Introduction to Innovative Approaches to Longitudinal Data Analyses

Karina Weichold
University of Jena, Germany
E-mail: karina.weichold@uni-jena.de

Studying human development and modeling complex developmental processes over time requires well-planned studies to collect longitudinal data. To plan such studies and to analyze the resulting wealth of longitudinal data, however, often seems to be a challenge for developmental researchers. When conducting longitudinal studies, investigators must consider issues such as identification of sample participants, participant retention, and appropriate timing of data collections. Analyzing the resultant data sets, by statistical or other means, might become difficult when researchers are confronted with questions such as: Which analytic strategy will yield the best answers to my research questions (e.g., focusing on inter-individual differences in intra-individual change patterns)? What is the most appropriate software tool for my project? How do I most effectively demonstrate causality within my longitudinal study? This special section aims at bringing together outstanding experts to discuss innovative approaches to longitudinal data analyses. In addition, examples of excellence in longitudinal studies within behavioral science will be introduced.

Four feature articles focus on innovative approaches to longitudinal data analyses. The first one by Nesselroade and Molenaar stresses the importance of investigating smaller samples over a long period of time (and several measurement occasions) rather than investigating very large samples over the short term. Von Eye in the second feature article discusses the meaning of change and the usefulness of statistical tools (e.g., ANOVA) to analyze change across time. The third paper by Bergman and Nurmi highlights the importance of a person-oriented framework to investigate human development over time, and, finally, Steyer et al. discuss the opportunities of intervention studies to investigate causal effects. These essays are accompanied by a series of Reports from the Labs exemplifying successful strategies for conducting longitudinal studies. These studies cover different research topics, age groups, time spans, and cultural and historical contexts. To begin with, Bynner reports on the opportunities and challenges of birth cohort studies. In addition, famous longitudinal studies are introduced by Poulton and Moffitt (Dunedin Multidisciplinary Health and Development Study, New Zealand), Tremblay (Montreal Longitudinal and Experimental Study, Canada), and Schaie and Willis (Seattle Longitudinal Study of Adult Cognitive Development, USA). In these reports, such methodological issues as sample selection, retention strategies, and analytic procedures to evaluate the data will be discussed. Finally, Sharma and Verma focus in their lab report on the usefulness of event sampling techniques to collect time-series of data across a short period.

In addition, this issue of the Bulletin presents a new section (Country Focus), wherein we aim at introducing in each issue developmental research foci of different nations around the globe. We are happy to start this section with a report from Guatemala by Ureta, Batz and Grazioso. This Bulletin also includes the Young Scholars' Corner by Joche Gayles, a report by Paul Oburu from the ISSBD workshop in Kenya, and a tribute to Professor Xiaojia Ge by Run Jin. The News Section includes the notes from our president Anne Petersen—her very last missive as president of ISSBD. We would like to thank Anne for her support of the editorial team, in particular for her encouraging support in its transition to becoming the ISSBD Bulletin. In addition, we are looking forward to a fruitful collaboration with the president-elect, Wolfgang Schneider.

We thank all authors of this Bulletin for their excellent contributions and investment of effort and time, and also thanks to Matthias Reitzle for expert consultation. By tackling the topic of developmental methodology in the special section of this ISSBD Bulletin, we aimed at meeting the needs of the members of the society, because a considerable number of respondents to the membership survey identified methodological aspects as a topic they wanted to see in a special section of the Bulletin. We are happy to respond to this input and hope that the readers of the Bulletin will find the papers interesting and stimulating for their own empirical research.

Finally, we are sad to announce that Bonnie Barber has left the editorial team. A special thanks to her for her extremely valuable input to the numerous issues of the ISSBD Bulletin (formerly ISSBD Newsletter) that we have worked on together since 2006.

When Persons Should Be Few and Occasions Should Be Many – Modeling a Different Kind of Longitudinal Data

John R. Nesselroade
University of Virginia, USA
E-mail: jrn8z@virginia.edu

and

Peter C. M. Molenaar
The Pennsylvania State University, USA
E-mail: pxm21@psu.edu

Longitudinal research featuring fewer participants measured on more occasions of measurement provides an important avenue to the study of processes. We discuss some of the design and analysis issues stemming from this orientation for the purpose of fostering awareness of innovations that we believe make the orientation and the tools the rightful concerns of developmentalists.

The collection and analysis of longitudinal data remain central to mounting an empirical attack on the essential riddles of developmental research. A key reason for the sustained importance of longitudinal data in developmental research is the access they uniquely offer to the identification of patterns of intraindividual variability and change. The study of intraindividual variability, in turn, supports two important foci: (1) identifying interindividual differences (and similarities) in patterns of intraindividual change (Baltes, Reese, & Nesselroade, 1977); and (2) emphasizing the individual as a primary unit of analysis–an available option that we believe has been taken too lightly in many instances. As we have argued elsewhere (Molenaar, 2004; Nesselroade & Molenaar, 2010), an emphasis on the individual ought to be one of the fundamental pillars of the empirical study of development. It is to this topic that we devote the remainder of our discussion.

A focus on intraindividual variability dictates the use of research designs and analytical tools that feature multiple repeated measurements. The occasions dimension of the data box (Cattell, 1966) must be liberally sampled if intraindividual variability is to be understood. Because of the need to measure participants often and, we believe, with multivariate measurement schemes, realistic constraints that always impinge on the conduct of research necessarily dictate the use of smaller sample sizes than is typical of, say, a one-occasion, cross-sectional study. Hence, the title of this article was chosen to call attention to the kind of research design that seeks deliberately to trade off participants for occasions of measurement. But, we believe that it is a trade-off that matters a lot! In what follows, we will explore why such a trade-off is desirable and what are some of its implications for developmental researchers.

The Individual as the Primary Unit of Analysis

Emphasizing the individual as the primary unit of analysis, and we are certainly not the first to do so (see e.g., Carlson, 1971), has helped to resurrect the old idiographic nomothetic debate in somewhat new, productive guises. For example, Molenaar (2004) used the ergodicity arguments of classical mechanics to build a rigorous case for focusing on the individual as the unit of analysis. His comments were especially pertinent to developmentalists because developmental phenomena are among the least likely behavioral phenomena to yield their secrets up to traditional individual differences approaches. In other words, the ways individuals differ from one another are not necessarily the ways people change as they develop. The converse has never been proven and, according to the ergodicity arguments presented by Molenaar (2004), the nature of developmental change (e.g., lack of stationarity) militates against the successful use of differences among individuals as a proxy for changes within individuals. The implication is that research designs featuring more occasions at the cost of fewer participants are highly pertinent to the study of development and deserve more consideration by developmental researchers.

In the arena of psychometrics a stronger focus on the individual as the chief unit of analysis has led to a novel, if somewhat controversial approach to the measurement of psychological constructs (Nesselroade, Gerstorf, Hardy, & Ram, 2007). Based initially on the modeling of individual-level, intensive measurement data, what Nesselroade et al. (2007) called "the idiographic filter" affords a way to put aside the idiosyncratic features of behavior that impinge in undesirable ways on one's data collection operations, especially the measurement of psychological constructs, in order to emphasize the underlying common features. Operationally, the approach amounts to abandoning the traditional imposition of invariance constraints on factor loadings (the links between observed and latent variables) and, instead, places them on the factor intercorrelations. The result of this change in the locus of invariant relations is that factors derived from the factor intercorrelations (second-order factors) do have invariant loadings from one individual to another but the pattern of loadings of either the first or second order factors on the observed variables is idiosyncratic to the individual. This arrangement allows one to argue that the same underlying construct is being measured but it is being measured in idiosyncratic ways appropriate to a given individual.

Thus, the gathering of intensive information at the individual level forms a rather natural alliance with certain general approaches to modeling observations. For example, multivariate time series models are a generally appropriate analysis form but we tend to favor those that also rest on the common factor analysis model because of its capacity for modeling with latent variables. P-technique factor analysis, which has been used for over a half-century, represents such an approach. More recently, elaborations of P-technique factor analysis–dynamic factor analysis–have begun to move into a more prominent role for the modeling of process (e.g., Nesselroade & Molenaar, 2003). We will briefly review and summarize the current status of this approach.

P-technique and Dynamic Factor Analysis

P-technique factor analysis involves applying the factor model to data representing one person's responses on many variables over many successive occasions of measurement. When one inter-correlates the measures over the occasions of measurement what is dramatized are the patterns

of intra-individual variation and covariation. Thus, the P-technique factor analytic model provides a way to characterize the structure of intra-individual variability for a given person. R. B. Cattell, the developer and prime exponent of P-technique factor analysis (see e.g., Cattell, Cattell, & Rhymer, 1947), promoted it as a way to identify individual traits of behavior. Bereiter (1963) identified P-technique as "the logical way to study the interdependence of measures" (p. 15). Despite its promise, the P-technique factor model has been the subject of substantial criticism, especially since the early 1960s. Still, the model has been used successfully in many different applications (see e.g., Luborsky and Mintz (1972) and Jones and Nesselroade (1990) for reviews). Recently, Molenaar and Nesselroade (2009) demonstrated that whatever its shortcomings, the traditional P-technique factor model can be relied on to provide an accurate estimate of the dimensionality of the latent factor series and for estimating the latent factor series.

Substantial improvements to P-technique factor analysis have come in the form of a class of models named dynamic factor analysis (DFA) models. The DFA models greatly extend the power of P-technique factor analysis by explicitly modeling the dynamics of change. For example, the modeling of auto- and cross-lagged relationships among latent variables, which would seem to capture much of what we regard as process, is made possible by DFA. Recent explorations involve mating the general dynamic factor modeling approach with the idiographic filter measurement approach in order to further strengthen the ability to model process at the individual level. Obviously, success with this type of modeling depends on having adequate numbers of occasions of measurement on the individual and this constrains research design possibilities, as was noted above.

Dynamical Systems Approaches

Also requiring large numbers of occasions of measurement are individually-oriented dynamical systems modeling approaches such as those being developed and implemented by Boker and his colleagues (e.g., Boker & Wenger, 2007). These models, which are cast in differential equations, are becoming more and more practical as workable estimation procedures make their way onto the scene. Combining living systems theorizing (e.g., Ford, 1987; Ford & Lerner, 1992; Overton, 2007) with powerful modeling tools promises to advance the rigorous study of developmental processes in important new directions over the coming decade. For the foreseeable future, success will rest heavily on collecting and analyzing data from intensively measured and, therefore, limited numbers of participants.

Generalizability and Prediction

In traditional interindividual differences designs, generalizability is often considered "accomplished" if one simply uses large, representative samples of participants. Thus, the matter of generalizability of findings inevitably arises when limits are placed on the sampling of individuals. Because research designs focused on gathering intraindividual variability information tend to involve smaller numbers of cases, they are subject to criticism for their lack of generalizability. We take a somewhat unorthodox stance with

regard to this matter. We believe that simply gathering information on large samples of participants and averaging over them as the way to develop general conclusions is not the best way to do developmental science. A likely outcome of basing averages on larger and larger samples is weaker and weaker relations among important variables as the averages accommodate more and more heterogeneity. Given what we see as the desirability of searching for similarities in patterns of intraindividual variability across individuals, the matter of generalizability becomes very much an incremental pursuit rather than a representative sampling matter.

As Molenaar (2004) indicated, there is much more to be concerned about in large aggregations than sheer numbers of participants. Heterogeneity of within-person structures in large numbers of participants can become hidden from the standard cross-sectional analyses. Indeed, under such conditions, plausible modeling outcomes can be obtained with no indication that the results are misleading. A thorough examination of intraindividual variability–analyses at the individual level–is necessary to detect this. This is why we favor aggregation based on information concerning similarities (and differences) rather than unquestioningly tossing as many cases into the sample as we can and then averaging. Morever, Molenaar (2004) has argued that in the future the technology for the kinds of intensive designs we are advocating will improve to the point that much larger numbers of cases can be intensively measured. As the numbers of cases on which such research can be conducted increases, the matter of generalizability becomes less problematic.

The prediction of future behavior is an oft-expressed goal of behavioral science. Traditionally, this has been approached by establishing correlations between stable attributes of individuals and any number of future outcome measures of interest. Large numbers make even weak predictive relations cost-effective, for example. In a further illustration of how emphasizing the individual leads us down a different path, the prediction focus shifts more toward predicting what an individual will do in the future based on his or her past behavior rather than on what one person will do relative to another. This, in turn, argues for more extensive information regarding a given individual's behavior; that is to say more occasions of measurement.

Conclusion

Many innovations have been made in the past few years in the modeling of longitudinal data. While, in general, this would seem to be a positive situation, we contend that there are some events the long-term implications of which bear much more heavily on the future of developmental research than others. Most important among these, we believe, are methods and techniques that elevate the individual to the status of primary unit of analysis. We have articulated many of the arguments for this orientation in a series of discussions (see e.g., Molenaar, 2004; Molenaar, Huizenga, & Nesselroade, 2003; Nesselroade, 2010; Nesselroade & Molenaar, 1999; Nesselroade & Molenaar, 2010) and will not enumerate them here. Suffice it to say that we believe that great riches await those prospectors who are willing to stake their claim and invest their efforts in what remains under-explored territory.

References

Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life–span developmental psychology: Introduction to research methods*. Monterrey, CA: Brooks/Cole.

Bereiter, C. (1963). Some persisting dilemmas in the measurement of change. In C. W. Harris (Ed.), *Problems in measuring change*. Madison, WI: University of Wisconsin Press.

