European Journal of Personality, *Eur. J. Pers.* **28**: 300–310 (2014)
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/per.1964

Moving Personality Development Research Forward: Applications Using Structural Equation Models

JOSHUA J. JACKSON[1]* and MATHIAS ALLEMAND[2]

[1]*Department of Psychology, Washington University in St. Louis, St. Louis, MO, USA*
[2]*Department of Psychology, University of Zurich, Zurich, Switzerland*

Abstract: As evidenced by this special issue, research in the field of personality development is moving beyond merely describing patterns of development to more complex questions, such as the mechanisms by which personality develops or remains stable. In this article, we briefly describe three useful longitudinal structural equation models: latent growth curve models, latent change models and growth mixture models. Various underexamined and novel questions in the field of personality development are then described, each in the context of the longitudinal models. In doing so, we describe shortcomings of standard models and highlight questions and designs that could continue to move personality development research forward. Copyright © 2014 European Association of Personality Psychology

Key words: personality development; personality change; latent growth curve model; latent change model; longitudinal methods; person–environment transactions; multilevel analysis

The field of personality development is quickly maturing, such that researchers are moving beyond merely charting patterns of development, to examining the mechanisms that drive development. This article provides a brief, non-technical overview of various longitudinal structural equation models (SEM). We first provide a brief introduction to latent growth curve modelling (LGCM) and then describe latent change and growth mixture models. This overview is conceptual, but we include references that provide more details about the models. Additionally, Mplus code is available in an online supplement[1] so that readers can easily apply the models to their own datasets. After discussing these SEM models, we identify a number of questions that are currently underexamined in the study personality development. In doing so, we provide examples of empirical studies that have used these models.

Longitudinal structural equation models

The basic focus of any developmental model is to estimate change and or stability in the construct of interest. For example, do people increase, decrease or stay the same in conscientiousness during adulthood? Researchers may ask whether the sample, pooled across all participants, changes across time and the degree to which, they do so uniformly, as well as how much any particular individual in the sample changes or stays the same. Examining each individual rather than just pooling across the entire sample is referred to as estimating interindividual variability in intraindividual change (Nesselroade & Baltes, 1974)—and is the basis for the longitudinal models discussed in this article. To accomplish this task, we wish to fit *trajectories* of change for each individual that describe how a person changes across a particular time period. In their most basic form, individual trajectories are linear and change monotonically, but they can take a variety of nonlinear forms to indicate growth spurts, accelerated declines and many other developmental patterns (Grimm & Ram, 2009; Ram & Grimm, 2007).

There are two common methods of estimating basic individual trajectories: multilevel models (synonymously referred to as hierarchical linear models and random effects models; Raudenbush & Bryk, 2002; Singer & Willett, 2003; West, Ryu, Kwok, & Cham, 2011) and SEMs (Bollen, 1998; Bollen & Curran, 2006). Under some conditions, these models are equivalent (Hertzog & Nesselroade, 2003). However, SEM has the ability to limit measurement error through construction of latent variables, which makes it especially attractive for personality research. Though the two methods are distinct, the same terminology is often used for aspects specific to them, resulting in confusion as to which is being applied and as well as which would be optimal to apply in any given situation. Individual trajectories are sometimes referred to as growth trajectories or curves or growth curves, but so are the overall trajectories, and both models may be referred to as either growth models, curve models or growth curve models. In the current article, we focus on estimating trajectories through SEM given the ability of the framework to fit advanced models especially suitable for questions concerning personality development (McArdle, 2009), in addition to accounting for measurement error.

*Correspondence to: Joshua J. Jackson, Department of Psychology Washington University in St. Louis, 1 Brookings Drive, St. Louis, MO 63130.
E-mail: j.jackson@wustl.edu
[1]http://pages.wustl.edu/personality

Received 30 October 2013
Revised 30 April 2014, Accepted 30 April 2014

Latent growth curve models

Fitting LGCMs is relatively straightforward using any SEM programme (e.g. AMOS, Mplus, lavaan and OpenMX,). In the most basic LGCM, two latent factors (intercept and change) are defined from a series of repeated measures (Figure 1). A typical minimal requirement to identify these latent factors is that there be at least three repeated measurements per individual (though this requirement may be relaxed; see Specht, Egloff, & Schmukle, 2011 for a treatment using two time points). Often, the latent *intercept* factor describes the starting value at the first wave of measurement (though this can be modelled differently, refer to the succeeding texts) whereas the latent *change* factor describes the change per unit of assessed time. The idea is that people start with some base level (the intercept) and then may change from one assessment to the next, the form of which is indicated by the change factor. To fit this, the latent intercept factor is scaled by fixing all loadings from the factor to the repeated measures at 1. The latent change factor is typically scaled by fixing the loading at the first wave of measurement to 0, at the second wave to 1, the third wave to 2 and so on in equally spaced intervals (assuming a linear trajectory). Thus, the latent intercept factor influences the repeated measure equally across time and implies an overall level, whereas the latent change factor represents a trajectory of change from that level. When the LGCM is scaled in this manner, the mean of the intercept factor reflects values at the first assessment, whereas the mean of the change factor reflects the average rate of change between two measurement occasions. Each of these mean estimates for the intercept, and change factors can be tested against a null hypothesis that the average value at the first repeated measure is zero (or any other value) and that there are no mean level changes, respectively. These estimates are often referred to as fixed effects, which can be contrasted with random effects that index the variability around the intercept and around the average change (refer to the succeeding texts).

The basic LGCM design depicted in Figure 1 can be elaborated in several ways, though many require an increase in the number of assessment points available. One may also be interested in knowing if changes occur more readily during certain time periods. Questions of nonlinear growth can be easily examined by introducing quadratic and cubic latent change factors to the model. Additionally, more *ad hoc* investigations of nonlinear growth can be examined by freely estimating the slope factor loadings, assuming the model is still identified (refer to McArdle & Bell, 2000 for more details). This technique allows the data to describe when the majority of the changes occur. A recent study employed this technique and found that in a three-wave study of young adults, most of the changes in personality traits occurred between the first and second waves rather than the changes in personality being equivalently spaced across measurement waves (Neyer & Lehnart, 2007). However, while a form of change may be found at the group level, this pattern might not hold for every participant. Use of LGCM requires the assumption that each participant changes in such a fashion, which may be untenable. It should be noted that any model-implied trajectory, including the linear trajectories, requires everyone follows the same path of development. Some of the variation from this model-implied pattern will be evidenced in the fit of the model and included in the variance of the change factor. Researchers should consider what pattern of change is most likely to represent both the overall sample trajectory as well as individual trajectories by examining the raw individual trajectories.

