volume 15 issue 3
september | 2013

www.ehps.net/ehp

The European Health Psychologist

Bulletin of the European Health Psychology Society original article

Longitudinal Methods in the Health Sciences: Four Recommendations

Gertraud Stadler
Columbia University
Megan L. Robbins
University of California
Jean-Philippe Laurenceau
University of Delaware
Niall Bolger
Columbia University

Longitudinal research with multiple time points has become more popular in health psychology, fueled by the rise of eHealth/mHealth studies. This article will address common challenges using longitudinal designs in the health sciences, including sources of variance and reliability of change, the difference between within-person effects and between-person effects, within-person mediation, and power. We make four recommendations: (1) to select change-sensitive measures and calculate variance components and reliability of change routinely as a starting point of data analysis, (2) to distinguish within-person process from between-person effects in data analysis, (3) to consider within-person mediation processes, and (4) to think of the different sources influencing power in longitudinal designs and to conduct power analyses. We will discuss how the use of advanced longitudinal designs could shape theory and research in the health sciences.

The Value of Longitudinal Designs in the Health Sciences

In the past decades, health scientists have begun more and more to study health and its correlates and determinants as they fluctuate and change over time. Intentions, self-efficacy, mood, behavior, health - all can fluctuate from day to day, week to week, while growing up from child to adolescent to adult, when acquiring healthy habits and shedding unhealthy ones, becoming sick and getting healthy again.

Longitudinal designs have a number of strengths. In longitudinal studies, researchers can minimize retrospective bias with appropriate assessment instruments, focus on within-person change versus between-person differences, get a better close-up picture of processes as they unfold, and examine how varying contexts influence affect, behavior, and health. However, longitudinal studies also present unique challenges. Therefore, we would like to present four recommendations for longitudinal research in the health sciences.

Integrating Theoretical Model, Temporal Design, and Statistical Model of Change

When health scientists study change over time it is helpful to consider how to best achieve “integration of theoretical model, temporal design, and statistical model” (Collins, 2006, p. 509). For coming up with a theoretical model of change, researchers need to know quite a bit about the phenomena of interest. How much evidence is there already about the speed of the process you want to study? How quickly do outcome and predictors fluctuate - across minutes, hours, days, weeks, years? What is the meaningful part of that variation in relation to random noise? What are the most important predictors of an outcome over time? In many cases, there is not much prior longitudinal evidence to answer these questions, particularly for a specific population of interest. In this case, pilot studies can help to make more informed guesses.

When researchers have come up with an - ideally evidence-based - theoretical model of change, they should match the temporal design and data analysis of their study with the hypothesized change in predictors and outcomes as closely as possible. Based on the theory of change, they will decide at what time to begin and end the study, how often to assess, and at what intervals. In a world of limited resources, design decisions often require tough compromises. If there are critical periods where most of the change occurs then most assessments should occur in that critical period and assessments before and after can be more spaced out. For example, researchers would measure more frequently right after the diagnosis of chronic illness, and more rarely later when patients have adapted and developed stable routines. But even with an ideal temporal design researchers still need to select measures and statistical models that fit their theory of change (see Recommendation 1), allow to distinguish within- and between-person variation (see Recommendation 2), and get at the processes of interest, including mediation (see Recommendation 3), all with adequate power (see Recommendation 4). We will visit each of these issues with four recommendations.

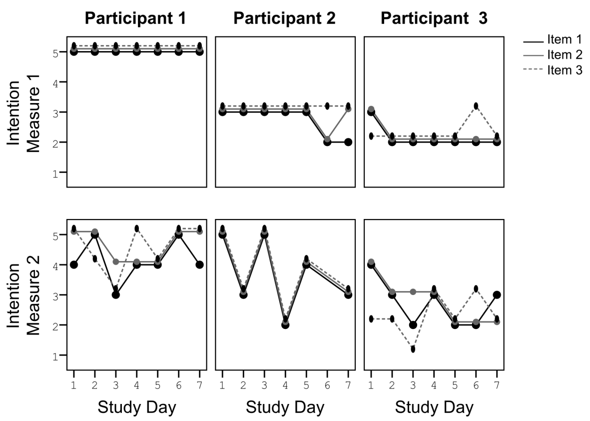


Figure 1: Panel plot of three participants' intentions per study day for 2 intention measures with 3 items each. Intention Measure 1 shows between-person variability, but little within-person variability, while Intention Measure 2 shows both between- and within-person variability.

Four Recommendations

Recommendation 1: Select Change-Sensitive Measures and Calculate Reliability of Change

Because most measures have been optimized for cross-sectional research rather than longitudinal research, finding appropriate measures with good psychometric properties for longitudinal studies can be quite challenging. A good starting point for integrating theoretical model, temporal design, and data exploration is to understand sources of variance and the reliability of each construct of interest in a longitudinal study. Ideally, for building a theory of change, researchers would be able to look up variance components and reliability of change in prior longitudinal studies and have conducted a pilot study in the population of interest.

To illustrate Recommendation 1, we will follow a research team interested in investigating changes in intention and physical activity in patients diagnosed with a chronic illness in a longitudinal intervention study. The researchers may have found two brief intention measures with three items each used in previous studies and tried them out in a brief pilot study in their population of interest. Looking at the data from their pilot study, the research team could start data exploration by drawing panel plots of individual participants' intentions (measured with three items each and the two intention measures) across study days. As Figure 1 shows for three exemplary participants, the

$$\mathsf{M_{pti}} = \mu + \mathsf{P_p} + \mathsf{T_t} + \mathsf{I_i} + (\mathsf{PT})_{\mathsf{pt}} + (\mathsf{PI})_{\mathsf{pi}} + (\mathsf{TI})_{\mathsf{ti}} + [(\mathsf{PTI})_{\mathsf{pti}} + \varepsilon_{\mathsf{pti}}] \qquad \text{(Eq. 1)}$$

two intention measures give different information. Intention Measure 1 (Figure 1, upper panels) captures differences in intention level between participants (between-person variability) while Intention Measure 2 (Figure 1, lower panels) captures intention fluctuations within person (within-person variability) in addition to differences in intention level between persons (between-person variability).