Boker, S. M., & Wenger, M. J. (Eds.) (2007). *Data analytic techniques for dynamical systems in the social and behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Carlson, R. (1971). Where is the person in personality research? *Psychological Bulletin, 75*(3), 203–219.

Cattell, R. B. (1966). The data box: its ordering of total resources in terms of possible relational systems. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology* (1st ed., pp. 67–128). Chicago, IL: Rand McNally.

Cattell, R. B., Cattell, A. K. S., & Rhymer, R. M. (1947). P-technique demonstrated in determining psychophysical source traits in a normal individual. *Psychometrika, 12*, 267–288.

Ford, D. H. (1987). *Humans as self-constructing living systems*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Ford, D. H., & Lerner, R. M. (1992). *Developmental systems theory: An integrative approach*. Newbury Park, CA: Sage.

Jones, C. J., & Nesselroade, J. R. (1990). Multivariate, replicated, single-subject designs and P-technique factor analysis: A selective review of the literature. *Experimental Aging Research, 16*, 171–183.

Luborsky, L., & Mintz, J. (1972). The contribution of P-technique to personality, psychotherapy, and psychosomatic research. In R. M. Dreger (Ed.), *Multivariate personality research: Contributions to the understanding of personality in honor of Raymond B. Cattell* (pp. 387–410). Baton Rouge, LA: Claitor's Publishing Division.

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology – this time forever. *Measurement: Interdisciplinary Research and Perspectives, 2*, 201–218.

Molenaar, P. C. M., Huizenga, H. M., & Nesselroade, J. R. (2003). The relationship between the structure of interindividual and intraindividual variability: A theoretical and empirical vindication of developmental systems theory. In U. M. Staudinger & U. Lindenberger (Eds.), *Understanding human development: Dialogues with lifespan psychology* (pp. 339–360). Norwell, Massachusetts: Kluwer Academic Publishers.

Molenaar, P. C. M., & Nesselroade, J. R. (2009). The recoverability of P-technique factor analysis. *Multivariate Behavioral Research, 44*, 130–141.

Nesselroade, J. R. (2010, in press). On an emerging third discipline of scientific psychology. In K. M. Newell & P. C. M. Molenaar (Eds.), *Individual pathways of change and development*. Washington, D. C.: American Psychological Association.

Nesselroade, J. R., Gerstorf, D., Hardy, S. A., & Ram, N. (2007). Idiographic filters for psychological constructs. *Measurement: Interdisciplinary Research and Perspectives, 5*, 217–235.

Nesselroade, J. R., & Molenaar, P. (2003). Quantitative models for developmental processes. In J. Valsiner & K. Connolly (Eds.), *Handbook of developmental psychology* (pp. 622–639). London: SAGE Publications.

Nesselroade, J. R., & Molenaar, P. C. M. (1999). Pooling lagged covariance structures based on short, multivariate time–series for dynamic factor analysis. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 224–250). Newbury Park, CA: Sage Publications.

Nesselroade, J. R., & Molenaar, P. C. M. (2010, in press). Emphasizing intraindividual variability in the study of development over the lifespan. In *Cognition, biology, and methods across the lifespan*. Volume 1 of the handbook of life-span development. Hoboken, NJ: Wiley.

Overton, W. F. (2007). A coherent metatheory for dynamic systems: Relational organicism-contextualism. *Human Development, 50*, 154–159.

The Many Parameters That Can Change

Alexander von Eye
Michigan State University, USA
E-mail: voneye@msu.edu

In this essay, the questions asked are: (1) what can the term change possibly mean? and (2) how can change be analyzed, statistically? These certainly are simple questions. However, only a very small number of the many possibilities that will be hinted at here are realized in current empirical research. Examples of concepts of change are given, and it is illustrated how statistical methods can be used to test hypotheses that correspond to specific concepts. Researchers are encouraged to pose detailed and specific questions concerning change.

Definition of Change

Change can be defined as variation in one or more parameters with reference to a context of change. Consider the following examples:

I. Change in mean. ANOVA is one method for analyzing changes in mean; structural models can be used also. The number of options is very large. One can ask whether a mean changes (1) from one observation to the next, (2) with reference to a single comparison category, (3) with reference to several comparison categories, (4) with reference to the grand mean, (5) with reference to all following means, (6) with reference to all prior means, or (7) in the form of developmental functions. Many other questions can be answered using ANOVA. Examples follow below.

II. Relative change in mean. For any of the examples listed under I, one can ask whether, in comparison to a reference, this change is larger, smaller, or the same. For example, one can ask whether change in one group of respondents differs from change in another group, whether later change is slower than earlier change, or whether change after treatment is faster than before treatment.

III. Change in variation. ANOVA requires that variability remains unchanged (homoscedasticity). However, developmental processes or treatments often have the (desired) effect that homoscedasticity no longer exists. Consider a training program. The result of this program can be that, in comparison to a control group, the training group improves performance. If training group members reach the learning criterion, the variability in the training group will be smaller than the variability in the control group. This may result in a violation of homoscedasticity requirements of such methods as regression or ANOVA. However, this can also indicate success of the training program.

IV. Predictability of change. Change in any parameter can be predictable from other parameters. For example, learning curves may be gender-specific. Here, we predict change from a time-stable variable (Gender). Change can also be predicted from time-varying variables. For example, change in the speed of healing may vary when a different treatment regimen is adopted. Methods of analysis depend on type of question asked. If the predictor is time-constant, ANOVA, discriminant analysis, log-linear, or structural models may be considered. Hierarchical linear modeling and structural modeling can be considered when covariates are time-varying. Configural Frequency Analysis (CFA) is the method of choice if statements are created in an exploratory context, and apply to sub-populations only. CFA can also be used to predict trajectories or end points (von Eye, Mair, & Mun, 2010).

V. Change in group membership. Over the course of adolescence, individuals may move into the group of experimental drug users at one point in time, and, at a later point in time, move into the group of abstainers. Similarly, in the labor market, individuals may move in and out employment categories, and, in the context of violence, individuals may move into and out of violent relationships. Methods of analysis include, for example, latent class analysis, event history analysis, structural modeling, and CFA.

VI. Differential change in group membership. In addition to asking questions concerning changes in parameters of group membership (see Item V), one can ask whether these changes are specific to sub-populations. One can, for example, ask whether moving in and out of employment status is gender-specific, or whether moving in and out of violent partnerships is age-specific.

VII. Change in group structure. Often, taxometric methods are used to create groups. This approach is of interest in particular when groups are formed that are described by similarities in behavior instead of organismic variables. The number of criteria that can be used to create groups is very large. von Eye, Mun, and Indurkhya (2004) discussed 10 decisions that must be made before selecting a clustering method. For example, one can create clusters based on measures of distance versus correlation, using probabilistic versus deterministic criteria, and with the goal of classifying everybody or only those who can be classified (cf. Bergman, 1988).

There are many more parameters than can change over the course of development or in response to treatment. In the following section, we illustrate statements that can be made based on the same data, when change in different
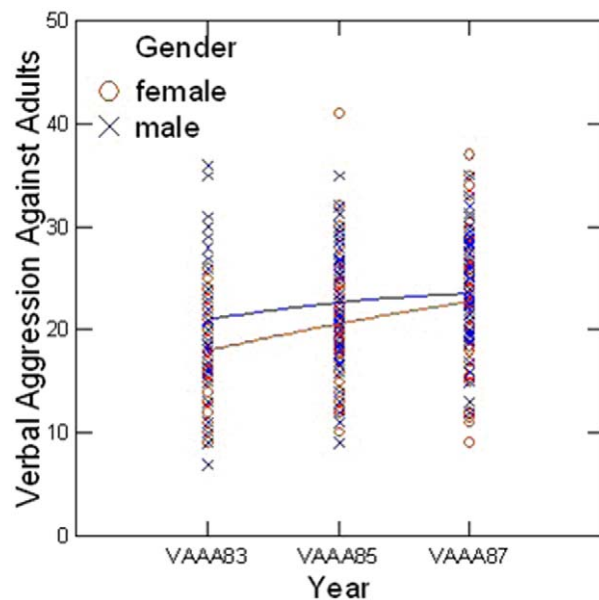

Figure 1. Gender-specific Development of Verbal Aggression during Adolescence

parameters is considered. Focus is on methods of ANOVA, structural modeling, and CFA.

Change in Parameters in the Same Data Set

In a study on the development of aggressive behavior in adolescence, Finkelstein, von Eye, and Preece (1994) asked how self-reported aggressive behavior changes over the course of adolescence. In addition, the authors asked whether this change is gender-specific. In this section, we present sample analyses of Finkelstein et al.'s (1994) data and illustrate that the same data can allow one to answer a plethora of questions concerning change. We select the variable Verbal Aggression Against Adults (VAAA). 114 adolescents (67 female) responded three times to questions concerning their VAAA, in the years 1983, 1985, and 1987. The three responses VAAA83, VAAA85, and VAAA87 resulted. The plot of the individual data points, along with the gender-specific trajectories over time (quadratic polynomials) appears in Figure 1.

The plot suggests that the gender-specific trajectories differ slightly and that there is a negatively accelerated trend of increasing verbal aggression against adults. In addition, boys seem to be verbally more aggressive than girls. We now illustrate a number of the above-mentioned possibilities of change in parameters.

Change in Mean (2.1). To analyze change in the mean parameter, we perform a repeated measures ANOVA with Gender as the between factor and the three observations of VAAA as the levels of the within factor. Effect coding was used, homoscedasticity was assumed, and equal spacing of observations over time was specified. Overall, boys were verbally significantly more aggressive against adults than girls ($F(1, 112) = 5.36$; $p = 0.022$), there was a significance increase in VAAA over time ($F(2,224) = 16.17$; $p < 0.01$), but the increase was not gender-specific ($F(2,224) = 1.81$; $p = 0.17$). The Huynh-Feldt å $= 1.00$ suggested only minimal

violations of compound symmetry. The Time effect explains 24% of the variance, and the Time x Gender interaction explains 3%.

As was indicated above, we can ask more specific questions than these, and still answer them using ANOVA. First, we ask whether the curve characteristics of the trajectory are as concluded from Figure 1. We, therefore, perform a polynomial decomposition and examine the first and second order polynomials, that is, the linear and the quadratic trends. We find a significant linear increase ($F(2, 112) = 21.64$; $p < 0.01$) and a non-significant curvature ($F(2, 112) = 0.21$; n.s.). The polynomials explain 28% of the variance.

Clearly, these are not the only parameters that can be examined. We can also ask whether time-adjacent and time-distant means differ from each other. To answer this question, we perform a profile analysis (Bock, 1975). Results suggest that all three mean differences are significant. Even the smallest difference, the one between 1985 and 1987 is, after Bonferroni correction for multiple tests, significant ($p = 0.03$). For the time-adjacent mean differences, we obtain for the difference between 1983 and 1985, $F(2, 112) = 7.54$; $p < 0.01$, and for the difference between 1985 and 1987, $F(2, 112) = 4.93$; $p = 0.01$). Clearly, time-adjacent differences are of major importance when the timing of intervention effects is of interest, or when bursts in development are studied, e.g., in adolescent development. Additional options for analysis were listed above.

Change in Variation (2.2). To determine whether variance homogeneity changes for the two gender groups, over the course of development, we perform a Levene test at each of the three observation points in time. We obtain, for 1983, $F(1, 112) = 8.50$, with $p < 0.01$, for 1985, $F = (1, 112) = 0.68$, with $p = 0.41$, and for 1987, $F(1, 112) = 0.01$, with $p = 0.94$. We conclude that, over the course of pubertal development, the two gender groups approach each other in their variability of VAAA. We also note that the heterogeneity in 1985 could be considered a violation of ANOVA assumptions.

Structuring Change (2.3). In ANOVA applications, we operate at the manifest variable level. In structural modeling, we ask whether superordinate, latent variables can be used to explain the covariation among variables. Here, we ask whether a polynomial structure can explain the covariation of the three VAAA observations with each other and with Gender (Bollen & Curren, 2006). Figure 2 displays the model that was tested. It was estimated using maximum likelihood in LISREL. The parameter for the intercept was freely estimated, but set equal for the three observation points. Model fit was excellent (chi2 (5) = 6.39; $p = 0.27$; RMSEA = 0.05; CFI = 0.98; and GFI = 0.97). The residual plot was near perfect.

This result suggests that the covariation among the four observed variables can be well explained based on the assumptions that (1) a system of orthogonal polynomials up to second order captures the developmental pattern, and (2) the two gender groups differ in elevation but not in the shape of their developmental trajectories.

The difference between the ANOVA results in Section 2.1 and the present modeling results are conceptual and
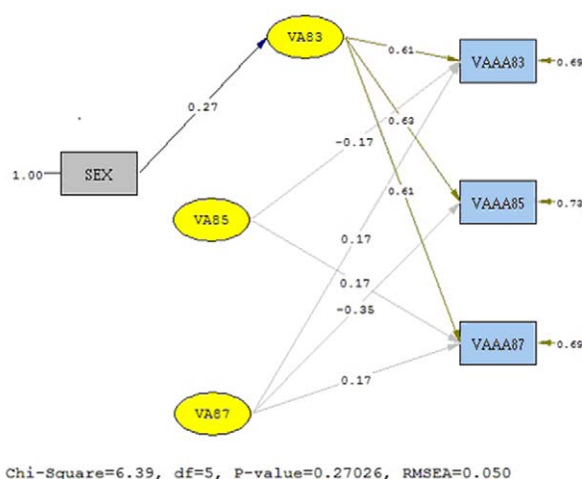

Chi-Square=6.39, df=5, P-value=0.27026, RMSEA=0.050

Figure 2. Latent Growth Curve Model of Adolescent Development of Verbal Aggression

of importance to theory. In Section 2.1, the observed scores were described using orthogonal polynomials. Here, we explain the covariation among the observed variables by unobserved, that is, latent variables that are defined by orthogonal polynomials. The definition of the latent variables can be seen as independent of the specific observed variables.