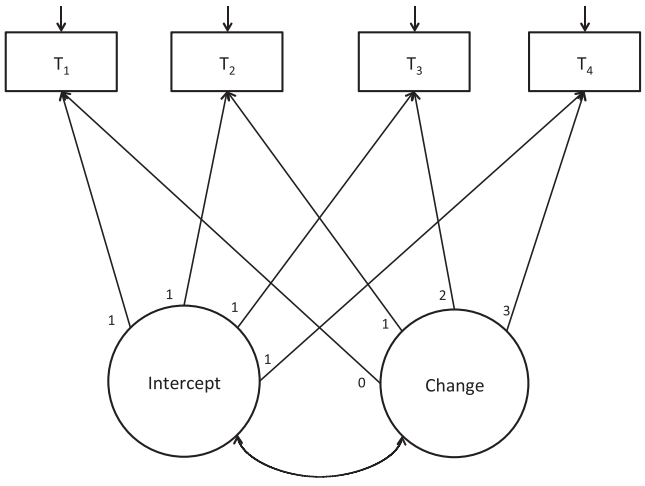

Figure 1. Univariate latent curve model.

Even if the model-implied trajectory fits the majority of the individual trajectories, that does not indicate that everyone in the sample changes in the same amount or direction (Roberts & Mroczek, 2008). Variance of the intercept and change factors reflects the extent of individual differences in level and change, and it is often these differences that are most interesting to personality researchers. For personality traits, numerous studies find significant variance in the change factor, indicating that sizeable proportions of individuals do not follow the normative patterns of personality trait change. These individual differences in change are found across a number of personality traits and appear to occur across the lifespan (e.g. Scollon & Diener, 2006; Small et al., 2003; Van Aken et al., 2006).

As in multiple regression models, covariates can be introduced into the model to estimate associations with the intercept and change factors. Models such as these are often referred to as *conditional models*, given that the intercept and slope factors now take into account these associations. For such a model, a factor must have significant variance given that the purpose is to explain individual differences in the factor. While dealing with covariates, it is important to differentiate between two different types. *Time-invariant* covariates consist of variables that do not change over time (e.g. date of birth) or variables that are only assessed once even though they may change across time (e.g. initial relationship satisfaction). In contrast, *time varying* covariates differ across time and may serve as developmental constructs themselves (e.g. relationship satisfaction assessed across multiple waves). Thus far most of the work in personality development has focused on time-invariant covariates. For example, timing of or circumstances surround discrete life

DOI: 10.1002/per

events such as one's first job, childbirth, divorce, starting therapy and studying abroad may be associated with individual differences in personality trait change (e.g. Ludtke et al., 2011; Specht et al., 2011).

It is also possible to estimate two or more separate LGCMs simultaneously. Figure 2 presents a simplified schematic of such a bivariate or dual-process model. The resulting bivariate LGCM can test various questions that are discussed in more detail later: are the starting values of each construct related, are initial levels of experience associated with subsequent changes in personality or *vice versa*, and, especially relevant to theories of person–environment transactions, are changes in experiences associated with changes in personality?

Despite the many promising questions that LGCM can address, there are numerous questions important to personality development that cannot be adequately addressed with LGCMs. Many of the limitations of LGCM stem from assumptions about the trajectory of change, namely that the trajectory reflects a relatively uniform long-term developmental process. A primary byproduct of this assumption is it is difficult to test whether there are time specific or directional associations between personality and experience (McArdle, 2009). The existence of correlated changes in two LGCMs, such as changes in personality and changes in environment, cannot identify whether the two are causally related to one another because of possible unmeasured processes. Some researchers attempt to overcome this limitation by using the intercept in a bivariate LGCM as a predictor of change in personality (or *vice versa*). This approach has shortcomings that mirror those of auto-regressive or cross-lagged designs (Rogosa, 1980) and likely extend to all observational trajectory models. For example, the direct relationship between experience and change in personality may be spurious because of unmeasured experiences or unmeasured growth processes. Use of the model also assumes that the initial measurement point is a meaningful point in time (Gollob & Reichardt, 1987). Instead, the selection of this initial point in time may have been arbitrary or convenient and thus caught in the middle of a more complex bidirectional or multidirectional process, resulting in biased estimates of supposed directional effects. Even if the initial time point is developmentally meaningful, it is difficult to establish the direction of effect given the multidetermined natures of our variables of interest. As a result, likely no observational longitudinal model will be able to isolate the direction of effect completely. To establish true causal relationships, it is necessary to rule out alternative hypotheses and replicate effects in different samples while using many different designs, ideally experimental.

Another byproduct of the assumption that LGCMs assess long-term trajectories that slowly change across months, years or decades is the inability to examine processes of short-term change. When discrete life events hypothesized to be predictors of change, the LGCM must accommodate the possibility the long-term trajectory of personality development is thrown-off-track, resulting in a new long-term trajectory. While this view of how environments influence personality development may be appropriate under some theoretical models, these models often neglect the potential for resiliency or coping and adjustment to new situations and experiences. Many experiences may affect personality (in terms of the manifestations of thoughts, feelings and behaviours), but effects may be short-lived, wane in importance over time or counteracted by another process, with considerable individual differences in the applicability of each (Luhmann, Orth, Specht, Kandler, & Lucas, 2014). LGCMs currently cannot adequately model such short(er)-term, time specific bidirectional processes. For many of these types of questions, multilevel model may be more appropriate to test these questions (e.g. Bolger & Laurenceau, 2013).