Shrout and colleagues have suggested using a generalizability theory framework (Cronbach, Gleser, Nanda, & Rajaratnam, 1972) to analyzing reliability in longitudinal data (Cranford, Shrout, Iida, Rafaeli, Yip, & Bolger, 2006; Shrout & Lane, 2012). Following this approach, researchers divide the available total variance for a certain measure into variance components linked to person, time, and item, and their combinations (Step 1) and then use these variance components to calculate reliabilities (Step 2).

In Step 1, the total variance is divided into variance components for person, time, and item, based on a three-way, crossed, analysis of variance model (person by time by item). The response of person p at time t to a certain item i, M_{pti}, can be understood as a combination of the nine components shown in Equation 1. The first component, μ, represents the population mean of the measure. The second component, P_p, captures that each person p can have higher or lower responses, regardless of items and time points; this effect reflects between-person differences in how persons respond to the measure of interest. The third component, T_t, captures that responses can be higher or lower at time point t compared to other time points, for all persons and all items. The fourth component, I_i, captures that item i can receive higher or lower responses than other items, for all persons and time points. The fifth component, $(PT)_{pt}$, captures that person p can give higher or lower responses at time point t, on all items. This component is particularly interesting for longitudinal research because it

Table 1: Sources of Variance and Reliabilities for Intention Measure 1 and Intention Measure 2 with three items each measured across 7 days.

Source of Variance		Intention 1 (3 items, 7 days)	%	Intention 2 (3 items, 7 days)	%
Variability across persons	σ^2_{PERSON}	0.45	33%	0.38	23%
Variability across days	σ^2_{TIME}	0.02	1%	0.01	1%
Variability across items	σ^2_{ITEM}	0.08	6%	0.07	4%
Person-by-time variability	$\sigma^2_{PERSON*TIME}$	0.03	2%	0.44	27%
Person-by-item variability	$\sigma^2_{PERSON*ITEM}$	0.17	12%	0.22	13%
Time-by-item variability	$\sigma^2_{TIME*ITEM}$	0.01	1%	0.01	1%
Residual variability	σ^2_{ERROR}	0.6	44%	0.49	30%
Total	σ^2_{TOTAL}	1.37	100%	1.62	100%
Reliabilities					
Between-person reliability	R_{KF}	0.95		0.95	
Within-person reliability	R_C	0.13		0.73	

$$R_{KF} = \frac{\sigma^2_{\text{person}} + \frac{\sigma^2_{\text{person*item}}}{i}}{\sigma^2_{\text{person}} + \frac{\sigma^2_{\text{person*item}}}{i} + \frac{\sigma^2_{\text{error}}}{t*i}} \quad \text{(Eq. 2)}$$

$$R_C = \frac{\sigma^2_{\text{person*time}}}{\sigma^2_{\text{person*time}} + \frac{\sigma^2_{\text{error}}}{i}} \quad \text{(Eq. 3)}$$

indicates systematic change over time: Some persons respond higher or lower at a certain time, regardless of the items used for the response. The sixth component, $(PI)_{pi}$, captures that person p can give higher or lower responses to item i than other items, at all time points. The seventh component, $(TI)_{ti}$, captures that item i can get higher or lower responses at time point t by all persons. The eight and ninth components, $(PTI)_{pti}$ and e_{pti}, capture that some persons have higher or lower responses on some items at specific time points. We would need repeated assessments of each item at a specific time point, to distinguish the systematic effect, $(TI)_{pti}$, from random error, e_{pti}. For most designs where each person provides only one response to each item at each time point, we cannot distinguish this error term from the three-way interaction effect of item, person, and time point, and therefore estimate them together with only one term, as indicated by the brackets around the two terms.

Following the generalizability theory approach for our example, the research team would conduct an analysis of the variance components for the two intention measures. Shrout and Lane (2012) provide code for conducting these analyses in SPSS and SAS. Figure 2 and Table 1 show sources of variance for two intention measures with three items each. Four variance components are of particular interest (see Table 1 and Figure 2 for the example data set): variability across persons, person-by-time variability, person-by-item variability, and residual variability. In the example, variability between persons accounted for about a third of the variance in Intention Measure 1 (33%) while it accounted for a quarter of the variance in Intention Measure 2 (23%). Person-by-time variability accounted for hardly any variance in Intention Measure 1 (2%) while it accounted for another quarter of the variance for Intention Measure 2 (27%). Because systematic change over time is often the main reason for conducting longitudinal research, the research team for our example should be excited to see that Intention Measure 2 seems to capture a good amount of this variance. Person-by-item variability was comparable between the two intention measures (Intention Measure 1: 12%, Intention Measure 2: 13%). Residual variability was larger for Intention Measure 1 (44%) than for Intention Measure 2 (30%)

In Step 2, we then use these variance components to calculate between-person reliability and reliability of change. Assuming fixed time points and items as in the example study, Cranford and colleagues (2006) calculate between-person reliability as shown in Equation

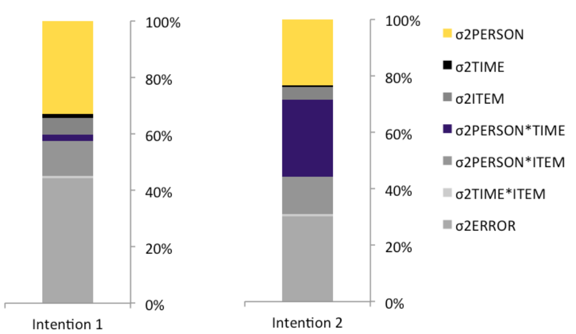


Figure 2: Sources of variance for Intention Measure 1 and Intention Measure 2 with three items each measured across 7 days.

$$\text{Predictor}_p t = \text{Predictor\_W}_{pt} + \text{Predictor\_B}_p \qquad \text{(Eq. 4)}$$

2. Between-person reliability is a ratio, with the numerator being the sum of variability across persons and person-by-item variability, divided by the number of available items, and the denominator this same sum plus residual variability, divided by the product of number of time points t by number of items i. Table 1 shows that both intention measures in our example show excellent between-person reliability.