The very same model can be used to model other variables of aggression development. For example, the exact same model for aggressive impulses is even better than the one for VAAA (chi2 = 2.34; RMSEA = 0.0), but describes physical aggression less well (chi2 = 31.23; RMSEA = 0.22). In other words, the latent variable structure has a more general abstract meaning than the comparable manifest variable model.

Configurations of Aggression Development (2.4). Taking a person-oriented perspective (Bergman & Magnusson, 1997), one can ask whether the statements that are made at the aggregate, variable-oriented level apply to every member of a population to the same degree. It has been shown that this is rarely the case (von Eye & Bergman, 2003). We now dichotomize the sample in the VAAA variables at the mean, and perform a Configural Frequency Analysis (CFA; Lienert, 1969; von Eye et al., 2010).

CFA allows one to identify sectors in the data space that contain cases at rates that are significantly different than expected. Indeed, we find that two sectors contain more adolescents than expected, thus constituting CFA types. More girls were found than expected who show below average VAAA over the entire observation span of four years ($z = 3.81$; $p < 0.01$). In addition, more girls were found than expected who show above average VAAA over the entire observation span of four years ($z = 3.31$; $p < 0.01$). In all other sectors of the data space, the hypothesis prevails that the four variables under study are unrelated. We conclude that the gender-specific trajectories and covariance patterns discussed in the earlier sections are carried by the individuals in just two density centers in the data space.

Discussion

In this essay, examples were given of questions that can be asked concerning parameters that are estimated in developmental research. Two general lines of arguments were followed. First, researchers were encouraged to ask many and detailed questions, even using the same data. Standard methods of data analysis such as ANOVA are far more flexible than their routine application may suggest.

Second, researchers were encouraged to ask whether group-level statements at the manifest or the latent level indeed apply to every member of a population to the same degree. More often than not, just a few subpopulations show the effects that are found at the group level.

References

Bergman, L.R. (1988). You can't classify all of the people all of the time. *Multivariate Behavioral Research, 23*, 425–441.

Bergman, L.R., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology, 9*, 291–319.

Bock, R.D. (1975). *Multivariate statistics in behavioral research.* New York: McGraw-Hill.

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

Finkelstein, J. W., von Eye, A., & Preece, M. A. (1994). The relationship between aggressive behavior and puberty in normal adolescents: A longitudinal study. *Journal of Adolescent Health, 15*, 319–326.

Lienert, G.A. (1969). Die "Konfigurationsfrequenzanalyse" als Klassifikationsmethode in der klinischen Psychologie. In M. Irle, (Ed.), *Bericht über den 16. Kongreß der Deutschen Gesellschaft für Psychologie in Tübingen*, 1968 (pp. 244–255). Göttingen: Hogrefe.

von Eye, A., & Bergman, L.R. (2003). Research strategies in developmental psychopathology: Dimensional identity and the person-oriented approach. *Development and Psychopathology, 15*, 553–580.

von Eye, A., Mair, P., & Mun, E.Y. (2010). *Advances in Configural Frequency Analysis*. New York: Guilford Press (in press).

von Eye, A., Mun, E.Y., & Indurkhya, A. (2004). Classifying developmental trajectories - a decision making perspective. *Psychology Science, 46*, 65–98.

Studying Patterns of Individual Development by I-States as Objects Analysis (ISOA)

Lars R. Bergman
Stockholm University, Sweden
E-mail: lrb@psychology.su.se

and

Jari-Erik Nurmi
University of Jyväskylä, Finland
E-mail: jari-erik.j.nurmi@jyu.fi

Human development can be viewed as a complex stochastic process partly unique to the individual, a process that is characterized by strong interactions and nonlinearity across individuals. From this perspective, a research paradigm has been developed in which individual functioning and development are regarded in holistic-interactionistic terms (Magnusson & Törestad, 1993). The holistic-interactionistic research paradigm has led to a person-oriented approach that has two main facets: (1) It focuses on individual development and (2) it relies on a system view in which whole-system properties are central. The key assumption is that the concepts involved cannot be primarily regarded as separate entities but that they all together form a meaningful meta unit, a *Gestalt* (Bergman, 2001; Bergman, Andershed & Andershed, 2009; Bergman, von Eye, & Magnusson, 2006; Magnusson, 1985). This approach provides a natural theoretical framework for the new developmental science (Cairns, Elder, & Costello, 1996).

At the level of statistical analysis, the person-oriented theoretical framework leads to an approach in which individuals are studied on the basis of their patterns of individual characteristics rather than on the basis of separate variables, as is the case in standard variable-oriented analyses. This focus is congenial to a dynamical systems perspective which suggests a key property of real-life processes is their tendency to crystallize in a limited number of attractor states, characterized by typical configurations of the involved components (Bergman & Magnusson, 1997). Methodological solutions for carrying out person-oriented research have to a large extent been based on classification procedures for finding typical patterns of values in the studied system components.

From a developmental view, a special interest of a person-oriented perspective is studying the development of typical individual patterns across time. Different methods are possible here and this article gives a short overview of one method, namely I-States as Objects Analysis (ISOA). The idea behind ISOA is borrowed from the study of dynamic systems (Barton, 1994; Boker, 2002; Granic & Hollenstein, 2006). In these kinds of studies a limited number of emerging attractor states are often found that are essentially the same across the studied time period. Following this idea, it is assumed in ISOA that the same typical patterns occur at the different ages (although individuals might change the typical pattern they belong to and the frequencies of the typical patterns may vary across age). The common classification structure is first identified and then used to study stability and change. This is what ISOA is designed to do (Bergman & El-Khouri, 1999).

Outline of ISOA

A central concept in ISOA is the *i-state*. It is defined as an individual's pattern of values at a specific age (=measurement occasion) in the variables that are to be used for the classification. Hence, each individual is characterized by one i-state at each age and his number of i-states is equal to the number of ages studied. An ISOA analysis involves the following steps:

Step 1. Longitudinal data are necessary where the same variables have been measured at all ages. On the basis of such data, a new data file is formed consisting of all i-states for all individuals. This constitutes the input for a classification analysis. Hence, the analytical unit in the classification

Table 1. Correlation matrix for Swedish and Mathematics at age 10 and 13 (n = 914).

		Age 10		Age 13			
		Swed.	Math.	Swed.	Math.	Mean	sd
Age 10	Swedish	−	.66	.75	.59	3.12	0.98
	Mathematics		−	.61	.72	3.18	0.94
Age 13	Swedish			−	.68	3.11	0.94
	Mathematics				−	3.14	1.04

analysis is the single i-state and the total number of units in the analysis is the number of individuals times the number of ages studied. For this purpose, cluster analysis can be used, e.g. based on the LICUR rationale (Bergman, Magnusson, & El-Khouri, 2003), but other methods are also possible. To illustrate the principle of the classification: Assume that four individuals are studied at three ages in five variables. This means that each individual provides three value profiles (=three i-states) to the classification analysis, one for each age, and the total number of analyzed profiles will in this case be 4 x 3 =12. A classification analysis is then made of the twelve analytical units, each with a value profile in the five variables. Let us assume that two classes (=typical patterns) were found. Then each i-state is given a class membership code that is 1 or 2.

Step 2. Each individual's sequence of class memberships (one for each age) is used as input for further analyses: (1) A description of the time-invariant classification in terms of the characteristics of the typical i-states, how frequent they are, and the quality of the classification (e.g. cluster homogeneity, replicability, explained variance, and so on (see Bergman et al., 2003); (2) an investigation of *frequency structure stability/change* by comparing the frequency of the different typical i-states between ages; and (3) an examination of *individual stability/change* in belonging to different typical i-states across ages, using exact significance tests to find typical developmental paths.

Example of an ISOA Analysis: Individual Development of School Grade Patterns

The purpose of the example is to show how an ISOA analysis can be carried out. What is presented is not a thorough treatment of the issues but an illustration of the potentiality of the ISOA approach. An important question for educators is the extent to which school achievement in the two most basic school subjects, the native language and mathematics, crystallizes and remains stable. It is well-known that school grades in these two subjects are rather stable during the school years (interindividual stability). However, the question can be reformulated in person-oriented terms as follows: To what extent are *patterns* of school grades in the native language and mathematics stable over time? From a *structural* point of view, we can ask whether the frequencies of different typical patterns are similar across age. From an individual viewpoint, we can ask to what extent individuals belong to the same typical patterns across age.

In the example, school grades in Swedish and Mathematics at age 10 and 13 were studied but ISOA can easily be extended to more than two time points and more than two school subjects. Data were used from the longitudinal

research program Individual Development and Adaptation (IDA; Magnusson, 1988) for 914 boys and girls with complete data from both ages. The school grades were given on a scale from one to five. Cluster analysis could have been used to identify an age-invariant classification structure but, for simplicity, a theoretically defined nine categories classification of school grade patterns was used (see Nurmi & Aunola, 2005) and hence, the classification analysis in Step 1 of ISOA was skipped. The nine categories were: *1. Low achievement, even grades, 2. Low Achievement, Swedish > Mathematics, 3. Low Achievement, Mathematics > Swedish, 4. Average achievement, even grades, 5. Average Achievement, Swedish > Mathematics, 6. Average Achievement, Mathematics > Swedish, 7. High achievement, even grades, 8. High Achievement, Swedish > Mathematics, and 9. High Achievement, Mathematics > Swedish.* There are four possible grade patterns not covered by the classification system but no child had any of these patterns.

Table 1 reports some results using standard variable-oriented analyses: the means, standard deviations, and correlations are very similar at both ages, suggesting stationarity. The stability coefficients are high.

Next, an ISOA analysis was carried out using the nine categories classification system. The results show that the frequencies of belonging to the different grade pattern categories are quite similar for the two grades (Figure 1). For instance, about half the sample has equal grades in Swedish and Mathematics at both ages with the most frequent pattern being average in both Swedish and Mathematics. Hence, structural frequency stability is high. There is, however, one clear frequency change between ages. Almost twice as many belong to the category *High Achievement, Mathematics > Swedish* at age 13 than at age 10 (8.0 % as compared to 4.4 %).

The individual category stabilities are also considerable with every grade pattern category at age 10 significantly linked to the same grade pattern category at age 13. Specifically, starting with a high achievement uneven pattern strongly increases the probability above chance to belong to the same uneven pattern category three years later. Three significant individual category changes were also found but are not interpreted here. Further analyses are, of course, possible. For instance, students who are characterized by a low achievement pattern at age 10 almost never end up belonging to a high achievement pattern at age 13 (only four children out of 292, not shown in Figure 1). These four children could be studied, looking for explanatory factors of this unexpected beneficial development.

An example of an empirical study using an ISOA procedure in clustering by cases is given by Nurmi and Aunola's (2005) study of patterns of task-motivation during the first school years. A "simple" ISOA can be performed using a

Grade pattern category, age 10		Grade pattern category, age 13
Low achievement Swe=Ma	**11.1** — 3.3x → **13.3**	Low achievement Swe=Ma
Low achievement Swe>Ma	2.2x 2.1x **8.1** — 3.0x → **11.4****	Low achievement Swe>Ma
Low achievement Ma>Swe	**12.8** — 3.6x → **9.2****	Low achievement Ma>Swe
Average achievement Swe=Ma	**22.2** — 1.8x → **21.9**	Average achievement Swe=Ma
Average achievement Swe>Ma	**8.2** — 3.8x → **8.4**	Average achievement Swe>Ma
Average achievement Ma>Swe	**10.3** — 3.4x → **8.5**	Average achievement Ma>Swe
High achievement Swe=Ma	**18.5** — 2.9x → **15.9**	High achievement Swe=Ma
High achievement Swe>Ma	**4.5** — 10.8x → **3.4**	High achievement Swe>Ma
High achievement Ma>Swe	2.6x **4.4** — 6.0x → **8.0*****	High achievement Ma>Swe

Figure 1. Grade pattern categories at age 10, 13. Frequencies (figures in bold) and significant longitudinal streams (arrows).

Note. An arrow indicates p<.05 when testing for a significant developmental stream (after Bonferroni correction for 81 tests). A figure on an arrow indicates the ratio between the observed and expected frequency for that developmental stream.
** p<.01, *** p<.001 when testing the proportion difference between ages.

standard statistical package but it is more convenient and powerful to instead use either SLEIPNER (2006) or ROPstat (2009), which both include beta versions of ISOA modules.

Discussion

If the person-oriented view of development is accepted as relevant for a specific research problem, then the development of patterns is of focal interest. However, there are different approaches for taking pattern information into account. What is particular for ISOA is the specific way in which the typical patterns are searched for, as well as the focus on stability and change in the patterns individuals show across time. A common classification structure is assumed that applies to all ages and it is arrived at either from theory, or, more commonly, from a classification analysis of all i-states.

There are two main advantages of the ISOA approach as compared to other classification approaches:

1. The use of a common classificatory grid at all ages much simplifies comparisons across age and allows for more easily interpretable findings. For instance, ISOA provides a basis not only to investigate the frequency of different typical patterns at different ages but it also provides a tool to investigate various developmental trajectories that individuals show in typical pattern membership across time. ISOA also provides a basis for investigating the direction of influence across different ages, as previous cluster membership can be statistically controlled for before entering other predictors. Moreover, when studying equifinality and multifinality (Cicchetti & Rogosch, 1996) using ISOA makes possible a straightforward identification of developmental trajectories that belong to one or the other category (Bergman et al., 2009).
2. In cases with many studied ages, the number of objects analyzed in the classification analysis becomes fairly large even for a fairly small sample of individuals, which decreases the sensitivity of the classification structure to sampling variation.

