students develop a more self-directed career attitude and prepare them for the demands of today's work environment.

Conclusions

The study of 'new careers' is still in its early stages (Inkson, et al., 2012), but scholars should have valid and rigorous measures available for assessing novel concepts in different languages. This would enable further developments in career studies and would help avoid conceptual ambiguity that hinders further developments in the field (Arthur, et al., 1989; Greenhaus, Callanan, & DiRenzo, 2008). In this paper, we have made two general contributions. First, we provided extensive evidence of the reliability and validity of a scale for assessing a protean and traditional career orientation. Second, we contributed to the literature by examining the role that a protean career orientation plays in the relationship between personality and career outcomes. In doing so, we answered calls for advancing career studies and the relevance of the so-called 'new careers' through rigorous empirical evaluations of newly emerging constructs (Arthur, 2008; Inkson, et al., 2012; Sullivan & Baruch, 2009).

References

Abele, A. E., & Spurk, D. (2009). The longitudinal impact of self-efficacy and career goals on objective and subjective career success. *Journal of Vocational Behavior, 74*(1), 53-62. doi: 10.1016/J.Jvb.2008.10.005

Arnold, J., & Cohen, L. (2008). The Psychology of Careers in Industrial and Organizational Settings: A Critical But Appreciative Analysis. In G. P. Hodgkinson & J. K. Ford (Eds.), *International Review of Industrial and Organisational Psychology* (pp. 1-44). London: Wiley.

Arthur, M. B. (2008). Examining contemporary careers: A call for interdisciplinary inquiry. *Human Relations, 61*(2), 163-186. doi: 10.1177/0018726707087783

Arthur, M. B., Hall, D. T., & Lawrence, B. S. (1989). Generating new directions in career theory: The case for a transdisciplinary approach. In M. B. Arthur, D. T. Hall & B. S. Lawrence (Eds.), *Handbook of career theory* (pp. 7-25). Cambridge Cambridge University Press.

Arthur, M. B., & Rousseau, D. M. (1996). *The boundaryless career*. New York: Oxford University Press.

Baruch, Y. (2008). *Development and validation of a measure for protean career*. Paper presented at the 68th Annual Meeting of the Academy of Management, Anaheim, CA.

Baruch, Y., Bell, M. P., & Gray, D. (2005). Generalist and specialist graduate business degrees: tangible and intangible value. *Journal of Vocational Behavior, 67*(1), 51-68. doi: 10.1016/j.jvb.2003.06.002

Baruch, Y., & Holtom, B. C. (2008). Survey response rate levels and trends in organizational research. *Human Relations, 61*(8), 1139-1160. doi: 10.1177/0018726708094863

Baruch, Y., & Quick, J. C. (2007). Understanding second careers: lessons from a study of US Navy admirals. *Human Resource Management, 46*(4), 471-491. doi: 10.1002/hrm.20178

Briscoe, J. P., & Finkelstein, L. M. (2009). The "new career" and organizational commitment: Do boundaryless and protean attitudes make a difference?

Career Development International, 14(2-3), 242-260. doi: 10.1108/13620430910966424

Briscoe, J. P., Hall, D. T., & DeMuth, R. L. F. (2006). Protean and boundaryless careers: An empirical exploration. *Journal of Vocational Behavior, 69*(1), 30-47. doi: 10.1016/J.Jvb.2005.09.003

Briscoe, J. P., Henagan, S. C., Burton, J. P., & Murphy, W. M. (2012). Coping with an insecure employment environment: The differing roles of protean and boundaryless career orientations. *Journal of Vocational Behavior, 80*(2), 308-316. doi: 10.1016/j.jvb.2011.12.008

Burkholder, G. J., & Harlow, L. L. (2003). An illustration of a longitudinal cross-lagged design for larger structural equation models. *Structural Equation Modeling, 10*(3), 465-486. doi: 10.1207/S15328007SEM1003_8

Byrne, B. M., Shavelson, R. J., & Muthen, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*(3), 456-466. doi: 10.1037/0033-2909.105.3.456

Cakmak-Otluoglu, K. O. (2012). Protean and boundaryless career attitudes and organizational commitment: The effects of perceived supervisor support. *Journal of Vocational Behavior, 80*(3), 638-646. doi: 10.1016/j.jvb.2012.03.001

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research, 1*(2), 245-276. doi: 10.1207/s15327906mbr0102_10

Cerdin, J. L., & Pargneux, M. L. (2009). Career and international assignment fit: Toward an integrative model of success. *Human Resource Management, 48*(1), 5-25. doi: 10.1002/hrm.20264