Cranford and colleagues (2006) calculate reliability of change as shown in Equation 3. Reliability of change is a ratio, with the numerator being person-by-time variability, and the denominator being the sum of person-by-time variability plus residual variability, divided by the number of items i. Table 1 shows that Intention Measure 1 in our example shows unacceptably low reliability of change (0.13), while Intention Measure 2 shows acceptable reliability of change (0.73) and would therefore be the measure of choice for further studies. An example write-up for Intention Measure 2 in a methods section would be: The measure showed outstanding between-person reliability (R_{KF} > .90) and acceptable reliability of change (R_C > .70).

For more details and syntax for calculating variance components and reliability for longitudinal designs, see Shrout and Lane (2012). Shrout and Lane (2012) give different examples for calculating appropriate reliabilities, depending on the design of the study. The generalizability theory approach presented by Cranford and colleagues (2006) assumes that items and assessment times can be distinguished and are thus fixed within person. That makes sense for items because we usually have a specific set of items and are not selecting randomly from a pool of items. For assessment times around a critical event, such as diagnosis of a chronic illness or an online intervention where all participants start at the same time, all participants have the same assessment time points and the assumption of fixed time points makes sense as well. However, in other designs, for example, studies with event-contingent assessment or experience sampling studies with random beeps, it makes sense to assume that assessment times are random and thus nested within person. For another approach within a multilevel framework, see also Wilhelm & Schoebi (2007).

Recommendation 2: Distinguish Within-Person Change From Between-Person Effects

Our second recommendation is to distinguish within-person processes from between-person effects. Longitudinal data, compared to cross-sectional data, provide the opportunity to observe and analyze changes over time within a person, facilitating the study of health and behavior in daily life (Mehl & Conner, 2012). Time-varying constructs, such as intention and behavior, contain two sources of variation, (a) within-person fluctuations around (b) each person's mean level that varies between persons. Figure 3 illustrates this distinction for the intention data used as an example. Person 1 has on average high intentions, Person 2 has moderate intentions, and Person 3 has low intentions. But all persons show at times higher and lower intentions than their typical level on the change-sensitive Intention Measure 2. Multilevel models for analyzing longitudinal data differentiate within- and between-person variability for outcomes, but not by default for predictors. Therefore, within- and between-person effects have been confounded in many

$$Y_{it} = \gamma_{00} + \gamma_{01}\text{Time}_{pt} + \gamma_{02}\text{Predictor\_W}_{pt} + \gamma_{03}\text{Predictor\_B}_{p} + \varepsilon_{pt} \qquad \text{(Eq. 5)}$$

prior analyses of longitudinal data but they need to be carefully distinguished to avoid biased conclusions (e.g., Allison, 2009; Bolger & Laurenceau, 2013; Curran & Bauer, 2011; Hamaker, 2012; Raudenbush & Bryk, 2002). The distinction of between- and within-person variability in the predictor variables by using person-level means and within-person deviation scores is one important contribution that many manuscripts have neglected to make so far.

In the theoretical model of change, it is helpful to distinguish between-person effects from within-person processes. Between-person effects reflect stable associations between predictor and outcome, and are prone to all alternative explanations that we are familiar with from cross-sectional research. For example, persons with higher intentions may be more physically active, but the causal mechanism behind this association could be in any stable third variable that is related to both intention and activity. Should the research team find that on days when participants show higher intentions they also show higher physical activity the number of alternative explanations shrinks to constructs that covary with intentions and activity from day to day. A last important theoretical question regarding within-person processes is if increases in a predictor have the same effects as decreases. Most theoretical models assume causal symmetry by default but increases and decreases could have differential effects. For example, increases in intention could have different effects on activity than decreases in intention. Stadler and colleagues have shown an approach to separate effects of increases and decreases in a predictor and found differential effects (Stadler, Snyder, Horn, Shrout, & Bolger, 2012). For the temporal design, it is important to select time frame and assessment frequency carefully and measure all variables as time-varying constructs with change sensitive measures, to allow for within-person fluctuations in the predictor as well as in the outcome, as delineated in the specific theory of change.

It is relatively straightforward to implement the distinction of within-person process and between-person effects in the statistical model. Each raw predictor score for person p at time t, $Predictor_{pt}$, can be split up into the time-varying within-person deviation, $Predictor\_W_{pt}$, from each individual i's average predictor level across all available time points, $Predictor\_B_p$ (see Equation 4).