The use of ISOA also includes some restrictions. For example, to use ISOA one must assume that approximately the same classification structure applies at all ages (although the proportion of individuals belonging to the different classes might vary between ages). This assumption is questionable if the studied age period is long and/or includes different developmental stages. However, for short-term development it can be expected that ISOA will often be suitable. It should then be recognized that it is not necessary to assume that there exists a single "true" time-invariant classification structure and that this structure must been found by the classification analysis. It is sufficient that the structure reasonably well reflects the pattern structure of the data, even if sound classification analyses performed at each age separately would reveal slightly different classification structures.

To the reader familiar with longitudinal configural frequency analysis (CFA, von Eye, 2002) it is clear that the analyses of the sequences of typical pattern memberships in Step 2 of ISOA can be carried out using methods developed within the CFA-paradigm (Bergman, 2000; Niedermeier & von Eye, 1999; von Eye & Bogat, 2006). Within this paradigm, pattern-based mediation analysis could also be carried out (von Eye, Mun, & Mair, 2009). Other methodological tools are also possible, such as applying methods for the study of Markov chains (Feller, 1958). Ongoing work concerning ISOA includes the development of methods for establishing the pattern of stability and change over time at the level of the single individual. Then the i-state classification structure is no longer sufficient and it is reverted to the original individual patterns. Each person's degree of self-similarity and change is estimated, and this information is related to information from other variables (of which one could be the typical i-state belongingness).

Author Note

of the collection of data before 1996 was David Magnusson. The data collections and database were supported by grants from the Swedish National Board of Education, the Swedish Committee for the Planning and Coordination of Research, The Bank of Sweden Tercentenary Foundation, and the Swedish Social Research Council.

References

Barton, S. (1994). Chaos, self-organisation, and psychology. *American Psychologist, 49*, 5–14.

Bergman, L. R. (2000). I-states as objects analysis (ISOA) – a way to generate sequences of categories for longitudinal (CFA) analysis. *Psychologische Beiträge, 42*, 337–346.

Bergman, L. R. (2001). A Person Approach in Research on Adolescence: Some Methodological Challenges. *Journal of Adolescent Research, 16*(1), 28–53.

Bergman, L. R., Andershed, H., & Andershed, A.-K. (2009). Types and continua in developmental psychopathology: Problem behaviors in school and their relationship to later antisocial behavior. *Development and Psychopathology, 21*, 975–992.

Bergman, L.R., & El-Khouri, B. M. (1999). Studying individual patterns of development using I-states as objects analysis (ISOA). *Biometrical Journal, 41*, 753–770.

Bergman, L. R. & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology, 9*, 291–319.

Bergman, L.R., Magnusson, D., & El-Khouri, B.M. (2003). *Studying individual development in an interindividual context: A person-oriented approach*. Vol. 4 in the series *Paths through life* (D. Magnusson, Ed.). Mahwah, NJ: Erlbaum.

Bergman, L.R., von Eye, A., & Magnusson D. (2006). Person-oriented research strategies in developmental psychopathology. In D. Cicchetti & D. J. Cohen, Eds., *Developmental psychopathology. Volume One: Theory and method*. (pp. 850–888). New York: Wiley.

Boker, S. M. (2002). Consequences of continuity: The hunt for intrinsic properties within parameters of dynamics in psychological processes. *Multivariate Behavioral Research, 37(3)*, 405–422.

Cairns, R. B., Elder, G. H., Jr., & Costello, E. J. (1996). *Developmental Science*. Cambridge: Cambridge University Press.

Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology, 8*, 597–600.

Feller, W. (1958). *An Introduction to Probability Theory and Its Applications, Vol. 1*. New York: Wiley.

Granic, I., & Hollenstein, T. (2006). A survey of dynamic systems methods for developmental psychopathology. In D. Cicchetti & D. J. Cohen, Eds., *Developmental psychopathology. Volume One: Theory and method* (pp. 889–930). New York: Wiley.

Magnusson, D. (1985). Implications of an interactional paradigm for research on human development. *International Journal of Behavioral Development, 8*, 115–137.

Magnusson. D. (1988). *Individual development from an interactional perspective*. Hillsdale, NJ: Erlbaum.

Magnusson, D., & Törestad, B., (1993). A holistic view of personality: A model revisited. *Annual Review of Psychology, 44*, 427–452.

Niedermeier, K. E., & von Eye, A. (1999). *Statistical Analysis of Longitudinal Categorical Data in the Social and Behavioral Sciences: An Introduction with Computer Illustrations*. Mahwah, NJ: Erlbaum.

Nurmi, J.-E., & Aunola, K. (2005). Task motivation during the first school years: A person-oriented approach to longitudinal data. Learning and Instruction, 15, Issue 2, 103–122.

ROPstat (2009). Statistical package developed by Andras Vargha and Peter Bansagi at Karoli Gaspar Reformed Church University, Hungary. ROPstat is found at www.ropstat.com

SLEIPNER (2006). Statistical package developed by Lars R. Bergman and Bassam M. El-Khouri, Stockholm University, Sweden. SLEIPNER, which is freeware, is found at www2.psychology.su.se/sleipner/

von Eye, A. (2002). *Configural Frequency Analysis: Methods, Models, and Applications*. Mahwah, NJ: Lawrence Erlbaum.

von Eye, A., & Bogat, G.A. (2006). Person orientation - concepts, results, and development. *Merril Palmer Quarterly, 52*(3), 390–420.

von Eye, A., Mun, E. Y., & Mair, P. (2009). What carries a mediation process? Configural analysis of mediation. *Integrative Psychological & Behavioral Science, 43*(3), 228–247.

Analyzing Total, Direct and Indirect Causal Effects in Intervention Studies

Rolf Steyer, Christiane Fiege and Norman Rose
Friedrich Schiller University, Jena, Germany
E-mail: rolf.steyer@uni-jena.de

In longitudinal studies we aim at describing certain attributes of a single individual or group of individuals over time, studying change between different time points. Sometimes there are interventions to which all individuals are exposed (e.g., schooling) and sometimes one group of individuals is exposed while another group is not. In both cases we often are interested in the effects of interventions to which individuals might ($X = 1$) or might not be exposed ($X = 0$) on one or several outcome variables Y.

At first sight, in randomized intervention studies, analyzing the effect of the intervention seems to be an easy task. In fact, if randomization does not fail (e.g., due to systematic attrition), we can estimate the *total effect* of the intervention on the outcome variable Y by the difference between the means of Y in the two groups represented by ($X = 1$) and ($X = 0$). Furthermore, we can test the hypothesis that the total effect of the treatment is zero via the *t*-test for independent groups. If we are interested in the *conditional effects* of the intervention — e.g., given high, medium, or low pre-test scores — in order to see if the intervention effects differ between these conditions, we still can use traditional techniques of regression and/or analysis of variance in order to estimate and test these conditional effects.

However, traditional statistical tools are not sufficient any more when it comes to analyzing *direct and indirect effects* of the treatment with respect to a mediator variable, say M, even if we presuppose a perfect randomized

experiment. Not only that, we need to know the distinction between total, direct and indirect effects; we also have to rely on certain assumptions that are not guaranteed to hold even in a randomized experiment! While randomization guarantees that the mean difference between the intervention groups is an unbiased estimate of the total effect of the intervention, *randomization does not insure* that the conditional mean differences between intervention groups given fixed values *m* of the mediator variable *M* are unbiased estimates of the direct effect of the intervention for that value *m* of the mediator *M.* However, this is what many authors seem to assume when analyzing direct and indirect intervention effects, e.g., using well-known path analysis techniques (see, e.g., MacKinnon, 2008, Ch. 3).

Biased estimates of intervention effects are often disastrous for our substantive conclusions. If there is bias, this can mean that the true effects are strongly positive while we have estimates indicating that they are strongly negative and vice versa. We may also have estimates indicating no intervention effects although the intervention effects are strongly positive or strongly negative. This can be disastrous in the analysis of undesirable side effects of interventions. Note that bias is not a phenomenon due to inevitable sampling error. Instead we are talking about *systematic bias* that would even occur in infinitely large samples. In other words, bias does not only pertain to sample means and their differences, but also to *population means* and their differences.

Beyond the randomized experiment, ordinary mean differences between intervention groups usually do not estimate any more the total treatment effect. Instead, there are two kinds of biases that distinguish this mean difference from an estimate of the total intervention effect: the baseline bias and the effect bias. The *baseline bias* is due to selection into the intervention groups determined by *severity of disorder* or by *motivation for treatment*, for instance. If those subjects with a severe disorder tend to be in the intervention group rather than in the untreated control group and treatment effects are only small or medium, then the mean differences in the outcome variable between treatment and control can still be negative — erroneously indicating a negative treatment effect — even though the treatment effect is positive for each and every individual. In contrast, the *effect bias* is due to selection into the intervention groups determined by the size of the expected treatment effect. If diagnostics suggest that the treatment considered would be beneficial for the subject, and if this determines treatment assignment, then the differences between treatment and control in the outcome variable will not unbiasedly estimate the effect of the intervention.

To conclude: Outside the randomized experiment, traditional statistical methods including regression, analysis of variance, and structural equation modeling, are not sufficient to estimate and test intervention effects. Within the perfect randomized experiment, these skills only suffice to estimate and test the total effect as well as the conditional effects given covariates such as pre-tests, gender, or educational status. The analysis of direct and indirect effects — even within the randomized experiment, just like the analysis of causal effects in quasi-experiments — inevitably involves the *theory of causal effects* and its implications for the *design of studies* and *techniques of data analysis* that aim at the analysis of causal effects. In the remainder of this paper, we will outline these points. This outline is a summary of *Probability and Causality* (Steyer, et al., in press). We refer the reader to this book, both for the details of the theory and for the references to the contributions of other authors on this topic.

The Theory of Causal Effects

The theory of causal effects is based on atomic stratification. As already noted above, mean differences in the total population and in subpopulations can be biased. Hence, the basic idea is to construct causal effects on the conditional expectations of Y given $X = x$ in the smallest subpopulations. Oftentimes, these smallest subpopulations are the observational units (e.g., students, clients with a particular disorder, etc.), and in this case we can built the theory on the conditional expectations $E_{X=x}(Y \mid U = u)$ of the outcome variable Y given the observational unit u in treatment condition x. However, since, in many cases, there is still systematic variability within the observational units, we base the general theory on the conditional expectations of Y given $X = x$ in the most fine-grained strata, the *atomic strata*. These atomic strata are obtained by conditioning on *all* potential confounders, i.e., on all those variables that are prior or simultaneous to the treatment variable X. Typical and important variables that are *prior* to treatment are pre-tests of the outcome variable (e.g., *achievement before treatment*, *aggressivity before treatment*, if the Y is achievement and aggressivity, respectively). An example for a variable that is *simultaneous* to the treatment variable X (e.g., child's training vs. no child's training) is a second treatment variable Z (e.g., mother's training vs. no mother's training) that varies simultaneously to X.

True Outcome Variables and True Total Effects

The random variable τ_1 whose values are these conditional expectations of Y in the atomic strata given treatment ($X = 1$) is called the true-outcome variable in treatment 1. A particular value of τ_1 is the expected value of the outcome variable Y under intervention ($X = 1$) given a particular combination of values on all potential confounders. Correspondingly, the random variable τ_0 whose values are the conditional expectations of Y in the atomic strata given control ($X = 0$) is called the true-outcome variable under control (see Steyer et al., in press, for mathematical details). The difference $\delta_{10} = \tau_1 - \tau_0$ is then called the *true total effect variable*. Its values are the *true total effects* in the most fine-grained strata of the potential confounders. By definition, these true effects are unbiased, because, defining them, we condition on *all* potential confounders that might induce bias.

Average Total Effect and Conditional Total Effects

The expectation $E(\delta_{10}) = E(\tau_1) - E(\tau_0)$ of these true total effect variables over the distribution of the strata defines the

average total effect of intervention 1 compared to intervention 0. Similarly, various kinds of conditional total effects are defined by the difference between various conditional expectations of the true-outcome variables. Each of these kinds of conditional total effects provides specific information that might be of interest in intervention studies. An example in case is the *conditional total effect* $E(\delta_{10} \mid Z = z) = E(\tau_1 \mid Z = z) - E(\tau_0 \mid Z = z)$ given the value z (e.g., *very severe*) of the covariate Z (e.g., *severity of symptoms*).

True Direct and True Indirect Effects. The definition of the *true direct effect variable* is very similar to the definition of the true total effect variable. The definition implies that the direct effect does not have to be a single value but can vary depending on other variables. However, this time we not only condition on *all* potential confounders, i.e., on all those variables that are prior or simultaneous to the treatment variable X, but also on a specified mediator variable M that is 'posterior' or 'subsequent' to the treatment and prior to the considered outcome Y. In the simplest case, M is univariate variable; however, in general it may also consist of several univariate variables, i.e., $M = (M_1, \ldots, M_K)$. Hence, the random variable whose values are the conditional expectations of Y given an atomic stratum, a value of M, and treatment $X = 1$ is denoted by $\tau_{1;\,M}$. Its counterpart for control $(X = 0)$ is $\tau_{0;\,M}$. The difference $\delta_{10;\,M} = \tau_{1;\,M} - \tau_{0;\,M}$ is then defined to be the *true direct effect variable* with respect to M, and the *true indirect effect variable* is the difference $\delta_{10} - \delta_{10;\,M}$.