The assumption that the basic trajectory is uniformly shaped is particularly limiting when the trajectory modelled is linear, as it often is due to limited numbers of assessments typical in personality research. The linear model is particularly

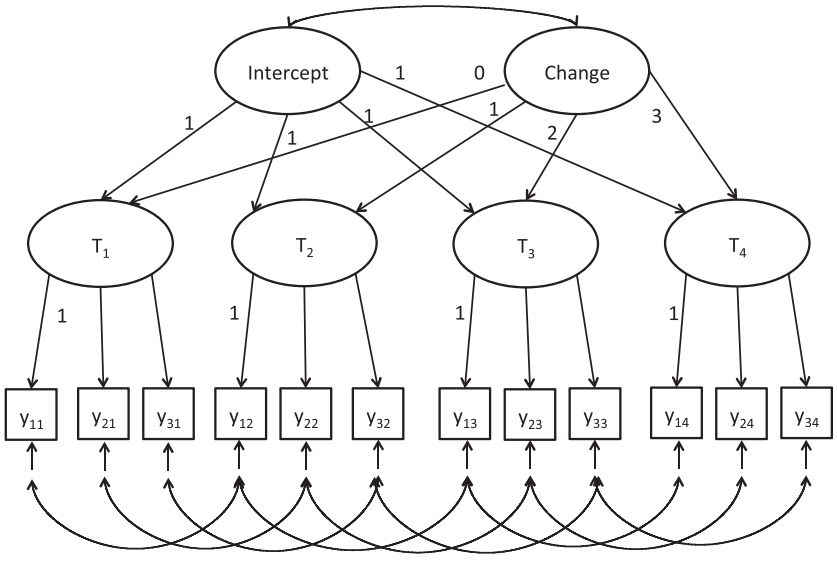

Figure 2. Second-order latent curve model.

DOI: 10.1002/per

poor at capturing situations where some individuals increase and others decrease or show not change. In an attempt to overcome these limitations, a number of techniques have recently been advanced that extend beyond the typical LGCM framework to address these complexities of development.

Latent change score models

Among the most promising techniques to overcome, some of these shortcomings are latent change score models (Ferrer & McArdle, 2010; McArdle & Hamagami, 2001). These models borrow many elements of LGCMs to create flexible systems that accommodate dynamic patterns of development. The unique components of these models are direct pathways between assessments points and a latent change score between each wave. The latent change scores between each wave allow for modelling of more complex patterns of development, such as increases followed by decreases in subsequent waves. Furthermore, these wave-to-wave change scores are used to create an overall latent change factor, similar to the latent change factor in LGCM in that it assesses overall change across the entire study period. As with a LGCM, there is an intercept parameter that reflects standing at the initial wave. The ability to model time specific changes and overall change across the study period allows much more complex nonlinear trajectories to be fit, than those that can be examined with LGCMs. As a result, these models are specifically useful to examine short-term processes or dynamic aspects of personality.

As with LGCMs, multiple development processes can be assessed simultaneously in the latent change score framework. In addition to modelling wave-to-wave changes and changes across the entire study period for the constructs, a cross-lagged model is included, where each assessment wave t is used to predict changes in the other construct(s) between assessment t and assessment wave $t+1$. This feature allows the ability to test the direct influence of personality on experience (or *vice versa*) but goes beyond standard cross-lagged models in making use of a latent change parameter, thus avoiding the confounding of regression error with actual change inherent in cross-lagged models (Rogosa, 1980). As a result, level of experience at each assessment wave can be used to explain subsequent changes in personality. Furthermore, because this model is an elaboration of a dual-process LGCM, the hypotheses outlined earlier for LGCMs can be tested, with some constraints imposed to make the models equivalent. Because of this flexibility, these types of models provide great promise for testing the various patterns and processes of personality development. Thus far, these models have been successfully employed within the area of cognitive development but have yet to catch on among more traditional personality researchers (Ferrer & McArdle, 2010). That is likely due to the many waves of assessment required to address more dynamic developmental processes.

Growth mixture models

Another extension of traditional LGCMs is to identify different classes of developmental trajectories (Ram & Grimm, 2009; Jung & Wickrama, 2008). As noted, in the standard LGCM, one trajectory form is assumed to be applicable to everyone in the sample, and the same is true of the relations between changes and any predictor variables. For naturally distinguishable populations, such as men and women, the assumptions of uniform trajectory (and other parameters) can be relaxed using multiple group models where each parameter is freely estimated in each group or constrained equal. However, often theory suggests that different classes or groups of people may have varying trajectories; for example, investment in life roles (e.g. occupational) may result in different developmental consequences (e.g. Roberts et al., 2005). Sometimes, these different groups or classes may be best identified based on their trajectory 'signatures' if groups or classes cannot be directly measured or identified. Growth mixture modelling combines latent class analysis with LGCM to identify such latent groups that evidence different trajectories.

Recent examples of this method in personality include Klimstra, Hale, Raaijmakers, Branje, and Meeus's (2009) identification of developmental differences in Block's three adolescent types. Extending this across adolescence to young adulthood, Johnson, Hicks, McGue, and Iacono (2007) found that three separate groups characterized personality development during this time period. Such groups likely do not exist in a strict sense such that there are qualitative or taxonomic differences among members of different groups. But the ultimately arbitrary division into groups can help to highlight underlying varying patterns of development, and these can be related to various indicators of functioning and experience, such as family problems and alcohol use, to characteristics that may explain the developmental patterns found within each group. Additional studies need to evaluate to what degree various associations noted were specific to the particular allocation of participants to groups that 'made sense' in any one sample. For example, it is likely that the experiences that drive personality do not influence everyone in the same manner or to the same degree; some experiences 'stick,' and affect personality, while other people are able to shrug these experiences off. Whether or not an experience 'sticks' may define a group, as judged from its distinctive postexperience trajectory.

UNDEREXAMINED QUESTIONS IN PERSONALITY DEVELOPMENT

Studies have utilized LGCM, and to a lesser extent growth mixture models, to describe patterns of personality development though only a few have gone beyond that to address the processes that drive change. Later, we discuss various questions that are currently underexamined in the study of personality development and discuss how the models described earlier may be used to address such questions.

Measurement invariance and latent repeated measures

In the basic LGCM represented in Figure 1, the repeated measures are comprised of scale scores, a composite measure or even individual items. Instead of using manifest variables as indicators, it is possible to treat the repeated measures as

DOI: 10.1002/per

latent constructs (Figure 3). This type of model is commonly referred to as a second-order LGCM, meaning that in addition to the latent factors of intercept and change, there are latent repeated measures. Currently, second-order LGCMs are underutilized in the field of personality development. While examples of models exist (e.g. Allemand, Steiger, & Fend, in press; Jackson, Hill, Payne, Roberts, & Stine-Morrow, 2012; Lüdtke, Roberts, Trautwein, & Nagy, 2011; Wagner, Lüdtke, Jonkmann, & Trautwein, 2013), the vast majority of LGCMs do not employ latent repeated measures. One of the main advantages of working in an SEM framework is the ability to have a theoretically error-free measure of the construct by modelling it as a latent variable. Doing so should result in greater precision in the estimates of associations between explanatory variables and change parameters, in addition to smaller standard errors for parameter estimates (Ferrer, Balluerka, & Widaman, 2008). These advantages also extend to latent change models, which also can incorporate second-order factors (Ferrer et al., 2008).