Thus, we first calculate person means in the predictor variable across all available time points, $Predictor\_B_p$. Then we subtract each person's mean from the raw predictor score, $Predictor_{pt}$, to arrive at deviation scores for the predictor for each time point, $Predictor\_W_{pt}$. To facilitate interpretation of the intercept, we calculate the mean of the person means and center the person means $Predictor\_B_p$ at the

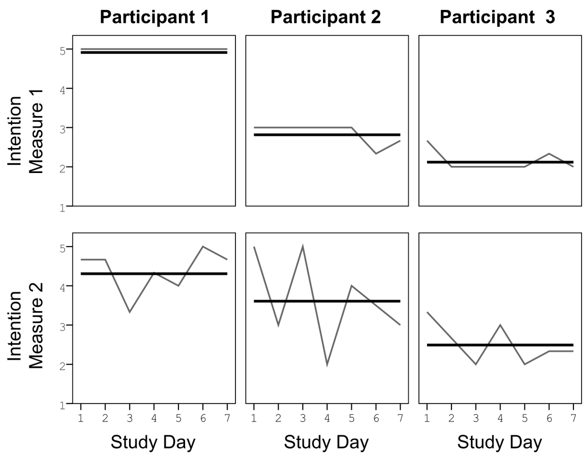


Figure 3: Three participants' intention fluctuations over 7 days around each person's mean for Intention Measure 1 and Intention Measure 2.

grand mean by subtracting the sample mean from the person means. Finally, we enter predictor deviations and person averages, the latter centered at the grand mean, into the multilevel model (Equation 5).

Equation 5 illustrates the data analysis approach for a continuous outcome Y_{it} predicted by time and the within- and between-person predictor. The coefficient γ_{02} tests whether at times when a participant is higher on the predictor than usual he or she is higher or lower on the outcome (within-person association); the coefficient γ_{03} tests whether persons who are higher in average predictor levels are also higher in the outcome (between-person association).

Interestingly, within- and between-person effects can differ considerably in size and even direction, and can differ in their causal processes. Neglecting these differences can obscure theory building (see Wilson, Stadler, Boone, & Bolger, under review). When researchers keep trying to find an effect on the between-person level that exists in the population on the within-person level and vice versa they will find mixed results (for more in-depth discussion, see Mehl & Conner, 2012; and Bolger & Laurenceau, 2013). Combined with a sound theory of change, the data analytic approach described above can facilitate health scientists' distinction between within- and between-persons effects, enhancing our understanding of temporal health processes.

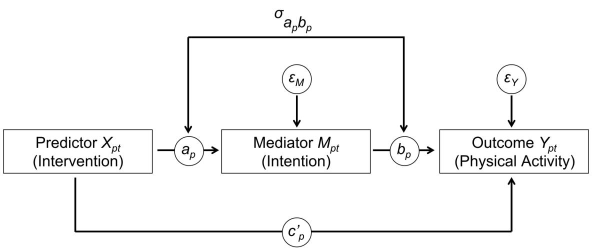


Figure 4: Example of within-person mediation: How much of the effect of the within-person predictor (randomly delivered daily intervention) on the outcome (daily physical activity) is mediated by the within-person mediator (daily fluctuation in intention around the person mean)?

Recommendation 3: Consider Within-Person Mediation Processes

Our next recommendation - to consider within-person mediation during theory building, design, and analysis – relies on the two prior recommendations to choose change-sensitive reliable measures and distinguish within-person processes from between-person effects. Given the great interest in developing and testing theories in health psychology, longitudinal researchers who want to understand causal influences on the within-person level can do so by using within-person mediation. This type of mediation is especially suited for intensive longitudinal data (Bolger & Laurenceau, 2013). Because participants are assessed repeatedly in an intensive longitudinal study, each participant can have his/her own mediation effect. Based on each person's mediation effect, researchers can then estimate an average within-person mediated effect as well as between-person heterogeneity around that average. For example, a research team conducting an intervention study aimed at increasing intentions to be physically active would want to see if the intervention actually increased intentions and if the increase in intentions explained the intervention's effect on physical activity. They could pursue these questions with a classic between-person mediation analysis (Baron & Kenny, 1986). But if cause, mediator, and outcome were measured repeatedly over time, they could test whether mediation occurs within each person and to what degree the causal chain explains the intervention effect for different persons. For example, if the intervention was delivered randomly on certain days (and assuming no carry-over effects), the research team would want to know if on intervention days - compared to control days – intentions were higher in most participants and if, in turn, this led to higher physical activity. Figure 4 provides an example of within-person mediation for a randomly delivered daily intervention aimed at increasing intentions and physical activity. Note that the three mediation coefficients c'_p as well as a_p and b_p are estimated for each person (indicated by a person-specific subscript p) allowing estimation of each person's mediation effect in addition to the average mediation effect. Finally, within-person mediation includes a new term, σ_{ab}, indicating how much the predictor-mediator link (i.e., a_p) covaries with the mediator-outcome link (i.e., b_p) and which must be included in the calculation of mediated effects if there are substantial *a* and *b* random effects (see Kenny, Korchmaros, & Bolger, 2003).

Within-person mediation provides health scientists with another tool for exploring causal mechanism. For hands-on guidance on how to conduct a within-person mediation analysis, see Bolger and Laurenceau (2013) who provide a detailed introduction including syntax and example data to run these analyses.

Recommendation 4: Pay Attention to Factors Influencing Power

Our last recommendation is to pay attention to factors influencing power in longitudinal studies. Power indicates the probability to detect a hypothesized effect with a given sample, if the effect actually exists in the population. A common threshold for acceptable power is .80, indicating that a study will detect the population effect with a probability of 80%. Many researchers are familiar with the five determinants of power in the cross-sectional context: effect size, sample size, variability of the predictor, unexplained variance in the outcome, and Type I error probability. Studies have more power if they investigate large effects, with large samples, maximize the variability of the predictor variable, minimize unexplained variance in the outcome, and choose more lenient probability levels (although there is little leeway to stray from the accepted .05 standard level).

Power to detect within-person effects in longitudinal studies has three additional determinants of power: the number of repeated time points, the amount of autocorrelation between time points, and how much the effect varies from person to person - in addition to the size of the within-person effect, sample size, variability of the within-person predictor, unexplained variance in the outcome, and Type I error probability. Studies with more time points per person, lower autocorrelation between time points, and relatively similar effects of the predictor on the outcome across participants have higher power.