Average and Conditional Direct and Indirect Effects

Again, the expectation $E(\delta_{10;\,M}) = E(\tau_{1;\,M}) - E(\tau_{0;\,M})$ of these true direct effect variables over the joint distribution of the strata and the mediator defines the *average direct effect* of intervention 1 compared to intervention 0. If this average direct effect is not informative enough, we may also be interested in the *conditional direct effect* $E(\delta_{10;\,M} \mid M = m) = E(\tau_{1;\,M} \mid M = m) - E(\tau_{0;\,M} \mid M = m)$ given the value m of the mediator M, or in the conditional direct effect $E(\delta_{10;\,M} \mid Z = z) = E(\tau_{1;\,M} \mid Z = z) - E(\tau_{0;\,M} \mid Z = z)$ given the value z of the covariate Z. Comparing $E(\delta_{10;\,M} \mid M = m_1)$ to $E(\delta_{10;\,M} \mid M = m_2)$ informs us about the difference in the conditional direct intervention effects between mediator values m_1 (e.g., high motivation after treatment) and m_2 (e.g., low motivation after treatment). In this way we can study if and how the direct intervention effect differs depending on the values of the mediator variable M. Similarly, comparing $E(\delta_{10;\,M} \mid Z = z_1)$ to $E(\delta_{10;\,M} \mid Z = z_2)$ informs us about the difference in the conditional direct intervention effects between covariate values z_1 (e.g., severe symptoms before treatment) and z_2 (e.g., no severe symptoms before treatment). In this way we can study how the direct intervention effect is modified by the values of the covariate Z. Of course, we may also consider the conditional expectations of $\delta_{10;\,M}$ given both m and z.

Similarly, we may consider the *average indirect effect* $E(\delta_{10} - \delta_{10;\,M}) = E(\delta_{10}) - E(\delta_{10;\,M})$ and different kinds of *conditional indirect effects*, the *conditional indirect effects* $E(\delta_{10} - \delta_{10;\,M} \mid M = m) = E(\delta_{10} \mid M = m) - E(\delta_{10;\,M} \mid M = m)$ given a value m of the mediator M, and the *conditional indirect effects* $E(\delta_{10} - \delta_{10;\,M} \mid Z = z) = E(\delta_{10} \mid Z = z) - E(\delta_{10;\,M} \mid Z = z)$ given the value z of a covariate Z.

Unbiasedness

The various kinds of causal effects defined above are of a purely theoretical nature, because they involve the theoretical variables τ_0 and τ_1 and/or $\tau_{0;\,M}$ and $\tau_{1;\,M}$. Nevertheless, they define what we *would like* to estimate. Furthermore, the theory provides knowledge about conditions that allow for a causal interpretation of estimable parameters. What *can* be estimated by the corresponding sample means are the conditional expected values $E(Y \mid X = x)$ of the outcome variable Y in treatment conditions, the conditional expected values $E(Y \mid X = x, Z = z)$ of the outcome variable Y given treatment x and value z of the covariate Z, the conditional expected values $E(Y \mid X = x, M = m)$ of the outcome variable Y given treatment x and value m of mediator M, and the conditional expected values $E(Y \mid X = x, Z = z, M = m)$ of the outcome variable Y given treatment x, value z of covariate Z, and value m of mediator M. Under certain conditions — some of which can be created by random assignment of units to treatment conditions and/or by careful selection of the covariates in the (possibly multivariate) covariate $Z = (Z_1, \ldots, Z_Q)$ — these conditional expected values can be *unbiased*.

The conditional expected value $E(Y \mid X = x)$ of Y in treatment x is called *unbiased* if $E(Y \mid X = x) = E(\tau_x)$. Similarly, the conditional expected value $E(Y \mid X = x, Z = z)$ of Y given treatment x and value z of covariate Z is called *unbiased* if $E(Y \mid X = x, Z = z) = E(\tau_x \mid Z = z)$. Furthermore, the conditional expected value $E(Y \mid X = x, M = m)$ of Y given treatment x and value m of mediator M is called *unbiased* if $E(Y \mid X = x, M = m) = E(\tau_{x;\,M} \mid M = m)$, and finally, the conditional expected value $E(Y \mid X = x, Z = z, M = m)$ of Y given treatment x, value z of covariate Z, and value m of mediator M is called *unbiased* if $E(Y \mid X = x, Z = z, M = m) = E(\tau_{x;\,M} \mid Z = z, M = m)$.

Unbiasedness of the conditional expectations $E(Y \mid X = x)$ can be created by *random assignment* of the unit to one of the treatment conditions. Unbiasedness of the conditional expectations $E(Y \mid X = x, Z = z)$ can be created by *conditional random assignment* of the unit to one of the treatment conditions given value z of Z and/or the appropriate selection of the covariates in $Z = (Z_1, \ldots, Z_Q)$. In contrast, unbiasedness of the conditional expectations $E(Y \mid X = x, Z = z, M = m)$ can *only* be created by appropriate selection of the covariates in $Z = (Z_1, \ldots, Z_Q)$. *Selecting the appropriate covariates is the only option in quasi-experimental studies, but also in the analysis of direct and indirect effects in the randomized experiment.*

Identification of Causal Effects

If the conditional expected values $E(Y \mid X = x)$ are unbiased, then the *average total effect* of the treatment is $E(Y \mid X = 1) - E(Y \mid X = 0) = E(\tau_1) - E(\tau_0)$. It is this difference that we estimate by the mean difference between treatment groups in a randomized study.

Next, consider the equation

$$E(Y \mid X, Z) = g_0(Z) + g_1(Z)X, \qquad (1)$$

which always holds for $E(Y \mid X, Z)$ if X is dichotomous, where $g_0(Z)$ denotes the *intercept function* and $g_1(Z)$ the *effect function*. Both functions are usually unknown but estimable.

According to this equation, the regression of Y on X given value z of Z is linear with intercept $g_0(z)$ and slope $g_1(z)$, both being functions of Z. If the conditional expected values $E(Y \mid X = x, Z = z)$ are unbiased, then the values $g_1(z)$ are the *conditional total treatment effects* given the value z of the covariate Z, and the expectation $E[g_1(Z)]$ is the *average total treatment effect* $E(\tau_1) - E(\tau_0)$.

If M is a mediator, we may consider the equation

$$E(Y \mid X, Z, M) = h_0(Z, M) + h_1(Z, M)\, X, \qquad (2)$$

which always holds for the regression $E(Y \mid X, Z, M)$ if X is dichotomous. If the conditional expected values $E(Y \mid X = x, Z = z, M = m)$ are unbiased, then the values of $h_1(z, m)$ are the $(Z = z, M = m)$-conditional direct treatment effects, the conditional expectations $E[h_1(Z, M) \mid M = m]$ are the *conditional direct treatment effect*s given value m of the mediator M, the conditional expectations $E[h_1(Z, M) \mid Z = z]$ are the *conditional direct treatment effect*s given value z of the covariate Z and the expectation $E[h_1(Z, M)]$ is the *average direct treatment effect* $E(\tau_{1;\, M}) - E(\tau_{0;\, M})$.

The various kinds of *indirect effect*s are obtained by taking the differences between the corresponding total and direct effects. For instance, if both $E(Y \mid X, Z)$ and $E(Y \mid X, Z, M)$ are unbiased, $E[g_1(Z)] - E[h_1(Z, M)]$ yields the average indirect effect, whereas $E[g_1(Z) \mid Z = z] - E[h_1(Z, M) \mid Z = z]$ is the $(Z = z)$-conditional indirect effect and $E[g_1(Z) \mid M = m] - E[h_1(Z, M) \mid M = m]$ is the $(M = m)$-conditional indirect effect intervention effect.

Conclusions

Within perfect randomized intervention studies, mean differences between treatment groups can be causally interpreted as the *average total treatment effect*. Similarly, conditional mean differences between treatment groups given the value z of a covariate Z can be causally interpreted as the $(Z=z)$-*conditional total treatment effect*. Whenever feasible, randomization should be used when the total effect of an intervention or conditional total effects given covariates are of interest. If we want to study average and/or conditional total effects *beyond the randomized experiment*, i.e., in quasi-experimental and observational studies, we can also estimate various total effects provided that: (a) we select the appropriate covariate vector Z, (b) adequately estimate the regression $E(Y \mid X, Z)$ – which is not always linear – and (c) estimate and adequately test the various conditional and unconditional expectations such as $E[g_1(Z)]$ and $E[g_1(Z) \mid Z = z]$.

If we want to study direct and indirect effects in the randomized experiment or in quasi-experimental and observational studies, we have to (a) select an appropriate covariate vector Z such that the regression $E(Y \mid X, Z, M)$ is unbiased, (b) adequately estimate the regressions $E(Y \mid X, Z, M)$ – which again might be nonlinear – and (c) estimate and adequately test the various conditional and unconditional expectations such as $E[h_1(Z, M)]$, $E[g_1(Z, M) \mid Z = z]$, $E[g_1(Z, M) \mid M = m]$ or $E[g_1(Z, M) \mid Z = z, M = m]$. For these analyses of causal effects we recommend the program *EffectLite* (Steyer & Partchev, 2008).

A careful selection of all relevant covariates is essential for estimating the various causal effects mentioned above. Otherwise we cannot hope for unbiasedness. This also applies to the analysis of direct and indirect effects *within the randomized experiment*. Whenever possible always *include the pre-tests of the outcome variable and the pre-test of the mediator as covariates* in the vector Z of covariates! In many applications this will already prevent large biases in the estimation of the various causal effects. In a quasi-experiment, never trust an analysis of conditional or average total effects that does not include the pre-test of the outcome variable and even in a randomized experiment, never trust an analysis of conditional or average direct effects that does not include the pre-test of the mediator variable as a covariate.

Of course, what has been presented in this paper is just a short note on what needs and deserves a much longer and more detailed presentation. Many important topics such as propensity scores (see, e.g., Rosenbaum, 2002), instrumental variables (Greenland, 2000), and models for the analysis of individual effects (see e.g., Steyer, 2005) were not presented at all. Again, for these and other important issues we refer the reader to Steyer et al. (in press).

References

Greenland, S. (2000). An introduction to instrumental variables for epidemiologists. *International Journal of Epidemiology*, 29, 722–729.

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

Rosenbaum, P. R. (2002). Observational studies (2nd ed.). New York: Springer.

Steyer, R. (2005). Analyzing Individual and Average Causal Effects via Structural Equation Models. *Methodology: European Journal of Research Methods in the Behavioral and Social Sciences*, 1, 39–54.

Steyer, R. & Partchev, I. (2008). *EffectLite: User's manual: A program for the uni- and multivariate analysis of unconditional, conditional and average mean differences between groups.* University of Jena, Jena: Germany (available at www.causal-effects.de).

Reports from the Lab

Birth Cohort Studies: Opportunities and Challenges

John Bynner
Centre for Longitudinal Studies, Institute of Education
and Director, Longview, London, UK
E-mail: Johnbynner@longviewuk.com

Cohort studies involving the follow-up of a sample, or samples, from a specified population across time have a long history in social and behavioral science and medicine, with origins in a series of classic studies in the US, Scandinavia and more recently the UK and New Zealand (e.g., Phelps, Furstenberg, & Colby, 2002; Martin et al., 2006). Such studies range from birth cohorts following individuals from birth – or increasingly, from registration of pregnancy – to studies starting at different ages corresponding to key transitions, e.g., the early teens, school entry, school to work, retirement. Another variation is to start data collection much later than birth. The classic Swedish studies (Janson, 2000) typically began at age 10 or older, relying on government registers to supply the data about cohort members up to that age. Yet another option is to give immediate access to later developmental stages rather than wait for the birth cohort to mature by starting a series of 'quasi-sequential' cohorts, e.g., beginning at birth, 5, 10 and 16. But the 'gold standard' for all of them is the long term prospective birth cohort study.

At the simplest level, birth cohort studies enable developmental scientists to test hypotheses about the consequences for outcomes later in life related to circumstances, experience and individual attributes measured earlier (e.g., Robins & Rutter, 1990). Although there is no random allocation of subjects to different environmental conditions as in the 'true experiment', the temporal sequencing of data collection and the use of statistical controls for selection bias gives a 'quasi experimental' basis for causal inference that is missing from the cross-sectional-survey (Shadish, Cook, & Campbell, 2001; Piesse, Judkins, & Kalton, 2009). Such possibilities make the birth cohort study a key research instrument for scientific disciplines ranging from medicine to psychology, sociology and economics. They also have much policy value in enabling policy makers to forecast the likely success of different intervention options and assess their effects.

The Millennium year prompted a huge expansion of birth cohort studies including the US National Institutes of Health-funded National Children' Study (NCS) based on a probability sample of 100,000 pregnancies in 105 sites phased in over a number of years starting in 2007 (www.nationalchildrensstudy.gov). More recently, *Etude longitudinale française depuis l'enfance (ELFE)*, the French birth cohort study, modeled in part on the NCS and based on 20,000 births after many years in preparation has now received government go-ahead for a 2010 start (www.elfe-france.fr). EUCCONET is a European Science Foundation-funded network chaired by the director of London's Centre for Longitudinal Studies (CLS) – see below – that shares information and experience between the new and older birth cohort studies in Europe and liaises more widely with other birth cohort studies across the world (www.esf.org/nc/activities/research-networking-programmes/medical-sciences-emrc/current-esf-research networking-programmes/the-european-child-cohort-network-eucconet.html).