Another advantage of these second-order models is that they are set up to test measurement invariance. When dealing with questions of development and change, regardless of the statistical model one employs, it is necessary that the construct of interest be measured in the same way across measurement occasions. Frequently, it is implicitly assumed that a scale score at one occasion can be interpreted as reflecting the construct identically at another measurement occasion. For example, liking to go to parties is assumed to mark extraversion for 25 year olds in the same way as it does for 40 year olds. This may not be the case, however. For example, 40 year olds are more likely to have children at home and may much prefer to socialize in ways that involve them and other parents rather than attending parties. There may be many ages when the items tend to be interpreted and mark the construct much differently than they do at others, thus altering the meaning of the responses (Zimprich, Allemand, & Lachman, 2012). If this is the case, then mean level differences between young adults and middle-aged adults cannot be simply interpreted as differences in extraversion because the item does not assess the latent construct in the same manner. The term *measurement invariance* is used to signify when the construct is measured similarly at all time points and is an essential prerequisite for strict comparison of mean levels over time and/or across age groups.

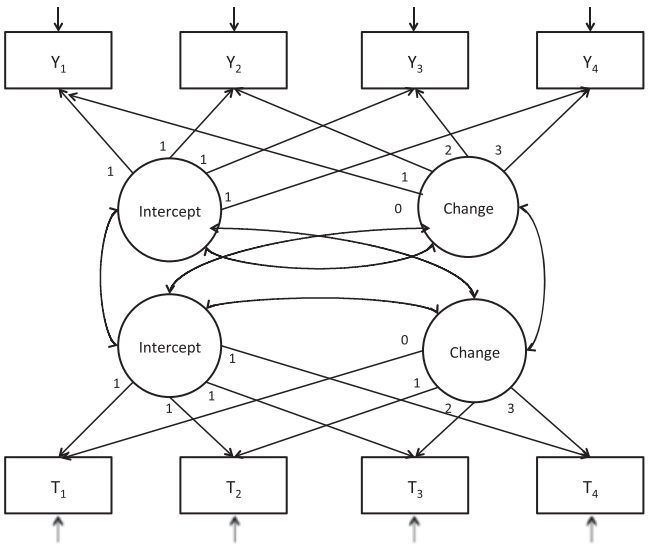

Figure 3. Bivariate latent curve model.

Measurement invariance over time can be tested through standard confirmatory factor analysis by examining whether the factor loadings, intercepts and residual variances are the same across measurement occasions. To test measurement invariance, one constructs a series of nested models—constraining the loadings, intercepts and variances to be equal in succession across successive models—starting with the least restrictive model (no constraints imposed) and ending with the most restrictive model (total invariance) in which all parameters are constrained to be the same across time. If the introduction of increasingly stringent invariance constraints results in little or no change in goodness of fit, there is evidence in support of measurement invariance.

Measurement invariance over time may be tested separately from the LGCM or may be tested and then imposed within the model (Widaman, Ferrer, & Conger, 2010; see Allemand et al., in press for a treatment of longitudinal measurement invariance with categorical variables). A deciding factor is whether or not the researcher desires to model the repeated measures latently in a second-order model or as manifest variables. If measurement invariance is established prior, one could fit a simplified model where the repeated measures are not assessed latently. However, fitting a second-order LGCM model allows one to include measurement invariance constraints, plus offers the advantages that go along with latent measures (Ferrer et al., 2008). Additionally, if the latent repeated measures came from a combination of sources such as self, peer and parent, a latent method factor could be added to the model to account for possible differences among the raters.

A relatively unexplored area for personality development researchers is what to do when measurement invariance does not hold. In a completely unsystematic survey with various personality development researchers, lack of measurement invariance is not uncommon. Typically, measurement invariance is viewed as a hurdle to clear and, once cleared, quickly forgotten. Failing to find invariance is viewed as a negative attribute, where changes in the construct of interest cannot be easily interpreted. Studies are discarded or findings are brushed under the table. However, as Edwards and Wirth (2009) pointed out, many theories predict that the manifestation of a construct will change over time—and that these changes are developmentally interesting! For example, items likely change in meaning across adolescent development when numerous physiological, sociological and psychological changes occur. Glossing over these differences or deeming the items that reflect them as poor and invalid is disregarding substantive information.

In situations such as these where invariance conditions are not met, we urge researchers to test whether certain a certain subset of the items demonstrate invariance. One can then use these items to identify the factor, though it needs

DOI: 10.1002/per

to be remembered that dropping items will change the interpretation of the factor when compared to the original full set of items. Then examine whether the ill-fitting item can inform theory, given that the misfit may stem from developmental changes. Furthermore, items that are not invariant across time can be included in the overall model to reflect the changing relationship across time through incorporating time specific item-weightings (Edwards & Wirth, 2009). Doing so allows one to see when and how these items change their association with the factor. An additional consideration is whether misfit across models reflects meaningful invariance rather than just statistically significant invariance. Traditional tests of measurement invariance may be overly sensitive and indicate non-equivalence when the differences are relatively minor, though they also often suffer from lack of power that hides important violations of invariance. We suggest the use of effect sizes estimate of the invariance to inform whether the invariance is meaningful (see Nye & Drasgow, 2011 for additional details). Furthermore, it is important to note where the invariance is occurring, as variances in the factor loadings are likely more problematic for meaningful interpretation than variance in the residuals.

For tests of measurement invariance and for fitting second-order latent growth models, the latent repeated measures do not necessarily have to use items for their indicators. Often parcels—which are composites of items—are used instead of items to create the latent variable. Parcels tend to be more reliable and more normally distributed than single items and are thus better able to meet the assumptions of maximum likelihood estimation, though there is some disagreement about the appropriateness of using parcels—for example, it can be difficult to establish that item allocations to parcels are appropriate and that results robust across different allocations (Little, Rhemtulla, Gibson, & Schoemann, 2013; Marsh et al., 2013). Parcels reduce the number of estimated parameters and, therefore, reduce the complexity of the measurement model, resulting in better model fit. Parcels have been successfully used in a number of studies in personality to test mean level changes in personality traits (e.g. Allemand, Zimprich, & Hertzog, 2007; Jackson et al., 2009) and self-esteem (Wagner et al., 2013), though further studies need to test the applicability of these models during times such as adolescence and old age where measurement invariance might not hold.