In addition to the traditional ways for increasing power discussed above, the three additional determinants can also be addressed through research design. First, one straightforward way to increase power is to add additional time points to the temporal design. However, adding persons is a better way to increase power (see Figure 5). Second, choosing a temporal design that spaces time points not

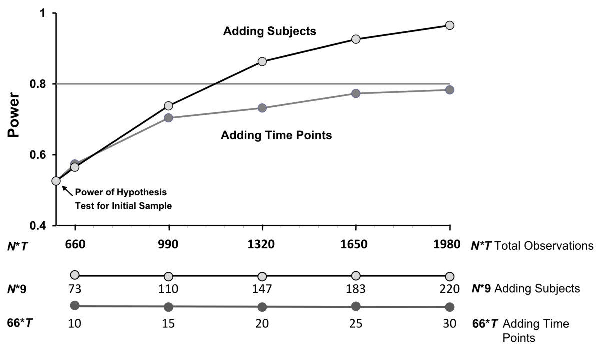


Figure 5: Power curves for the within-person fixed effect of physical activity on depression: What is the benefit of adding persons versus time points to the sample? (reprinted from Bolger, Stadler, & Laurenceau, 2012, p. 299)

too closely together diminishes autocorrelation between time points. Last, researchers can reduce the variability in the predictor-outcome link between persons by using interventions tailored to participants' needs and with standardized implementation or choosing time intervals where the predictor occurs relatively uniformly across participants.

Considering power as part of theory building, study design, and data analysis is a practice that pays off especially in the longitudinal context to optimize the allocation of resources. Bolger, Stadler, and Laurenceau (2012) give more details on conducting power analyses for within-person effects, and provide syntax and an example data set. For conducting power analyses across a wide range of research designs, see Bolger & Laurenceau (2013).

Discussion

The future of longitudinal research in the health sciences is very promising. With a growing evidence base, researchers can achieve better fit between theory, study design, and data analysis. So far, we have only a vague picture of how health and its determinants change over time for many populations, and we need to rely on pilot studies that are necessarily giving limited information to inform larger studies. Without knowing the dynamics of change, it is hard to know how to best allocate resources. We may need to rely on a more fine-grained temporal resolution, keeping in mind that we can always aggregate measures if change is slower than we thought while we cannot retrieve more details that we have not collected. With more evidence, future research will increasingly become more efficient and sophisticated, relying on measures geared towards capturing change and within-person effects, and even allowing us to look at mediating and moderating processes with enough power. Longitudinal research, and particularly intensive longitudinal studies and data-burst designs, give us a chance to gain a more complete understanding of stability and change in health across the life course and its causes.

References

Allison, P. D. (2009). *Fixed effects regression models*. Thousand Oaks, CA: Sage.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.

Bolger, N., & Laurenceau, J-P. (2013). *Intensive longitudinal methods*. New York: Guilford Press.

Bolger, N., Stadler, G., & Laurenceau, J.-P. (2012). Power analysis for intensive longitudinal studies. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 285–301). New York: Guilford Press.

Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, 57, 505–528.

Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin*, 32, 917–929.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, 62, 582-619.

Hamaker, E. L. (2012). Why researchers should think "within-person": A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 43–61). New York: Guilford Press.

Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower level mediation in multilevel models. *Psychological methods*, 8, 115-128.

Mehl, M. R., & Conner, T. S. (Eds.). (2012). Handbook of research methods for studying daily life. New York: Guilford Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

Shrout, P. E., & Lane, S. P. (2012). Psychometrics. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 302–320). New York: Guilford Press.

Stadler, G., Snyder, K. A., Horn, A. B., Shrout, P. E. & Bolger, N. (2012). Close relationships and health in daily life: A review and empirical data on intimacy and somatic symptoms. *Psychosomatic Medicine*, 74, 398-409.

Wilhelm, P., & Schoebi, D. (2007). Assessing mood in daily life: Structural validity, sensitivity to change, and reliability of a short-scale to measure three basic dimensions of mood. *European Journal of Psychological Assessment*, 23, 256-267.

Wilson, P. A., Stadler, G., Boone, M. & Bolger, N. (under review). Fluctuations in well-being and sexual risk episodes among HIV-positive gay men.

Gertraud Stadler
is Associate Research Scientist at the Department of Sociomedical Sciences and Department of Psychology, Columbia University, United States

stadler@psych.columbia.edu

Megan L. Robbins
is Assistant Professor at the Department of Psychology, University of California, Riverside, United States

mrobbins@email.arizona.edu

Jean-Philippe Laurenceau
is Professor of Psychology at Department of Psychology, University of Delaware, United States

jlaurenceau@psych.udel.edu

Niall Bolger
is Professor of Psychology at Department of Psychology, Columbia University, United States

bolger@psych.columbia.edu

letter

Transparent Reporting, the Foundation for Full Disclosure

A letter to Peters, Abraham, & Crutzen (2012) and Hagger, Conner, & O'Connor (2013)

Thomas E. Fuller
Mark Pearson
Jaime L. Peters
University of Exeter Medical School

The recent discussion between Peters, Abraham, & Crutzen (2012) and Hagger, Conner, & O'Connor (2013) are timely and welcome additions from a health psychology perspective to the broader issue of improving the reporting of research in a transparent and accurate manner. Addressing authors' perceived barriers (such as those described by Peters et al.) to share and fully disclose data sets, syntax and output is a complex and challenging task; it will require significant commitment and sustained effort from all parties involved. Crucially, a balance will also need to be attained between the needs of those publishing datasets and those wishing to examine them.