The British Birth Cohort Studies

Britain is unique in having established and maintained through the cohort members' lives a series of birth cohort studies (Bynner & Joshi, 2007). The series began with the 1946 birth cohort study (National Survey of Health and Development) followed by comparable studies starting in 1958 (National Child Development Study), 1970 (British Cohort study), 1992 (Avon Longitudinal Study of Parents and Children), 2000 (Millennium Cohort Study). All these studies with one exception – the area-based ALSPAC) – are based on some form of national probability sample, and enable national population parameters to be estimated. A new birth cohort study will begin in 2012. Figure 1 shows the studies' progress from birth onwards related to age and period (year) of data collection.

However, the designs have shifted over time to match changing analytic priorities beginning, in the case of the first three, with samples comprising all births in a single week averaging 17,000 (in the case of the 1946 study for cost reasons reduced in the first follow-up to 5,362). In the later studies, geographically defined clustered samples of all 14,000 births over two years in one area (1992 study) and all 21,000 births over one year in 400 electoral wards (2000 study) were followed. The new birth cohort study plans to take the clustering ('ecological') component further with a national probability sample of a year's births accompanied by area studies in contrasting localities. Data collection at the beginning of the study will also be more intensive, compared with the earlier studies, with three planned before the cohort members reach the age of 2. Frequency of data collection in the other studies comprises 1 to 2 year intervals up to age 15 in the 1946 and 1992 studies, up to age 9 in the 2000 study and, as dictated by financial constraints, 4 to 8 year intervals in the 1958 and 1970 studies.

The Centre for Longitudinal Studies (CLS) located in London's Institute of Education[1] is responsible for three of these studies, the 1958, 1970 and 2000 (Millennium) birth cohort studies – known respectively as:

National Child Development Study (NCDS) – with data collected at birth, 7, 11, 16, 23, 33, 37, 42, 46, 50 including a survey of the cohort members' children at cohort age 33 (one third sample), a basic skills assessment (10% sample) at age 37, and a biomedical assessment of the whole sample at age 44.

1970 British Cohort Study (BCS70) – with data collected at 5, 10, 16, 26, 30, 34, 38, including a basic skills assessment
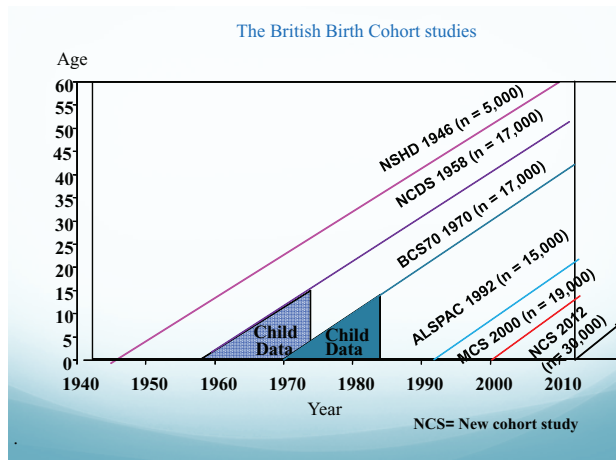
Figure 1. Gender-specific Development of Verbal Aggression during Adolescence

at age 21 (10% sample), and at age 34 (whole sample) when there was also a study of cohort members' children (one third sample).

Millennium Cohort study (MCS) with data collected at birth, 9 months, 3, 5, 7.

Under the terms of funding[2], all the *anonymised* data collected in these studies are available for scientific use from the UK Economic and Social Data Service through download from the Data Archive at the University of Essex. The CLS also has close connections with the 1946, 1992 studies enabling analysis to be conducted collaboratively across a series of birth cohort studies spanning a 54-year period up to 2000 and involving over 60,000 cohort members. This short paper sets out the scientific opportunities and challenges for users of the data.

Design and Analysis

Analysis and design of longitudinal data go hand-in-hand because the design of the study reflects the changing analytic targets to which the research using the data is likely to be directed. Through the period of early childhood from birth (or pregnancy) to school entry, the focus will be on the medical conditions associated with pregnancy and birth, family circumstances and early developmental indicators in cognitive, behavioral and emotional domains. Later the focus will shift towards schooling and its effects, and the influence of the peer group and local community institutions such as hospitals and libraries to which the child relates. Further on, the transition from school to work or further and higher education becomes a critical issue. Training experience and the developing occupational identity, then partnership and family formation will follow as will exercising the rights and responsibilities of citizenship. In the later stages of life, health once more comes strongly into the picture as biological and physical functions decline as part of old age.

Rationale and design principles. Increasing interest in the field of bio-social ecology development, including 'gene-environment interplay' and 'accumulated risk' (Bronfenbrenner, 1977; Lerner, 1998; Shanahan & Boardman, 2009; Schoon, 2006), 'linked lives', 'timing' of events, 'location' (in time and space) and 'the exercise of personal agency',

Intergenerational Continuities

including turning points in shaping the life course (Giele & Elder, 1998; Elder & Giele, 2009; Heinz, Huinink & Weyman, 2009) is reflected in the move towards multilevel design and analysis and ever-earlier data collection. 'Linked lives' are captured by data collection from both parents and siblings in the family context; then the inquiry is extended to 'significant others' such as, teachers, employers, peers and spouses, who serve as key factors in development and such phenomena as turning points. Timing is reflected in the intersection of the three confounded 'extrinsic effects' on life course development: Age reached (A), Period of data collection (P) and Cohort born into (C), where $A=P+C$ (Schaie, 2000). Holding one of the three effects constant, and assuming constancy in the other, identifies the influence of the third. Thus the comparison of health outcomes and the processes giving rise to them at roughly the same age, (early thirties) across the 1946, 1958 and 1970 cohorts, enables the effect of societal changes (cohort effect) on morbidity to be assessed (Ferri, Bynner, & Wadsworth, 2003). Location in time *and* space (history and geography) brings in the fourth of what Alwin (1995) defines as the 'space-time coordinates' of longitudinal study encompassing culturally based institutional differences between and across regions and within and between countries (in the case of the British studies, Scotland, Wales and England). Finally, the exercise of agency comprising the capabilities through which individuals construct their lives and contribute to the life of the next generation, argues for intergenerational enquiry whereby data collection is extended to cohort members' own children. A one third representative sub-sample at age 33/34 of the 1958 and one half of the 1970 cohort members (to accommodate the cohort shift of later childbearing in the latter cohort) offers the opportunity for study of 5,000 children in each of them.

The attraction of the life course perspective is that it attempts to embrace all the foregoing elements of developmental processes and outcomes in and across different life course domains at different ages and stages. Salience of data for particular theoretical purposes is therefore a key issue for secondary researchers. Another is relevance of data for scientific perspectives that are changing, as new theoretical developments and methodology replace the old. Such considerations establish a set of principles for the design and use of birth cohort study data that underpin the British studies (Bynner, 2005).

- The longer data collection continues over cohort members' lives the greater the potential for understanding life course processes and their short term, medium term and long term outcomes, e.g., the influence on psychological well-being in adulthood of poor educational experience and qualifications.
- The larger the sample the more scope there is for examining sub-groups such as ethnic minorities or individuals with disabilities. The alternative is over-sampling the populations these sub-groups represent, but this risks a shift of scientific focus later on, e.g., the 1946 study over-sampled the best educational performers, whereas the modern emphasis tends to be on the disadvantaged. Through supplementary funding the size of the Millennium Cohort Study samples in Scotland, Wales and Northern Ireland (included for the first time) was doubled to give scope for separate national analyses and cross-national research.
- To capture the effects of socio-political and cultural change, successive cohorts are needed at regular intervals across time and cross-nationally.
- The larger the sample of parents, and the longer the period when new births to cohort members can be entered into the study for follow-up, the greater the opportunity for creating analytically viable four generation uncensored datasets – grandparents, parents, cohort members, (all) cohort members' children.
- The more frequently data are collected, especially in the early years, the less subject the prospective data are to bias, but the value of frequent inquiry needs to be set against the risk of overburdening respondents who in adolescence particularly may be difficult to persuade that they should continue in the study.

Analysis strategy. Longitudinal data collection supports all the classic statistical techniques designed for continuous and discrete data (Magnusson, Bergman, Rudinger, & Törestad, 1991; Lynn, 2009). Thus for continuous outcome variables OLS regression-based methods are the starting point leading to more comprehensive methods of testing and evaluating models of life course processes. The development of structural equation modeling with latent variables (Jöreskog & Sörbom, 1979) counts as one of the major methodological breakthroughs for life course study, offering the means of testing competing theories about developmental processes across different life domains while taking account of measurement error. The methodology can be applied not only within a single study, but can be given equivalent variables across time – i.e., using successive cohorts – and cross-nationally (Blossfeld, 2009). 'Person-centered' approaches

such as latent class analysis offer the complementary approach of identifying different types of individuals whose life courses can be tracked (Magnusson et al., 1991).

Extensions reflect the increasing complexity of designs, with multi-level methodology matching the effects of clustering in the sample design and analysis. Researchers use these methods to focus on the variance components of outcome variables at different levels of a multi-level structure, e.g., family effects, school effects and so on. For discrete outcome variables, there is much use of logistic regression techniques, again within what is increasingly a multivariate and multi-level framework. Where censoring of outcome data occurs, as in mortality and morbidity studies, survival analysis will be the preferred option using hazard models. Finally, when quasi-experiments are constructed using longitudinal data, as in case control studies of the effects of medical interventions, analysis of variance will be the preferred option.

Challenges

Users of birth cohort study data face a number of challenges in realizing their full analytic potential (Magnusson et al., 1991). In the case of the CLS studies we have a series of longitudinal datasets extending in practice, or in conception, across the whole life course and compiled over a considerable period by teams with different theoretical perspectives and analytic interests. Although the excellent documentation available makes the data sets relatively easy to use to use, some validity checks and cautions are needed when modeling life course processes with the datasets supplied.

Coverage. The phrase 'fading relevance' (Janson, 2000) means that measures taken at the beginning of a birth cohort study may no longer be optimal for contemporary research purposes. Those variables that are needed are often absent from the dataset and those that attract little interest often predominate. The analysis of a school's effect on educational attainment, for example, places importance currently on school climate or ethos, but such concepts were not prominent at the time of the earlier studies so there is no operationalization of them in the data. More subtle questions arise in connection with the validity of measurements that may no longer have the same salience and meaning in the contemporary context. It is important to determine whether repeated measures of such an attribute as cognitive ability, assessed over many years with an instrument that is changing to accommodate age, is really measuring the same thing over time.

Sample attrition and missing data. In a longitudinal study, attrition begins from the very minute a sample is constructed. Thus there will be loss of cohort members between the time of selecting the sample and carrying out the first wave of data collection. From then on in every subsequent wave numbers will decline as individuals refuse to take part or are no longer contactable through relocation to unknown addresses or through changing their names following marriage. Sample loss reduces efficiency of estimation as sample error increases with reduced sample size. But the more serious problem is that of bias: those individuals who remain in the survey may no longer constitute the representative sample with which the study began. As in longitudinal studies everywhere, British cohort studies

tend to lose disproportionately the less educated and more economically disadvantaged cohort members. Also, more women than men remain in the studies over time.

Frequency and mode of data collection. The greater the frequency of contact with cohort members the less reliance needs to be placed on retrospection in building the longitudinal record e.g., job and housing histories in between surveys. This has to be set against the burden placed on respondents of too frequent participation in surveys. The mode of data collection moves between CAPI (computer aided personal interview face-to-face or by telephone), postal questionnaire or moves increasingly to online contact. Each will have its own costs and benefits. The decisions involved in the early years of the studies are, however, often inconsistent. This raisies questions about the equivalence of the data for the same variables within and across studies.

These challenges can be met in a number of ways. Missing variables may require recourse to original sources of the data like schools or hospitals (if they are still in existence and if permission can be gained) for additional information or recoding of raw data to conform to more up-to-date protocols. Administrative records are another potential source, but raise questions of accuracy as well as access issues concerned with data protection and disclosure. The incidence of missing cases through non-contact or refusal can be reduced by continuous liaison with cohort members to retain their interest in the project through such 'panel maintenance' devices as a newsletter, annual birthday card, and a website, with telephone follow-up to update address records. Careful management of contacting procedures and the use of highly trained and experienced interviewers can also improve retention. The other approach is statistical: Capitalizing on the statistical record to restore earlier distributions of key variables through weighting the data, or as a basis for multiple imputation of missing values for particular variables or for the follow-up survey as a whole.

To overcome biases brought about by design features concerned with frequency and mode of data collection requires a combination of statistical validation work prior to analysis, drawing on the fund of advice available through the study team and the expert analysts to whom they relate, and the 'grey literature' of the study that only the team have access to. This makes the point that longitudinal resource management is a vital but rather under-recognized aspect of quality assurance. A continuing survey team and good leadership are essential to ensure that the necessary tasks are undertaken to perfection at the different stages of the survey cycle: design, data collection, data editing and cleaning, data release and advice on use. This is apart from the analysis that the team themselves will clearly want to undertake. The whole issue is one of balancing one set of needs against another. Research benefits not only those who are doing it, but all analysts thereafter. In many respects a cohort study is an evolving scientific instrument that while maintaining critical continuities with the past requires continual testing, evaluating and reinventing to gain the maximum returns from the substantial investment that goes into it.