Can interventions be used to change personality?

Recently there have been many discussions concerning possible interventions to change personality traits (Heckman, Pinto, & Savelyev, 2013), with some studies already demonstrating that interventions or training may lead to such change (e.g. Blonigen, Timko, & Moos, 2013; Jackson, Thoemmes, Jonkmann, Lüdtke, & Trautwien, 2012; Tang et al., 2009). LGCMs are well suited to deal with these group designs and can handle this situation two different ways. First, treatment conditions can be used as a predictor variable, much like any other discrete life experience. A second way to test the effectiveness of interventions or experimental designs is to perform a multiple group analysis. Instead of predicting the change factor by group membership, one could instead fit a multiple-group LGCM that explicitly tests whether particular parameters differ across the control and experimental groups. To do so, two unconditional LGCMs are estimated, in which parameters can be constrained to be the same or allowed to vary freely across groups. If the fits of less constrained models are significantly better than that of the constrained model, then the two groups differ significantly on the affected parameters of interest (e.g. change factor).

It is helpful to examine differences between groups in all parameters of interest. When not using a multiple-group design (and instead using a dichotomous predictor variable), an assumption is that the other model parameters are the same across groups. Often, however, experimental designs can influence parameters other than change, such as the variance of a particular group or the response to a particular question that may be related to the treatment, and the multiple-group design can reveal pretreatment group differences that could impact apparent treatment effects. Not accounting for these group differences in model parameters can lead to biased results, which is why a multiple group design is quite useful. In addition to experimental designs, this multiple-group method could be especially useful in situations where preexisting groups exist. For example, romantic couple status (single versus in a relationship) could be associated with differences in both mean trajectories, as well as variance around those trajectories.

Can changes in personality be used as an explanatory variable?

While events may influence personality development, individual differences in change in personality may also influence outcomes—after all, attainment of some outcome goal is one of the main reasons people give for changing behaviour in the first place to predict outcomes. For example, increasing evidence suggests that changes in personality traits predict health and longevity above and beyond their respective initial levels (e.g. Magee, Heaven, & Miller, 2013; Mroczek & Spiro, 2007; Steiger, Allemand, Robins, & Fend, 2014). Changes in personality traits are thought to influence health because changes in personality likely reflect changes in personal behaviours associated with health such as diet and exercise (Takahashi, Edmonds, Jackson, & Roberts, 2013). Similarly, Hill, Jackson, Roberts, Lapsley, and Brandenberger, 2011 examined how changes in life goals during college can have downstream effects on one's satisfaction measured at age 35. Those who increased importance on goals for family and for their careers were happier at midlife than those for whom importance of these goals remained steady or decreased. These findings suggested that the process of changing one's goals influenced pathways to happiness separate from initial levels. Our understanding of personality development and its consequences would benefit from considering change not only as an outcome but also as an important process that can have downstream influences.

DOI: 10.1002/per

Person–environment transactions

Many theories of personality development propose that person–environment transactions play prominent roles in shaping development. For example, one theory posits that most personality change results from prolonged behavioural changes associated with investments in life roles. As new behaviours become habitual, they can lead to changes in the way people view themselves and the ways others view them, as reflected in personality trait ratings (Roberts & Jackson, 2008). Thus, most personality change should occur from continued experiences that unfold over long time intervals (Wrzus & Roberts, 2014), rather than single isolated events. Adequately testing such theories of person–environment transactions necessitates repeated assessments of both personality and experience. For example, a recent study looked at the student role and found that changes in academic investment leading up to a graduation exam were associated with increases in conscientiousness (Bleidorn, 2012). Similarly, another recent study found that increases in investment in one's occupational role were associated with increases in conscientiousness, whereas those that performed a greater number of counterproductive work behaviours across time were more likely to decrease in conscientiousness (Hudson, Roberts, & Lodi-Smith, 2012).

Studies such as these that identify correlated changes between personality and experience are thought to provide evidence of reciprocal effects between experience and personality. However, correlated changes do not necessarily reflect reciprocal effects; instead, changes may be correlated because an outside third variable causes change in both constructs. Or, change in one construct may lead to changes in the other but not *vice versa*. Or, correlated change may reflect an initial association between the variables that persists through time (Klimstra, Bleidorn, Asendorpf, van Aken, & Denissen, 2013). Unfortunately, LGCMs are not ideally structured to tease apart these possibilities.

In contrast to LGCM, latent change score models may be better suited to examine person–environment transactions given their stronger ability to test reciprocal relationships. One question concerning person–environment transactions is the extent to which such corresponsive effects are involved in personality development. Corresponsive effects are said to exist when social selection processes lead individuals into experiences that act to accentuate the same trait that led them to have the experience in the first place (Caspi, Roberts, & Shiner, 2005). Corresponsive effects relate two common processes within personality development: social selection and socialization processes. Selection effects are found when personality predicts a future experience and socialization effects are identified when that experience is associated with changes in personality. For example, neurotic individuals are more likely to experience negative life events. and in turn, these negative life events may cause further increases in neuroticism (Lüdtke et al., 2011). Under the LGCM, such corresponsiveness is suggested when an experience variable is associated with both latent personality intercept and change factors. The idea of corresponsive relationship however requires a reciprocal pattern of development, where the personality change in turn increases the likelihood of the experience recurring. With a LGCM, this may be addressed using correlated changes; however, these analyses do not provide any insight about whether selection was involved in the first occurrence of the experience if it is measured at the same time as the first assessment of personality.

With a bivariate latent change score model, however, these sequencing processes may be explored in more detail, by focusing on coupling parameters. Coupling parameters are the cross-lagged predictions from the assessment wave t in variable X to change scores in variable Y at assessment wave $t+1$ and *vice versa*, and they can differ from the analogous ones applying to waves $t+1$ and $t+2$. The coupling parameters can begin to distinguish whether levels of experience predict changes in personality and if personality also predicts changes in experience, much like in a typical cross-lag model. However, the difference here is that the latent change model takes into account absolute levels of change, which the cross-lag model does not. As a result, these models are better suited to test relative strengths of influence between variables that have correlated changes. While these models offer much promise for the field of personality development, they alone cannot tease apart causal directions, as many of the problems with cross-lagged models also apply to latent change score models (see above). Incorporation of these models along with an experimental designs or group differences may further be able to investigate direction of effects.