Chan, K. Y., Ho, M. H. R., Chernyshenko, O. S., Bedford, O., Uy, M. A., Gomulya, D., . . . Phan, W. M. J. (2012). Entrepreneurship, professionalism, leadership: A framework and measure for understanding boundaryless careers. *Journal of Vocational Behavior, 81*(1), 73-88. doi: 10.1016/j.jvb.2012.05.001

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464-504. doi: 10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255. doi: 10.1207/S15328007SEM0902_5

Creed, P., Macpherson, J., & Hood, M. (2011). Predictors of "new economy" career orientation in an Australian sample of late adolescents. *Journal of Career Development, 38*(5), 369-389. doi: 10.1177/0894845310378504

De Vos, A., & Soens, N. (2008). Protean attitude and career success: The mediating role of self-management. *Journal of Vocational Behavior, 73*(3), 449-456. doi: 10.1016/j.jvb.2008.08.007

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human development*. New York: Plenum.

Enache, M., Sallan, J. M., Simo, P., & Fernandez, V. (2011). Examining the impact of protean and boundaryless career attitudes upon subjective career success. *Journal of Management & Organization, 17*(4), 459-473. doi: 10.5172/jmo.2011.17.4.459

Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology, 47*(6), 1231-1236. doi: 10.1016/j.jesp.2011.05.007

Frese, M., Fay, D., Hilburger, T., & Leng, K. (1997). The concept of personal initiative: Operationalization, reliability and validity of two german samples. *Journal of Occupational and Organizational Psychology, 70*(2), 139-161. doi: 10.1111/j.2044-8325.1997.tb00639.x

Fuller, B., Jr., & Marler, L. E. (2009). Change driven by nature: A meta-analytic review of the proactive personality literature. *Journal of Vocational Behavior, 75*(3), 329-345. doi: 10.1016/j.jvb.2009.05.008

Gerber, M., Wittekind, A., Grote, G., & Staffelbach, B. (2009). Exploring types of career orientation: A latent class analysis approach. *Journal of Vocational Behavior, 75*(3), 303-318. doi: 10.1016/j.jvb.2009.04.003

Gerbing, D. W., & Anderson, J. C. (1984). On the Meaning of Within-Factor Correlated Measurement Errors. *Journal of Consumer Research, 11*(1), 572-580. doi: 10.1086/208993

Gill, C. M., & Hodgkinson, G. P. (2007). Development and validation of the five-factor model questionnaire (FFMQ): An adjectival-based personality inventory for use in occupational settings. *Personnel Psychology, 60*(3), 731-766. doi: 10.1111/j.1744-6570.2007.00090.x

Greenhaus, J. H., Callanan, G. A., & DiRenzo, M. (2008). A boundaryless perspective on careers. In C. L. Cooper & J. Barling (Eds.), *Handbook of organizational behavior* (pp. 277-299). Thousand Oaks: Sage.

Greenhaus, J. H., Parasuraman, S., & Wormley, W. M. (1990). Effects of race on organizational experiences, job performance evaluations, and career outcomes. *The Academy of Management Journal, 33*(1), 64-86. doi: 10.2307/256352

Hall, D. T. (1996). Protean careers of the 21st century. *The Academy of Management Executive, 10*(4), 8-16. doi: 10.5465/AME.1996.3145315

Hall, D. T. (2004). The protean career: A quarter-century journey. *Journal of Vocational Behavior, 65*(1), 1-13. doi: 10.1016/j.jvb.2003.10.006

Hall, D. T., & Mirvis, P. H. (1996). The new protean career: Psychological success and the path with a heart. In D. T. Hall & Associates (Eds.), (pp. 15-45). San Francisco, CA: Jossey-Bass.

Hall, D. T., & Moss, J. E. (1998). The new protean career contract: Helping organizations and employees adapt. *Organizational Dynamics, 26*(3), 22-37. doi: 10.1016/S0090-2616(98)90012-2

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*(5), 967-988. doi: 10.1177/014920639502100509

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1*(1), 104-121. doi: 10.1177/109442819800100106

Hirschi, A., Freund, P. A., & Herrmann, A. (in press). The career engagement scale: Development and validation of a measure for proactive career behaviors. *Journal of Career Assessment*.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179-185. doi: 10.1007/bf02289447

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*(3), 117-144. doi: 10.1080/03610739208253916

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55. doi: 10.1080/10705519909540118

Ilies, R., Fulmer, I. S., Spitzmuller, M., & Johnson, M. D. (2009). Personality and citizenship behavior: The mediating role of job satisfaction. *Journal of Applied Psychology, 94*(4), 945-959. doi: 10.1037/a0013329

Inkson, K., Gunz, H., Ganesh, S., & Roper, J. (2012). Boundaryless careers: bringing back boundaries. *Organization Studies, 33*(3), 323-340. doi: 10.1177/0170840611435600