Conclusion

Large-scale longitudinal surveys are increasingly recognized as making a major contribution to understanding human life course processes and their outcomes within and across generations in different settings. The extension of such surveys across time (sequential cohorts), across countries (comparative cohorts), and intergenerationally is a natural development and a key reflector of their value. The Millennium year gave a boost in many countries to building from the foundations, and replicating the findings, of long-standing cohort studies of the British kind in a range of national settings. Unlike the early studies, however, the newest have a huge fund of experience to draw on. Such collaboration offers scientific benefits in numerous directions not only for the developmental scientist who designs and implements the longitudinal study but for the wider set of secondary users who will increasingly be gaining access to it. The whole development is exemplified by the launch of the international journal, Longitudinal and Life Course Studies, www.journal.longviewuk.com and the Society for Longitudinal and Life Course Studies, www.longstudies. longviewuk.com

The Centre for Longitudinal Studies exemplifies both the history and the current practice of the cohort study enterprise. It will continue to be a standard bearer for what is now an ever-growing international program.

Notes

1. From 1985 to 1998 located in London's City University under the name of Social Statistics Research Unit.
2. Currently the Economic and Social Research Council and a consortium of government departments sponsor the studies.

References

Alwin, D. F. (1995). Taking Time Seriously: Studying Social Change, Social Structure, and Human Lives, in Moen, P., Elder, G. H. and Lüscher, K. (Eds.). *Linking Lives and Contexts: Perspectives on the Ecology of Human Development*. Washington, D.C.: American Psychological Association.

Blossfeld, H-P. (2009). Comparative Life Course Research: A Cross-National and Longitudinal Perspective. In Elder, G. H. and Giele, J. Z. (Eds). *The Craft of Life Course Research*, pp. 28-306. London: Guilford Press.

Bronfenbrenner, U. (1977). *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge: Harvard University Press.

Bynner, J. (2005). Longitudinal Cohort Designs. *Encyclopaedia of Social Measurement*. San Diego: Elsevier.

Bynner, J., & Joshi, H. (2007). Building the Evidence Base from Longitudinal Data: the Aims, Content and Achievements of the British Birth Cohort Studies. *Innovation: The European Journal of Social Research, 20* (2), 159-180.

Elder, G. H., & Giele, J. Z. (Eds.) (2009). *The Craft of Life Course Research*. London: Guilford Press.

Ferri, E., Bynner, J., & Wadsworth, M. (Eds.) (2003). *Changing Britain, changing lives: three generations at the end of the century*. London: Institute of Education.

Giele, J. Z., & Elder, G. H. (1998). *Methods of Life Course Research*. London: SAGE Publications.

Heinz, W. R., Huinink, J. and Weyman, A. (Eds.) (2009). *The Life Course Reader: individuals and Societies across time*. Frankfurt /Main: Campus Verlag.

Janson, C-G. (2000). *Seven Swedish Longitudinal Studies in the Behavioural Sciences*. Stockholm: Swedish Council for Planning and Coordination of Research.

Lerner, R.M. (Ed.), (1998) *Handbook of Child Psychology, Volume 1: Theoretical Models of Human Development*. New York: Wiley.

Jöreskog, K. G., & Sörbom, D. (1979). *Advances in Factor Analysis and Structural Equation Models*. Boston: Abt Associates.

Lynn, P. (Ed.) (2009). Methods for Longitudinal Surveys in Lynn, P. (Ed.). *Methodology of Longitudinal Surveys*, 1-18. Chichester: John Wiley and Sons.

Magnusson, D., Bergman, L. R., Rudinger, G., & Törestad, B. (Eds.) (1991*). Problems and Methods in Longitudinal Research*. Cambridge: Cambridge University Press.

Martin, J., Bynner, J., Kalton, G., Boyle, P., Goldstein, H., Gayle, V., Parsons, S., & Piesse, A. (2006). *Strategic Review of Panel and Cohort Studies: Report to the Research Resources Board of the Economic and Social Research Council*. http://www.longviewuk.com/pages/documents/ESRCstrategicreviewwebsite2bedit27.03.06a.pdf

Phelps, E., Furstenburg, F.F. and Colby, A. (2002). Looking at Lives: *American Longitudinal Studies of the Twentieth Century* (Sage Foundation, New York).

Piesse, A., Judkins, D., and Kalton, G. (2009). Using Longitudinal Surveys to Evaluate Interventions, in Lynn, P. (Ed.) *Methodology of Longitudinal Surveys*, 1-18, Chichester: John Wiley and Sons.

Robins, L., & Rutter, M. (1990). *Straight and Devious Pathways from Childhood to Adulthood*. Cambridge: Cambridge University Press.

Schaie, K. W. (2000). The impact of longitudinal studies on understanding development from young adulthood to old age. *International Journal of Behavioral Development, 24*, 257-266.

Schoon, I. (2006). *Risk and Resilience: Adaptations in changing times*. Cambridge: Cambridge University Press.

Shadish, K. W., Cook, T. D., & Campbell, D. (2001). *Experiments and Quasi-Experiments for Generalised Causal Interference*. Boston, MA: Houghton Mifflin.

Shanahan, M. J.,& Boardman, J. D. (2009). Genetics and Behaviour in the Life Course: A Promising Frontier. In Elder, G. H. & Giele, J. Z. (Eds.) *The Craft of Life Course Research* (pp. 215-235). London: Guilford Press.

The Dunedin Multidisciplinary Health and Development Study: Tips and traps from a 40-year longitudinal study

Richie Poulton
University of Otago, Dunedin, New Zealand
E-mail: richie.poulton@dmhdru.otago.ac.nz

and

Terrie E. Moffitt
Duke University, Durham, NC, USA, and King's College, London, UK
E-mail: terrie.moffitt@duke.edu

The Dunedin Multidisciplinary Health and Development Study (hereafter the Dunedin Study) is a longitudinal investigation of health and behavior in a complete birth cohort. The cohort of 1037 children (52% male) was constituted at age 3 when pediatric investigators followed up and enrolled 91% of births between April 1, 1972 and March 31, 1973 in Dunedin, New Zealand. At enrollment cohort families represented the full range of socioeconomic status in New Zealand, as compared to the New Zealand census. Cohort members are primarily white; 7.5% self-identify as being Maori which matches the ethnic distribution of the South Island (Silva & Stanton, 1996; Moffitt, Caspi, Rutter, & Silva, 2001). Day-long assessments (called phases) have been conducted at ages 3, 5, 7, 9, 11, 13, 15, 18, 21, 26 and most recently at age 32 years (in 2004-2025). Our next assessment, at age 38, begins on June 1, 2010.

The Study is known for its low attrition rate. As can be seen in the table below, all but one of the 12 assessments conducted thus far have enjoyed participation rates well above 90%.

The Dunedin Study is led by a Director, an Associate Director, and a team of 'Theme Leaders' who share responsibility for the following six broad research themes: mental health and cognition, cardiovascular health, respiratory health, sexual and reproductive health, oral health, and psychosocial functioning. Another theme explores how the multidisciplinary database can address issues of concern to Maori.

Research Strategy

The Dunedin Study uses a prospective-longitudinal, correlational design. This enables several different types of studies including: (1) prediction studies of the childhood correlates of later health and behavioral outcomes; (2) developmental studies of onset, course, continuity and change in health and behavior; (3) epidemiological studies of the prevalence and incidence of health and behavior problems, and associations among problem types; and (4) methodological studies of reliability and sampling biases. Thus the basic strategy of the Dunedin Study involves testing for causal relations within this correlational design. We adopt a stepwise approach. First we document that a basic association between two constructs exists, striving for gold-standard (for the period) measurement. Second, we document temporal sequence; whether the putative causal variable precedes the outcome variable. Third, we look for a dose-response contingency between the putative cause and outcome. Fourth, we attempt to rule out as many rival causal explanations as possible by introducing control variables into the analyses from our extensive database. Fifth, we test whether putative causal experiences are associated with intra-individual change. Sixth, we try to establish specificity by substituting alternative dependent, then independent variables. Finally, we posit a plausible explanatory process and test this using mediation analyses.

While the Dunedin Study is correlational (precluding definitive causal statements), there are several features of the Study that help strengthen casual inference. These include: (i) a low attrition rate enhances the accuracy of our effect size estimates; (ii) use of supplementary data sources (self-reports, informant reports, records, tests) permits replication checks across sources, and allows

Table 1.

		Design of Dunedin Multidisciplinary Health and Development Study			
Year	Phase	Birthday	Number Eligible	Number Seen	% Seen
1972-73	I	birth record			
1975-76	III	3	1139*	1037	91
1977-78	V	5	1037	991	96
1979-80	VII	7	1035[#]	954	92
1981-82	IX	9	1035[#]	955	92
1983-84	XI	11	1033[#]	925	90
1985-86	XIII	13	1031[#]	850	82
1987-88	XV	15	1029[#]	976	95
1990-91	XVIII	18	1027[#]	993	97
1993-94	XXI	21	1020[#]	992	97
1998-99	XVI	26	1019[#]	980	96
2003-05	XXXI	32	1015[#]	972	96
2010-12	XXXVII	38	1010[#]		

* Number resident in Otago (province); [#] Surviving Study members.

construction of multiple-indicator measurement models; (iii) our long history of proven confidentiality and non-intervention enhances frank reports on delicate topics; and (iv) face-to-face interviewing conducted at the Unit insures privacy, avoids falsification, reduces nay-saying to hasten the assessment process, and facilitates breadth of data collection (e.g., biological data can be collected under controlled conditions, using state-of-the-art technology).

We prefer face-to-face interviews because this helps avoid problems with literacy and comprehension (between 5-10% of a general population sample will have at least mild difficulties). We present the assessment modules in a counter-balanced fashion to avoid order effects, and to maintain participants' interest during the 8-hour assessment day (e.g., a cognitively taxing assessment is followed by a physiological assessment with low cognitive demand). We hire appropriately qualified interviewers (e.g., a clinical psychologist for the mental health assessment, a dentist for the oral health assessment, a nurse for the cardiovascular assessment) and provide extensive training on the research protocols. Finally, interviewers are kept blind to the Study member's performance in other sessions, and past waves of the study. This helps reduce the potential for interview bias and inflated associations from shared method variance. Importantly, we have published evidence that half-a-lifetime of research participation has not improved Study members' mental or physical health as compared to same-aged participants in the New Zealand National Health and Nutrition Survey (e.g. BMI, smoking, visits to the Doctor (Poulton et al., 2006)).

Research Findings

To date there have been 1100 publications and reports from the Dunedin Study, with approximately two thirds of these appearing in peer-reviewed journals. Space limitations preclude a lengthy discussion of findings. Rather, we highlight several types of research which take advantage of the design features described above. As a multidisciplinary study that nests laboratory-type assessments within an epidemiological sampling frame, the study is well positioned to marry biology with life experience. A recent series of studies modeling nature-nurture interplay in the behavioral domain (Caspi et al., 2002, 2003, 2005, 2007) illustrate the value of having a biorepository with DNA, as well as high quality repeated measures of environmental exposures from birth onwards. These studies have helped stimulate interest among behavioral scientists in GxE. One of the key messages from this work is that genes by themselves tell us very little; it is the combination of genes and environments that matters. The implication is that attempts to modify environments remain a very sensible intervention strategy for improving health and development.

Another approach deliberately emphasises cross-disciplinary studies. This is made easier because the Dunedin Study was founded as a multidisciplinary research project at a time when that was not the vogue. Subsequently funding agencies have become keen to invest in multi- or cross-disciplinary research because multiple physiological systems (e.g. CNS, gastrointestinal tract, the immune system) interact with each other, and with psychosocial influences 'outside the skin'. Thomson et al. (2008) examined the relation between lifetime cannabis use and gum disease at age 32, controlling for lifetime tobacco exposure (the major risk factor for periodontal disease), as well as several other covariates. With good cannabis exposure data collected prospectively on almost all our participants we showed that (i) smoking cannabis was an independent risk factor for developing gum disease even among non-smokers; and (ii) that the more cannabis a person used, the greater the risk for gum disease (i.e., a dose-response relation). It is interesting to note that our rates of reported cannabis use and dependence tend to be higher than those typically reported in cross-sectional national surveys (Poulton, Moffitt, Harrington, Milne, & Caspi, 2001; Moffitt et al., 2009), and that illicit substance use occurs at twice the rate among the 20% of participants hardest to recruit and most

often lost to follow-up but retained in the Dunedin Study (Poulton et al., in preparation).

Another study sought to understand how childhood maltreatment might get "under the skin' to cause physical disease (Danese, Pariante, Caspi, Taylor, &Poulton, 2007). In the first study in humans spanning childhood to adulthood, we showed that child maltreatment significantly increased the risk for clinical levels of systemic inflammation (a ubiquitous risk factor for chronic age-related diseases). This association remained after adjustment for co-occurring risks during childhood (low birthweight, SES, and IQ); stress in adulthood (low SES, depression, high perceived stress); and adult health and health behavior (high cardiovascular risk status, smoking, low physical activity, and poor diet). Establishing the robustness of this association between childhood maltreatment and adulthood inflammation was only possible because of the developmental database containing many potential confounding factors.

Sample Retention Strategies

Good systems for maintaining contact and locating study members are essential. At each assessment we ask the study members for contact details of people who know them well, and are likely to know their movements over time. We have requested this type of information at multiple phases and have consequently built up a substantive 'contacts' database for each study member. This information is augmented by change-of-address forms returned to us following the mail-out of a regular newsletter.

Good recruitment staff: Location data is used by a sample tracer who is employed at each assessment to track, locate and recruit Study members. A travel co-ordinator organizes travel plans for the Study member (and other family members if requested). One quarter of our sample lived outside of New Zealand at age 32, thus this can be a demanding role, and one we have supported via travel industry training for the staff member concerned.