Another option to examine person–environment transactions is to identify possible indirect pathways that relate the two change processes. For example, by examining more narrowly defined behaviours, one could better test how the two broader variables are associated. Various experience sampling methods (Bolger & Laurenceau, 2013) could be used to measure daily behaviours associated with personality and the experiences subjects encounter. Through this more detailed measurement, one could test whether changes in personality assessed at the broad level are associated with changes in theoretically relevant behaviours or experiences at the lower order level. These changes in behaviour or experience at the lower order level can then be associated with changes in the other variable of interest that was assessed more broadly, across months or years (Roberts, 2009). These variables then can be examined simultaneously using basic meditational methods (Selig & Preacher, 2009). For example, to examine the relationship between changes in personality and changes in work satisfaction, one could measure the work behaviours that would likely change if conscientiousness changes, such as markers of job performance. If changes in conscientiousness are related to changes in the job behaviours, one can then test to see if changes in behaviour are also associated with changes work satisfaction. While this design does not directly examine the direction of association, a stronger inference can be made that theoretically meaningful behaviours and experiences, rather than a lurking third variable, are related to changes.

Depending on when the intervening variables are assessed and whether they are developmental constructs, an

DOI: 10.1002/per

additional approach is to model three separate growth processes and use the three change factors for the IV, DV and mediator. For example, changes in conscientiousness were associated with changes in health partially because people who increased in conscientiousness also increased in the tendency to perform healthy behaviours (Takahashi et al., 2013). Another recent application of this model used changes in drinking motives to explain that the relationship between changes in personality and 'maturing out' of problem drinking during young adulthood (Littlefield, Sher, & Wood, 2010). Ideally, the three change trajectories would occur during different time periods—as in standard mediation models where the IV precedes the mediator, which precedes the DV—so as to rule out reverse causation.

Are changes correlated across traits and methods of assessment?

In addition to associations between changes in personality and experience, researchers might be interested in whether changes in one personality construct correspond with changes in another, as this may indicate common or unique drivers of development. To date, there is mixed evidence that traits change relative in concert. For example, in one of the first studies, Allemand et al. (2007) observed a number of medium effect-sized correlated changes in the Big Five personality traits across a 4-year period in middle-aged and older adults. Moreover, in follow-up study, and Allemand, Zimprich, and Martin (2008) found correlated personality change over 12 years, except for neuroticism. These findings imply that personality changes share a substantial amount of commonality. In contrast, Soto and John (2012) found only minimal correlations among changes in personality traits over 40 years. Such coordinated changes may be more common at various stages in the lifespan than others, such as during young adulthood or older age, or perhaps the inconsistent results were due to use of different analytic models, the measures under study or sample demographics. Moreover, when interpreting correlated changes, it is also important to compare the size of the initial association between the two variables with the correlated change estimate. Correlated changes could reflect initial relationship between variables due to neurophysiology, for example, persisting through time rather than novel mechanisms at work (Klimstra et al., 2013).

Examinations of correlations between two different modes of assessment offer important tests of whether changes in personality manifest in observable ways. For example, do personality change estimates based on parent-reports correspond to change estimates based on self-reports? Sizeable correlations would indicate that personality change manifests not just in self-perceptions but also in observable behaviours. Few studies of this kind exist. A recent study examined the association between changes in personality within romantic partners and found that self reported personality trait change modestly corresponded with partner reports of personality (Jackson, Fraley, Vicary, & Brumbach, 2014). The association between the two methods suggests that changes in behaviour can be observed by close others and meaningfully integrated into their overall perceptions of other people, though the different perspectives are not synonymous, much like the research on self-reports and other-reports of personality (Vazire, 2010). Future research needs to extend these types of correlated change questions to look at different age groups, length of study, measures and measurement modality to see whether and to what degree change estimates are related across these factors.

SUMMARY

In this article, we provided a non-technical introduction to longitudinal SEMs and described open questions in the field of personality development. Given the nature of this brief review, we only touched on a number of designs, neglecting important variants such as dyadic and biometric latent curve models that take into account dependencies among multiple LGCMs (e.g. Bleidorn et al., 2010; Kashy et al., 2008) or latent-state trait models (e.g. Donnellan, Kenny, Trzesniewski, Lucas, & Conger, 2012).

Though the models we reviewed cannot answer every important question concerning personality development, especially those concerning the causal processes involved in development, many questions for which they are well suited remain unaddressed. One major limiting factor to successfully employing many of the models discussed earlier is a thorough assessment of the experiences that ultimately shape one's personality. Currently, there is a paucity of longitudinal data that assess the proximal mechanisms of development; instead, assessments favour broader experiences (e.g. divorce) that are easier to assess and interpret. Rather than examining the impact of a broad event like divorce, it would be more advantageous to look at the experiences of each couple member and their interpretation of experiences prior to and after divorce. Many such fine-grain, daily experiences are currently being assessed in cross-sectional studies but have yet to make the leap to longitudinal studies, though there are exceptions in the works.

Furthermore, experiences should be assessed at multiple times so as to begin to better understand the extent to which the subjective and objective environment one surrounds themselves with is stable or changing. Most studies test whether static experiences are associated with changes in personality; however, it is just as possible that the environment can act to create continuity across time, resulting in fewer changes in one's personality (Sameroff, 1995). Unfortunately, too few studies exist to know how stable external environments are, as well as how this impacts personality development. With the correct data, both LGCMs and latent change models can begin to isolate experiences (either lower order experiences or long-term environmental presses) associated with personality change and/or stability.

While the SEMs outlined earlier can begin to address many important questions of personality development—such as what experiences are associated with change and/or stability—their ultimate downfall is the inability to detect causal direction with certainty. This inability stems from the difficulty of teasing apart selection and socialization processes in passive longitudinal designs. Given that almost all life

DOI: 10.1002/per

experiences are not completely random, selection processes play a role in shaping the experience. Because of this, socialization may actually be selection in disguise, as it is unclear whether changes thought to be due to a particular experience would occur if selection processes were not present. In the personality development equivalent of asking whether the chicken or the egg came first, we can ask: did the person change as a result of the experience or did the change occur because of the characteristics that lead the person to the experience in the first place? The conundrum occurs because one cannot have the experience without the person, much in the same way that having an egg depends on having a chicken, and *vice versa*. Without a way to tease the characteristics of the person (i.e. selection) apart from the experience (socialization), it cannot be certain that other, unmeasured experience related to the characteristics of the person were actually responsible for change.