The disclosure and sharing of data is an important aspect of improving transparency in research but should be considered as a necessary complement to the full and accurate reporting of what was planned and done. Without this, a data set loses meaning as readers cannot assess whether or not it was obtained in a methodologically sound way. Furthermore, fully reporting a completed study satisfies the ethical obligation researchers have to research users, the scientific community and the public who fund research through taxation.

A large number of guidelines, designed to support the reporting of studies using a wide range of designs and/or specialist fields of research, currently exist. (See the EQUATOR network website for more information: http://equator-network.org.) In some instances the use of a reporting guideline is a requirement of the journal despite - with the exception of CONSORT - there being a lack of data on the effectiveness of this as a strategy to improve the reporting of health research. By failing to establish whether or not reporting guidelines (like any intervention) are effective, an opportunity is missed to potentially refine and enhance a strategy that could improve the transparency of reporting of health research.

To this end, our group is currently conducting an evaluation of the Transparent Reporting of Evaluations with Non-randomised Designs reporting guideline (TREND; Des Jarlais, Lyles, Crepaz, & TREND Group, 2004). Whilst imperfect, TREND's focus on behavioural and public health interventions and external validity has the potential to be relevant and useful to health psychologists' reporting of research. We have found some evidence to suggest more complete reporting and better study quality with TREND users. Further analyses are currently being conducted and we plan to submit these for publication before the end of 2013.

Reporting guidelines and policies requiring full disclosure are unlikely to be the only interventions to improve research reporting, but they may well form a solid foundation on which to build. It is likely that additional initiatives (e.g. All Trials: www.alltrials.net), strategies at a range of levels (e.g. author, editor, journal, publisher, funding agency, regulatory body), possibly involving a degree of enforcement, will be required to facilitate change in reporting behaviours and policies. Establishing the role played by each of these components will contribute to our understanding of effective strategies to improve the reporting of research in health psychology and related fields.

References

Des Jarlais, D. C., Lyles, C., Crepaz, N., & Group, T. (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health*, 94(3), 361-366. doi: 10.2105/ajph.94.3.361

Hagger, M., Conner, M., & O'Connor, D. (2013). Response to Peters, Abraham, & Crutzen (2012) from the EHPS journal editors. *European Health Psychologist*, 15(1), 6-7.

Peters, G.-J., Abraham, C., & Crutzen, R. (2012). Full disclosure: doing behavioural science necessitates sharing. *European Health Psychologist*, 14(4), 77-84.

Competing interests

The authors declare that they have no competing interests.

Funding

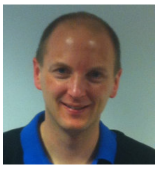
Funding for the project is provided by the Peninsula Collaboration for Leadership in Applied Health Research and Care (PenCLAHRC) from a competitive grant awarded by the National Institute for Health Research (NIHR).

Acknowledgement

This letter presents independent research funded by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) for the South West Peninsula. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health in England.

Thomas E. Fuller
is Associate Research Fellow at Peninsula Collaborations for Leadership in Applied Health Research and Care (PenCLAHRC), University of Exeter Medical School, United Kingdom

t.fuller@exeter.ac.uk



Mark Pearson
is Senior Research Fellow at Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, United Kingdom

mark.pearson@exeter.ac.uk

Jaime L. Peters
is Research Fellow at both PenCLAHRC and PenTAG, University of Exeter Medical School, United Kingdom

j.peters@exeter.ac.uk

commentary

Special Issue of *Health Psychology* Highlights Interface of Health and Social Psychology

Angela D. Bryan
University of Colorado Boulder

Some of the most robust theory development, theory testing, and empirical work in health psychology is done by or in collaboration with social and personality psychologists. The longstanding tradition of the merging of broadly applicable theories and methods of social and personality psychology with the domain of health has origins that date back more than half a century. As one example, in the early 1950's, the United States Public Health Service was keenly interested in why more people did not avail themselves of preventive screening and vaccination. Three social psychologists, Godfrey Hochbaum, Irwin Rosenstock, and Stephen Kegels, were tasked with determining how best to increase the use of these services. Hochbaum, Rosentock, and Kegels were trained in the Lewinian tradition, and thus it is not surprising that "the orientation of the work would be toward developing a theory not only useful in explaining a particular problem, but also adaptable to other problems" (Rosenstock, 1974, p. 329). Indeed, as the "father" of social psychology, Lewin himself conducted research in a health context, working during World War II to encourage the eating of non-traditional meats. Although these examples reflect early work in the United States, perhaps nowhere has the integration of social and health psychology been stronger than in Europe, where leading social psychologists bring their strong empirical and theoretical traditions to bear on questions of health significance.

Understanding how far the rich collaborative relationship between social and health psychology has come, where it stands now and the challenges and opportunities of the future was the underlying motivation for a recently published special issue of Health Psychology entitled "Theoretical Innovations in Social and Personality Psychology and Implications for Health" [Volume 32, Number 5, May 2013]. Guest editors William Klein, Alex Rothman, and Linda Cameron have structured a compendium that brings together three distinct types of articles, including excellent work by European social psychologists. First is a section of Conceptual Articles that highlight the state of the science in the broad fields of social/personality psychology and judgment and decision-making. The goal of this section was to elucidate current theories that are either already informing work in a health context or, perhaps more importantly, are ripe for extrapolation to the area of health. For example, Paschal Sheeran, along with colleagues Peter Gollwitzer and John Bargh highlight the relevance of nonconscious processes to health behavior. A second section features some outstanding empirical articles. In this section, Natalie Schüz, Benjamin Schüz and Michael Eid highlight the role of self-affirmation in mitigating defensive reactions to threatening health information in the context of skin cancer prevention. Finally, commentaries from leading scholars view the intersection of health and social/personality psychology through the lens of the future. Included is a commentary by Susan Michie, Robert West, and Bonnie Spring exploring the fertile ground and important challenges of realizing the potential of the theory to practice cycle in social, personality and health psychology.