Jin, J., & Rounds, J. (2012). Stability and change in work values: A meta-analysis of longitudinal studies. *Journal of Vocational Behavior, 80*(2), 326-339. doi: Doi 10.1016/J.Jvb.2011.10.007

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The Core Self-Evaluations Scale. Development of a measure. *Personnel Psychology, 56*(2), 303-331. doi: 10.1111/j.1744-6570.2003.tb00152.x

Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology, 87*(3), 530-541. doi: 10.1037/0021-9010.87.3.530

Judge, T. A., & Kammeyer-Mueller, J. D. (2011). Implications of core self-evaluations for a changing organizational context. *Human Resource Management Review, 21*(4), 331-341. doi: 10.1016/j.hrmr.2010.10.003

Kahn, J. H. (2006). Factor Analysis in Counseling Psychology Research, Training, and Practice: Principles, Advances, and Applications. *The Counseling Psychologist, 34*(5), 684-718. doi: 10.1177/0011000006286347

Konrad, A. M., Ritchie, J. E., Lieb, P., & Corrigall, E. (2000). Sex differences and similarities in job attribute preferences: A meta-analysis. *Psychological Bulletin, 126*(4), 593–641. doi: 10.1037/0033-2909.126.4.593

Kossek, E. E., Roberts, K., Fisher, S., & Demarr, B. (1998). Career self-management: A quasi-experimental assessment of the effects of a training intervention. *Personnel Psychology, 51*(4), 935-960. doi: 10.1111/j.1744-6570.1998.tb00746.x

MacKenzie, S. B., & Podsakoff, P. M. (2012). Common Method Bias in Marketing: Causes, Mechanisms, and Procedural Remedies. *Journal of Retailing, 88*(4), 542–555. doi: 10.1016/j.jretai.2012.08.001

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly, 35*(2), 293-334.

Mainiero, L. A., & Sullivan, S. E. (2005). Kaleidoscope careers: An alternate explanation for the" opt-out" revolution. *The Academy of Management Executive, 19*(1), 106-123. doi: 10.5465/AME.2005.15841962

Park, Y. (2009). An integrative empirical approach to the predictors of self-directed career management. *Career Development International, 14*(6-7), 636-654. doi: 10.1108/13620430911005690

Plouffe, C. R., & Grégoire, Y. (2011). Intraorganizational employee navigation and socially derived outcomes: Conceptualization, validation, and effects on overall performance. *Personnel Psychology, 64*(3), 693-738. doi: 10.1111/j.1744-6570.2011.01223.x

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42*(1), 185-227. doi: 10.1080/00273170701341316

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209. doi: 10.1037/a0020141

Purcell, J., & Hutchinson, S. (2007). Front-line managers as agents in the HRM-performance causal chain: theory, analysis and evidence. *Human Resource Management Journal, 17*(1), 3-20. doi: 10.1111/j.1748-8583.2007.00022.x

Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment, 29*(4), 347-363. doi: 10.1177/0734282911406661

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514. doi: 10.1007/Bf02296192

Segers, J., Inceoglu, I., Vloeberghs, D., Bartram, D., & Henderickx, E. (2008). Protean and boundaryless careers: A study on potential motivators. *Journal of Vocational Behavior, 73*(2), 212-230. doi: 10.1016/j.jvb.2008.05.001

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*(4), 422-445. doi: 10.1037//1082-989X.7.4.422

Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*(1), 78-90. doi: 10.1086/209528

Strauss, K., Griffin, M. A., & Parker, S. K. (2012). Future work selves: How salient hoped-for identities motivate proactive career behaviors. *Journal of Applied Psychology, 97*(3), 580-598. doi: 10.1037/a0026423

Stumpp, T., Hülsheger, U. R., Muck, P. M., & Maier, G. W. (2009). Expanding the link between core self-evaluations and affective job attitudes. *European Journal of Work and Organizational Psychology, 18*(2), 148-166. doi: 10.1080/13594320802281110

Sullivan, S. E., & Baruch, Y. (2009). Advances in career theory and research: a critical review and agenda for future exploration. *Journal of Management, 35*(6), 1542-1571. doi: 10.1177/0149206309350082

van de Vijver, F. J., & Leung, K. (1997). *Methods and data analysis for cross-cultural research* (Vol. 1). Thousand Oaks, CA: Sage.

Vandenberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods, 3*(1), 4-70. doi: 10.1177/109442810031002

Verbruggen, M., & Sels, L. (2008). Can career self-directedness be improved through counseling? *Journal of Vocational Behavior, 73*(2), 318-327. doi: 10.1016/j.jvb.2008.07.001

Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under- and overextraction on principal axis factor analysis with varimax rotation. *Psychological Methods, 1*(4), 354-365. doi: 10.1037/1082-989X.1.4.354