Of course, researchers cannot ethically randomly assign many major life events, such as divorce, that would be necessary to tease apart selection and socialization processes. Even if one could, there are many experiences that would be near impossible to reliably manipulate, such as perceptions of a divorce. As a result, we are at a fork in the proverbial road to understanding the processes that drive personality development. One could choose to invest effort into designs involving interventions, training programmes, natural disasters or policy changes to obtain stronger internal validity. Or, researchers could employ many of our designs discussed earlier that are inadequate in terms of identifying a causal association but can be used as a first step to rule out alternative hypotheses. We feel that both are necessary and, when combined with improved measurement, both capable of moving personality development research forward.

ACKNOWLEDGEMENTS

The authors wish to thank Jaap Denissen, Wendy Johnson, Ulrich Orth, Stefan Schmukle, Jule Specht and two anonymous reviewers for their many invaluable suggestions.

REFERENCES

Allemand, M., Steiger, A. E., & Fend, H. A. (in press). Empathy development in adolescence predicts social outcomes in adulthood. *Journal of Personality*. DOI:10.1111/jopy.12098

Allemand, M., Zimprich, D., & Hertzog, C. (2007). Cross-sectional age differences and longitudinal age changes of personality in middle adult- hood and old age. *Journal of Personality*, *75*, 323–358. DOI:10.1111/j.1467-6494.2006.00441.x

Allemand, M., Zimprich, D., & Martin, M. (2008). Long-term correlated change in personality traits in old age. *Psychology and Aging*, *23*, 545–557. DOI:10.1037/a0013239

Bleidorn, W. (2012). Hitting the road to adulthood short-term personality development during a major life transition. *Personality and Social Psychology Bulletin*, *38*, 1594–1608.

Bleidorn, W., Kandler, C., Hülsheger, U. R., Riemann, R., Angleitner, A., & Spinath, F. M. (2010). Nature and nurture of the interplay between personality traits and major life goals. *Journal of Personality and Social Psychology*, *99*, 366.

Blonigen, D. M., Timko, C., & Moos, R. H. (2013). Alcoholics anonymous and reduced impulsivity: a novel mechanism of change. *Substance Abuse*, *34*, 4–12.

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: an introduction to diary and experience sampling research*. New York, New York: Guilford.

Bollen, K. A. (1998). *Structural equation models*. New York, New York: John Wiley & Sons, Ltd.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: a structural equation perspective* (Vol. *467*). Hoboken, NJ: Wiley.

Caspi, A., Roberts, B. W., & Shiner, R. (2005). Personality development. *Annual Review of Psychology*, *56*, 453–484.

Donnellan, M. B., Kenny, D. A., Trzesniewski, K. H., Lucas, R. E., & Conger, R. D. (2012). Using trait–state models to evaluate the longitudinal consistency of global self-esteem from adolescence to adulthood. *Journal of Research in Personality*, *46*, 634–645.

Edwards, M. C., & Wirth, R. J. (2009). Measurement and the study of change. *Research in Human Development*, *6*(2–3), 74–96.

Ferrer, E., Balluerka, N., & Widaman, K. F. (2008). Factorial invariance and the specification of second-order latent growth models. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *4*, 22.

Ferrer, E., & McArdle, J. J. (2010). Longitudinal modeling of developmental changes in psychological research. *Current Directions in Psychological Science*, *19*(3), 149–154.

Grimm, K. J., & Ram, N. (2009). Nonlinear growth models in Mplus and SAS. *Structural Equation Modeling*, *16*, 676–701.

Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development*, *58*, 80–92.

Heckman, J. J., Pinto, R., & Savelyev, P. (2013). Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, *103*(6), 2052–2086.

Hertzog, C., & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: an overview of methodological issues. *Psychology and Aging*, *18*, 639.

Hill, P. L., Jackson, J. J., Roberts, B. W., Lapsley, D. K., & Brandenberger, J. W. (2011). Change you can believe in changes in goal setting during emerging and young adulthood predict later adult well-being. *Social Psychological and Personality Science*, *2*, 123–131.

Hudson, N. W., Roberts, B. W., & Lodi-Smith, J. (2012). Personality trait development and social investment at work. *Journal of Research in Personality*, *46*, 334–344.

Jackson, J. J., Bogg, T., Walton, K., Wood, D., Harms, P. D., Lodi-Smith, J. L., & Roberts, B. W. (2009). Not all conscientiousness scales change alike: a multi-method, multi-sample study of age differences in the facets of conscientiousness. *Journal of Personality and Social Psychology*, *96*, 446–459.

Jackson, J. J., Fraley, R. C., Vicary, M., & Brumbach, C. (2014). *The honeymoon effect in couples: personality trait development from two perspectives*. Manuscript under review

Jackson, J. J., Hill, P. L., Payne, B. R., Roberts, B. W., & Stine-Morrow, E. A. (2012). Can an old dog learn (and want to experience) new tricks? Cognitive training increased openness to experience in older adults. *Psychology and Aging*, *27*, 286.

Jackson, J. J., Thoemmes, F., Jonkmann, K., Lüdtke, O., & Trautwien, U. (2012). Military training and personality trait development: does the military make the man or does the man make the military? *Psychological Science*, *23*, 270–277.

Johnson, W., Hicks, B. M., McGue, M., & Iacono, W. G. (2007). Most of the girls are alright, but some aren't: personality trajectory groups from ages 14 to 24 and some associations with