In sum, this Health Psychology special issue is an excellent compendium of current empirical work and theorizing at the intersection of social, personality, and health psychology. It was partially sponsored by the National Cancer Institute, and a limited number of free printed copies of the journal are available (to request a copy, please contact Juanita Cox at coxj@efdb.nci.nih.gov). As readers of The European Health Psychologist are well aware, the major causes of morbidity and mortality in the western world are chronic conditions that are the direct result of human behavior; either lack of healthy behavior or excess in unhealthy behavior (Fisher et al., 2011). This Health Psychology special issue is thus timely and important. If we hope to have a meaningful impact on public health, the science of health behavior change—merging the theory and methods of social and personality psychology with the applied questions and contexts at the core of health psychology—will most certainly be at the forefront.

References

Rosenstock, I.M. (1974). Historical origins of the health belief model. *Health Education Monographs*, 2, 328-335.

Fisher, E.B., Fitzgibbon, M.L., Glasgow, R.E., Haire-Joshu, D., Hayman, L.L., Kaplan, R.M., Nanney, M.S., Ockene, J.K. (2011). Behavior matters. *American Journal of Preventive Medicine*, 40(5), e15-e30.

Angela D. Bryan
is Professor of Psychology and Neuroscience at Department of Psychology and Neuroscience, University of Colorado Boulder, United States

angela.bryan@colorado.edu

debate paper

Innovation or Provocation? A Swedish Scenario on Future Welfare Services

Sven Ingmar Andersson

Lund University

More empathy and more high tech in the health and elder health care system in Sweden. That is one conclusion drawn in a recent governmental report analyzing the long-term demand for welfare service by 2050. The report, translated into English (Socialdepartementet, 2013), argues that health and elder care services cannot be produced much longer the way they are done now as the price the patients and clients will pay will be too high in terms of suffering and antiquated working methods (all the time it takes to get an appointment, getting to the clinic, sitting in the waiting room, for example). According to the report, individual patients, clients, caregivers and businesses have already taken the lead to implement innovative approaches and technologies which will enable people to do self-testing and automatic screening at home or at labs located in shopping centers or at metro platforms. Having the individuals themselves managing and monitoring their preventive health care through computerized online decision making by means of online scanners, for example, to analyze the vitality of bodily organs and cells ensures, it is argued, that serious maladies will be detected at an early stage and that treatment will be more cost-effective. The computerized diagnoses available through expert systems in the home are seen as more reliable than those provided by the human brain and will even make the health care centers of today rather superfluous. For such prevention to be successful, the argument continues, it is essential to synchronize the contact between patients and the computer simulated care system.

To accomplish this, the report introduces a system of mentors. A mentor uses the expert systems to assess what the individual wishes and needs and to make an accurate diagnosis based on the information provided by the automated systems. The accurateness thus accomplished will, according to the report, be the case in 87% of the time. Otherwise, specialists will be called in to deal with rare or, especially in regard to mental matters, serious disease. Generally, a person should have up to three mentors, one until the age of 35, another until they retire at the age of 65-80, and finally one for the rest of life. It is important, it is stated, that both parties accept each other, otherwise the system will not work. The mean number of mentees for a mentor is estimated to 50 or so. The skills for many mentors are described as those traditionally referred to as psychologists, primary care physicians, and physiotherapists, for example. Even patients with a long-term disease can, when "hyperlinked", manage most of their own health care through access to networks and technical resources such as robotics and exoskeletal machines, for example, thereby increasing autonomy and reducing personal costs.

Comments: The health care sector in Sweden has undergone great changes in recent years. From a health psychology point of view, changes have been both positive and negative. It is now quite common that psychologists are employed in community health care centers. Cognitive behavioural therapy (CBT) has developed dramatically and is a familiar concept to the public. Its popularity has resulted in education in CBT, generally through short-term courses, being offered to various groups of non-psychologists within the public and private health care sector. Such short education of psychological therapists acting in the health care market has been met by criticism. The governmental report and the recent development must be seen as an urgent need to reflect on the status of health psychology in the future

Reference

Ministry of Health and Social Affairs, Sweden (2013). *Empathy and high tech. Interim results of the LEV project.* http://www.government.se/sb/d/17184/a/208550

Sven Ingmar Andersson
is Professor of Psychology at the Departmenty of Psychology at Lund University, Sweden

sven_ingmar.andersson@psy.lu.se

network grant report

Bringing Researchers Together: Report on EHPS Network Grant

Richard Cooke
Aston University
Peter Harris
University of Sussex
Benjamin Schüz
University of Tasmania
Urte Scholz
University of Zurich
Guido van Koningsbruggen
VU University Amsterdam

This article describes the process of applying for an EHPS network grant ('Exploring when and how self-affirmation works') and the progress made in building a research network across several European countries. There are three aims of the article, to (1) outline the application process and consider how to manage a network grant, (2) highlight the benefits of network grant funding, and (3) outline the activities achieved during the grant period. We begin by outlining the background to the network grant application.

Background to grant application

The start point for the grant application was a shared interest in self-affirmation among the researchers involved in the bid. Self-affirmation involves asking individuals to focus on a valued aspect of their self-concept (e.g., honesty) prior to receiving threatening information (e.g., a health warning). At the time of the grant application, research had shown that self-affirmation was effective at reducing defensive processing of health warnings (see Harris & Epton, 2009, 2010), but had less impact on health behaviour change (Harris & Napper, 2005, Harris et al., 2008). In particular, there was a lack of research on the processes underpinning and modifying self-affirmation effects, and we thought it would be interesting to conduct research to explore these processes.

EHPS conferences provided a useful meeting point for discussions on self-affirmation, with symposia at the 2009 conference in Pisa and the 2010 conference in Cluj bringing together researchers interested in this topic. Richard Cooke, Peter Harris and Benjamin Schüz participated in both symposia and after the Cluj conference, Richard and Benjamin discussed the possibility of applying for an EHPS network grant. Independently, Urte and Peter were also discussing applying for an EHPS network grant, so they joined forces and recruited Guido to round off the team.

When considering making an application for an EHPS network grant it is a good idea to discuss your proposals with colleagues at conferences and to consider which European researchers are working in your area of interest. All applications to the network scheme need to have researchers based in three different European countries. Our submission included five researchers based in four countries (RC & PH in the UK; BS then in Germany, US in Switzerland and GvK in the Netherlands). Setting up networks in this way is a good idea if you plan to subsequently apply for funding from EU funding schemes, which typically require applicants from multiple countries.

Network grant application

Our proposal was to conduct research into the processes underpinning and modifying self-affirmation effects; each individual had different ideas about contributions to the project. Benjamin wrote an initial outline for the application, and then everyone added their own proposals for research projects: We settled on the idea that four researchers would conduct research studies into related, but independent, processes, and that one researcher would oversee the projects. As part of the application you need to nominate a coordinator who will liaise with the EHPS, and oversee the delivery of the network grant. This is an important post not only with regard to communicating with the EHPS, but also because it is always good to have a clear distribution of responsibilities. Moreover, we found, as with other grant applications, that it is well worth setting early deadlines for applicants to submit work so that the different pieces can be edited into a coherent application.

Managing the network grant

After getting funding, you are faced with several practical issues in managing the grant. The main issue to address is the timing of the research meetings. We proposed that we would have four research meetings, based in different countries, including a meeting at the first EHPS conference taking place once the award has been made, which is requirement of the scheme. We received funding (€5000) in November 2010, so the EHPS conference in Crete in 2011 was to be our first conference meeting. To progress the grant we met before Crete, in June 2011 in Berlin. Our second meeting was held in Crete in September 2011, and our third meeting was held in Amsterdam in May 2012. We held our last meeting at the EHPS conference in Bordeaux in July 2013.

Activity during the funding period

The work programme for our project was to conduct four studies. We have completed these projects and presented the initial results of these in a symposium at the 2013 EHPS conference in Bordeaux. Our results suggest that self-affirmation decreases resistance regardless of the health message’s threat level and to genuine emotive warnings, and may do so by increasing anticipated regret. However, some backfire effects occur and further studies are clearly needed and are currently being discussed in our networking group to understand more about how and why self-affirmation works. The study conducted by Richard showed that self-affirmation promoted physical activity regardless of threat level. The study conducted by Benjamin found that self-affirmation increased intentions to reduce alcohol consumption, with the strongest effects on heavy drinkers, and that these increased intentions lead to a subsequent reduction in alcohol consumption. The research by Guido showed that self-affirmation increased anticipated regret and intentions, and that regret mediated the affirmation effects on intentions. In addition, the results suggest that anticipated regret and intentions are serial mediators linking self-affirmation and behavior. Finally the study conducted by Urte found that self-affirmation decreased healthy intentions, which may have been due to the overall low level of defensiveness observed in this particular sample (cf. earlier research showing negative effects of self-affirmation among non-threatened participants).

So, thus far we have created a network of researchers across Europe, delivered a symposium on the work conducted, and generated data for four independent peer review papers. We are currently discussing future joint research projects and other funding options.

Complications

We planned to complete our research meetings in 12 months as required by the scheme. However, it quickly became apparent that this timeframe was impossible to meet. We recommend being realistic about your proposal: in hindsight we were too ambitious about what we wanted to achieve in 12 months, and we recommend that you factor your proposals into your day-to-day activities when applying for a network grant.

Another issue that we did not appreciate at the time of our application is what happens to your plans when your team members all decide to move jobs during your project. By the time of our first meeting in June 2011, we knew Benjamin was off to Australia. In 2012 Urte was moving from Bern to Konstanz and is now moving from Konstanz to Zurich. Also in 2012 Peter moved to Sussex and Guido is now working in Amsterdam. So, as with all projects, bear in mind that you need to factor in the unexpected – add some room for manoeuvre in your timings.

Benefits of EHPS network grant

When you apply for a network grant you know that it will benefit your research career, but we believe it is difficult to fully comprehend these benefits until you receive the funding and organize your meetings. The scientific benefits alone are worth the time and effort needed to put together an application: through the network grant we have been given the opportunity to collaborate with colleagues in other countries, learning about their approach to research, their ways of working, and forging links for future research. You are being given funds to discuss science, and there is nothing better, especially given the increasing demands on academic time. Creating an international network also opens up funding schemes that would be inaccessible to researchers working in one country. So, by proposing a network you can access not only EU funding schemes (e.g., Horizon 20:20) and national schemes that focus on working across Europe. For example the ORA scheme funds applications for bilateral projects between France, Germany, the Netherlands and the UK. Beyond the scientific benefits, there are also many social benefits from travelling around Europe meeting colleagues and discussing shared research interests!

Richard Cooke
is Senior Lecturer at the Department of Psychology at Aston University, United Kingdom
r.cooke@aston.ac.uk

Peter Harris
is Professor of Psychology at the School of Psychology, Sussex University, United Kingdom
p.r.harris@sussex.ac.uk

Benjamin Schüz
is Senior Lecturer at School of Psychology, University of Tasmania, Australia
benjamin.schuez@utas.edu.au

Urte Scholz
is Professor of Psychology at the Department of Applied Psychology, University of Zurich, Switzerland
urte.scholz@psychologie.uzh.ch

Guido van Koningsbruggen
is Assistant Professor at the Department of Communication Science, VU University Amsterdam, Netherlands
g.m.van.koningsbruggen@vu.nl