DOI: 10.1002/per

outcomes. *Journal of Personality and Social Psychology*, *93*, 266–284.
Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, *2*, 302–317.
Kashy, D. A., Donnellan, M. B., Burt, S. A., & McGue, M. (2008). Growth curve models for indistinguishable dyads using multilevel modeling and structural equation modeling: The case of adolescent twins' conflict with their mothers. *Developmental Psychology*, *44*(2), 316.
Klimstra, T. A., Bleidorn, W., Asendorpf, J. B., van Aken, M. A. G., & Denissen, J. J. A. (2013). Correlated change of Big Five personality traits across the lifespan: a search for determinants. *Journal of Research in Personality*, *7*(6), 768–777.
Klimstra, T. A., Hale III, W. W., Raaijmakers, Q. A., Branje, S. J., & Meeus, W. H. (2009). Maturation of personality in adolescence. *Journal of Personality and Social Psychology*, *96*, 898.
Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods*, *18*, 285–300.
Littlefield, A. K., Sher, K. J., & Wood, P. K. (2010). Do changes in drinking motives mediate the relation between personality change and "maturing out" of problem drinking?. *Journal of Abnormal Psychology*, *119*, 93.
Lüdtke, O., Roberts, B. W., Trautwein, U., & Nagy, G. (2011). A random walk down university avenue: life paths, life events, and personality trait change at the transition to university life. *Journal of Personality and Social Psychology*, *101*, 620–637.
Magee, C. A., Heaven, P. C. L., & Miller, L. M. (2013), Personality change predicts self-reported mental and physical health. *Journal of Personality*, *81*: 324–334. DOI: 10.1111/j.1467-6494.2012.00802.x
Luhmann, M., Orth, U., Specht, J., Kandler, C., & Lucas, R. E. (2014). Studying changes in life circumstances and personality: it's about time. *European Journal of Personality*, *28*(3), 256–266.
Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J., & Von Davier, M. (2013). Why item parcels are (almost) never appropriate: two wrongs do not make a right—camouflaging misspecification with item parcels in CFA models. *Psychological Methods*, *18*(3), 257.
McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, *60*, 577–605.
McArdle, J. J., & Bell, R. Q. (2000). An introduction to latent growth models for developmental data analysis. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: practical issues, applied approaches, and specific examples* (pp. 69–107). Mahwah, NJ: Erlbaum.
McArdle, J. J., & Hamagami, F. (2001). Linear dynamic analyses of incomplete longitudinal data. In L. Collins, & A. Sayer (Eds.), *Methods for the analysis of change* (pp. 139–175). Washington, DC: American Psychological Association.
Mroczek, D. K., & Spiro, A. (2007). Personality change influences mortality in older men. *Psychological Science*, *18*, 371–376.
Nesselroade, J. R., & Baltes, P. B. (1974). Adolescent personality development and historical change: 1970-1972. *Monographs of the Society for Research in Child Development*, *39*, 1–80.
Neyer, F. J., & Lehnart, J. (2007). Relationships matter in personality development. Evidence from an 8-year longitudinal study across young adulthood. *Journal of Personality*, *75*, 535–568.
Nye, C. D., & Drasgow, F. (2011). Effect size indices for analyses of measurement equivalence: Understanding the practical importance of differences between groups. *Journal of Applied Psychology*, *96*, 966–980.
Ram, N., & Grimm, K. (2007). Using simple and complex growth models to articulate developmental change: matching theory to method. *International Journal of Behavioral Development*, *31*, 303–316.
Ram, N., & Grimm, K. J. (2009). Methods and measures: Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development*, *33*(6), 565–576.
Raudenbush, S. W., & Bryke, A. S. (2002). *Hierarchical linear models: applications and data analysis methods* (Vol. *1*). Thousand Oaks, CA.: Sage.
Roberts, B. W. (2009). Back to the future: personality and assessment and personality development. *Journal of Research in Personality*, *43*, 137–145.
Roberts, B. W., & Jackson, J. J. (2008). Sociogenomic personality psychology. *Journal of personality*, *76*(6), 1523–1544.
Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science*, *17*, 31–35.
Roberts, B. W., Wood, D., & Smith, J. L. (2005). Evaluating five factor theory and social investment perspectives on personality trait development. *Journal of Research in Personality*, *39*(1), 166–184.
Rogosa, D. R. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, *88*, 245–258.
Sameroff, A. J. (1995). General systems theories and developmental psychopathology. In D. Cicchetti, & D. J. Cohen (Eds.), *Developmental psychopathology* (pp. 659–695). New York: Wiley.
Scollon, C. N., & Diener, E. (2006). Love, work, and changes in extraversion and neu- roticism over time. *Journal of Personality and Social Psychology*, *91*, 1152–1165.
Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development*, *6*, 144–164.
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: modeling change and event occurrence*. New York: Oxford university press.
Small, B. J., Hertzog, C., Hultsch, D. F., & Dixon, R. A. (2003). Stability and change in adult personality over 6 years: Findings from the Victoria Longitudinal Study. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *58*(3), 166–176.
Soto, C. J., & John, O. P. (2012). Development of Big-Five domains and facets in adulthood: mean-level age trends and broadly versus narrowly acting mechanisms. *Journal of Personality*, *80*, 881–914
Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: the impact of age and major life events on mean--level and rank--order stability of the Big Five. *Journal of Personality and Social Psychology*, *101*, 862.
Steiger, A. E., Allemand, M., Robins, R. W., & Fend, H. A. (2014). Low and decreasing self-esteem during adolescence predict adult depression two decades later. *Journal of Personality and Social Psychology*, *106*, 325–338.
Tang, T. Z., DeRubeis, R. J., Hollon, S. D., Amsterdam, J., Shelton, R., & Schalet, B. (2009). Personality change during depression treatment: a placebo-controlled trial. *Archives of General Psychiatry*, *66*, 1322.
Takahashi, Y., Edmonds, G., Jackson, J. J. & Roberts, B. W. (2013). Longitudinal correlated changes in conscientiousness, preventative health-related behaviors, and physical health. *Journal of Personality*, *81*, 417-427
Van Aken, M. A., Denissen, J. J., Branje, S. J., Dubas, J. S., & Goossens, L. (2006). Midlife concerns and short-term personality

DOI: 10.1002/per

change in middle adulthood. *European Journal of Personality*, *20*(6), 497–513.
Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of personality and social psychology*, *98*(2), 281.
Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: measuring the same construct across time. *Child Development Perspectives*, *4*, 10–18.
Wagner, J., Lüdtke, O., Jonkmann, K., & Trautwein, U. (2013). Cherish yourself: longitudinal patterns and conditions of self-esteem change in the transition to young adulthood. *Journal of Personality and Social Psychology*, *104*, 148.
West, S. G., Ryu, E., Kwok, O. M., & Cham, H. (2011). Multilevel modeling: current and future applications in personality research. *Journal of Personality*, *79*, 2–50.
Wrzus, C., & Roberts, B. W. (2014). *TESeRa: short- and long-term processes of personality development in adulthood*. Unpublished manuscript.
Zimprich, D., Allemand, M., & Lachman, M. E. (2012). Factorial structure and age-related psychometrics of the MIDUS personality adjective items across the life span. *Psychological Assessment*, *24*, 173

DOI: 10.1002/per