Remove barriers to participation: The most obvious barrier is financial, thus we provide transportation to and from the Dunedin Research Unit, accommodation, and meals. We also provide a letter to the Study member, addressed to the member's employer, explaining the value of the Study and seeking the employer's co-operation in allowing our Study member to attend the assessment without loss of income. In rare circumstances (approximately 2%) people who wish to participate but cannot travel to Dunedin (e.g., work commitments, incarceration) are visited by two interviewers in the field.

The assessment needs to be engaging for study members and well-balanced across the day to avoid fatigue or boredom. The assessment staff must be well-trained and professional at all times. If only one grumpy staff member deals poorly with a study member, this risks alienating that study member, and potentially their friends who are also in the study, forever. To ensure staff members remain fresh, week-long breaks are programmed into the assessment phase, and social events occur regularly to keep staff morale high.

Research outputs are noticed by Study members. Thus, they gain a greater appreciation of the value of their contributions, and they feel respected because the researchers have worked hard and produced good work.

Good resources are critical. Although good funding is tremendously helpful, it is also worth noting that the Dunedin Study was run on the 'smell of an oily rag' in the first 10 years. Volunteers from the local community made the difference in those early years. An equally important resource is the researchers' commitment. Longitudinal research takes a long time. Researcher must be committed and patient to fully enjoy the fruits of this unusual labour. Also, Study members appreciate continuity in key staff positions, strengthening their ties to the study.

All of these factors are necessary, but not sufficient for high retention. In other words, it is not so much that you *do* these things (the list above would not surprise anyone), rather, it is *how* you do these things. This largely comes down to a combination of professionalism, courtesy and persistence (x3). Paying close attention to the researcher-participant relationship - in the most simple terms 'treat people like you would like to be treated' - will go a long way to earning sample members' loyalty.

Challenges for Launching New Longitudinal Research

Funding agencies grimace when they see applications for large longitudinal studies that will run for decades! These studies are costly, in terms of both dollars and time. Successfully launching a new cohort study will probably require identifying a novel niche, a need that no one else is addressing. Diversifying your funding sources should be a priority, albeit with care to avoid having too many contributors with unrealistically high expectations of return and/or rights to influence the research agenda. Data linkage can add great value at relatively low cost, so all opportunities must be explored.

Demonstrating utility is increasingly expected by public-good research funders. Being able to provide concrete evidence of impact on policy and/or practice enhances the likelihood of continued support, plus evidence, if impact is good, to feed back to study members. This suggests that longitudinal researchers need to engage more directly with policy makers. This can be challenging insomuch as policy-makers tend to respond to different contingencies than do researchers, and the demands of the policy world (e.g., short-term results) are often antithetical to carefully conducted, rigorous research.

Links to other studies: One of the most critical needs in behavioral and health sciences is replication of findings across different nations and population groups. We must document the robustness of each finding, and we must also delimit when and where it does not apply. Thus, communication among research teams and harmonization of data collection and analyses is necessary if we are to translate findings into improved health care. Here we have described the Dunedin Study, as an invitation to contact with other similar studies.

References

Caspi, A., McClay, J., Moffitt, T. E., Mill, J. S., Martin, J., Craig, I., Taylor, A., & Poulton, R. (2002). Role of

genotype in the cycle of violence in maltreated children. *Science, 297*, 851-854.

Caspi, A., Moffitt, T. E., Cannon, M., McClay, J., Murray, R. M., Harrington, H. L., Taylor, A., Arseneault, L., Williams, B. S., Braithwaite, A., Poulton, R., & Craig, I. (2005). Moderation of the effect of adolescent-onset cannabis use on adult psychosis by a functional polymorphism in the COMT gene: Longitudinal evidence of a gene x environment interaction. *Biological Psychiatry, 57*, 1117-1127.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I., Harrington, H. L., McClay, J., Mill, J. S, Martin, J., Braithwaite, A., & Poulton, R. (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science, 301*, 386-389.

Caspi, A., Williams, B. S., Kim-Cohen, J., Craig, I., Milne, B. J., Poulton, R., Schalkwyk, L. C., Taylor, A., Werts, H., & Moffitt, T. E. (2007). Moderation of breastfeeding effects on cognitive development by genetic variation in fatty acid metabolism. *PNAS (Proceedings of the National Academy of Sciences of the USA), 104*(47), 18860-18865.

Danese, A., Pariante, C. M, Caspi, A., Taylor, A. & Poulton, R. (2007). Childhood maltreatment predicts adult inflammation in a life-course study. *PNAS (Proceedings of the National Academy of Sciences of the USA), 104*, 1319-1324.

Moffitt, T. E., Caspi, A., Rutter, M., and Silva, P. A. (2001). *Sex Differences in Antisocial Behaviour: Conduct Disorder, Delinquency, and Violence in the Dunedin Longitudinal Study*. Cambridge: Cambridge University Press.

Moffitt, T. E., Caspi, A., Taylor, A., Kokaua, J., Milne, B. J., Polanczyk, & Poulton, R. (2009). How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. Psychological Medicine. *doi:10.1017/S0033291709991036.*

Poulton, R., Ely, M., Baxter, J., Caspi, A., Dickson, N., Hancox, R.J., Moffitt, T. E., Paul, C., Sears, M., & Thomson, W.M. (in preparation). Quantifying the effects of non-random attrition in a 32-year longitudinal birth cohort study.

Poulton, R., Hancox, R. J., Milne, B. J., Baxter, J., Scott, K., & Wilson, N. (2006). The Dunedin Multidisciplinary Health and Development Study: are its findings consistent with the overall New Zealand population? *New Zealand Medical Journal*, 119, U2002. *Available online 16751825.*

Poulton, R., Moffitt, T. E., Harrington, H.L., Milne, B.J., & Caspi, A. (2001). Persistence and perceived consequences of cannabis use and dependence among young adults: implications for policy. *New Zealand Medical Journal*, 114, 544-547.

Silva, P. A., & Stanton, W. R. (Eds.) (1996). *From Child to Adult: The Dunedin Multidisciplinary Health and Development Study*. Auckland: Oxford University Press.

Thomson, W. M., Poulton, R., Broadbent, J. M., Moffitt, T. E., Caspi, A., Beck, J. D., Welch, D., & Hancox, R. J. (2008). Cannabis smoking and periodontal disease among young adults. *JAMA*, 299(5), 525-531.

The Montreal Longitudinal and Experimental Study: Tracing the Developmental Trajectories of Behavior Problems and Assessing their Prevention

Richard E. Tremblay
International Network for Early Childhood Health Development, University College Dublin, Ireland; University of Montreal, Canada; INSERM, France
E-mail: trembler@GRIP.Umontreal.ca

A major focus of the Montreal Longitudinal and Experimental Study (MLES), which began in 1984, has been the behavior of disruptive kindergarten boys with a view to preventing delinquent behavior later on. With its 1,161 kindergarten boys from low SES environments, the MLES appears, at the beginning of the 21st century, to be the largest longitudinal and experimental study specifically designed to understand the development and prevention of delinquency with subjects first assessed before Grade 1.

The main instruments used in the MLES to assess behavior problems are mother and teacher ratings and self-reported delinquency. Behavior ratings were also obtained annually from mothers, and from classroom peers at ages 10, 11 and 12. A structured psychiatric interview was conducted with the boys and their mothers when the boys were 15 years of age. Annual questionnaires completed by mothers provided information on family background, life events, parenting behavior, domestic relationships, and social support. Annual interviews with the boys provided information on a variety of dimensions including personality, life events, perceptions of parenting, domestic relationships, friendships, attitudes toward school and the law, sleep, and leisure activities. Direct observations of social interactions were made at school, at home and in laboratory situations with sub-samples between ages 7 and 15. Psychophysiological and neuropsychological tests, as well as different biological assessments, were also made on sub-samples between ages 7 and 20.

Teachers and mothers rated the boys' behavior in the spring, from age 6, at the end of the kindergarten year, to age 15. The six scales derived from teacher and mother ratings were: Physical aggression (fights with other children; kicks, bites, or hits other children; bullies or intimidates other children); Opposition (doesn't share materials; irritable; disobedient; blames others; inconsiderate); Covert conduct problems (destroys property, lies, steals, truants); Anxiety (tends to be fearful or afraid of new things or new situations; cries easily; appears miserable, unhappy, fearful, or distressed); Inattention (has poor concentration or short attention span; inattentive); Hyperactivity (restless, runs about or jumps up and down, does not keep still; squirmy, fidgety); and Prosocial behavior (comforts upset child; helps sick child; helps hurt child;

Table 2. Overview of the MLES Research Design

Year	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Age	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Total Sample	TPT			PT	TPQ BQ PSS	TPQ BQ PSS	TPQ BQ PSS	TPQ BQ	TPQ BQ	TPQ BQ	BQ	BQ					Cr	BQ		Cr	Adn	BQ	
Group A		PLBLO BHO BSO TPQ PSS		PLBLO BHO BSO TPQ PSS PSL		PLBLO BHO BSO PSL		PL BLO	PL BLO	PL BLO													
Group B		TPQ PSS	TPQ EX TPQ PSS	EX TPQ PSS																			
Group C		PQ	PQ PQ	TPQ PSS																			
Group D								BLO PSL	BLO PSL	BLO PSL	BLO BLO			BLO PSL	BLO BLO	BLO			BLO				Pet

Group A: Observation group (N = 84); Group B: Intervention group (N = 84); Group C: Control group (N = 42); Group D: Bio-psycho-social study group (N = *). This group included subjects from Group A, B, C and from the total sample

* The N varied depending on the specific problems studied. N = 203 from 1991 to 1996; N = 336, from 1997 to 1999; N = 40, 2002-2003; N = 26, 2006

T = Teacher questionnaire; PT = Parent telephone; PQ = Parent questionnaire; PL = Parent laboratory; BLO = Boys laboratory observation; BQ = Boys questionnaire; BHO = Boys home observations; BSO = Boys school observations; PSS = Peers sociometric data at school; PSL = Peers sociometric data in laboratory; EX = Experimental intervention; Cr = Criminal records; Pet = Brain PET scan; Adn = ADN collection

praises other; helps with task difficulty; helps clear up mess; shows sympathy; invites bystander; stops quarrels; helps pick up objects).

Self-reported delinquency was assessed every year from ages 10 to 17 and around age 20. The boys were seen in small groups at their schools between March and May; they answered 27 self-reported delinquency items which were distributed in a questionnaire pertaining to school, family, friends, and leisure activities. The items were: steal from school; steal from store; steal from home; keep object worth less than $10; steal bicycle; sell stolen goods; keep object worth between $10 and $100; steal objects worth more than $100; breaking and entering; enter without paying; trespassing; take drugs; take alcohol; get drunk; destroy school material; destroy other material; vandalism at school; destroy objects at home; vandalize car; set a fire; strong-arm; gang fights; use weapon in a fight; fist fight; beat up someone; carry a weapon; throw objects at persons. The responses to all 27 items ranged from 1 (never) to 4 (very often).

Treatment

During the planning phase of the study in 1982-83, the Oregon Social Learning Center's work with parent (Patterson, 1982; Patterson, Reid, Jones, & Conger, 1975) appeared to us to be one of the most innovative training programs for parents with aggressive children. It was decided that techniques developed at the Oregon Social Learning Center would be a major component of the treatment program. Accordingly, the program coordinator was trained at the Center. In brief, the procedure involved: (1) giving parents a reading program; (2) training parents to monitor their children's behavior; (3) training parents to give positive reinforcement for prosocial behavior; (4) training parents to manage family crises; and (5) helping parents to generalize what they have learned. The parent training component was supplemented by eliciting cooperation from the teacher. Work with parents and teachers was carried out by two university-trained child-care workers, one psychologist, and one social worker, all working full-time. Each of these professionals had a caseload of 12 families. The team was coordinated by a fifth professional who worked on the project half-time. Work with the parents was planned to last for 2 school years with one session every 2 weeks. However, the professionals were free to decide that a given family needed more or fewer sessions at any given time. The maximum number of sessions given to any family was 46 and the mean number of sessions over the 2 years was 17.4, counting families that refused to continue.

For the school based social skills training component of our interventions, two types of training were given to the disruptive boys within a small group of prosocial peers in school. During the first year a prosocial skills program was devised based on other programs (Cartledge & Milburn, 1980; Michelson, Sugai, Wood, & Kazdin, 1983; Schneider & Byrne, 1987). Nine sessions were given on themes such as "How to make contact," "How to help," "How to ask 'why'," and "How to invite someone in a group." Coaching, peer modeling, role playing, and reinforcement contingencies were used during these sessions. During the second year the program aimed at self-control.

Using material from previous studies (Camp, Blom, Hebert, & Van Doorminck, 1977; Goldstein, Sprafkin, Gershaw, & Klein, 1980; Kettlewell & Kausch, 1983; Meichenbaum, 1977), 10 sessions were developed on themes such as "Look and listen," "Following rules," 'What to do when I am angry," "What to do when they do not want me to play with them," and "How to react to teasing." Coaching, peer modeling, self-instructions, behavior rehearsal, and reinforcement contingencies were also used during these sessions. This treatment too was offered by the professionals who provided parental training, although different workers assisted the parents and the child.

We also experimented with stimulating fantasy play and teaching the boys to be critical of television. However, because of lack of funds, only half the treated group ($n = 25$) received the fantasy play training and one fifth ($n = 9$) received the television intervention.

Statistical Analyses

The MLES was used to pioneer the use of group-based semi-parametric statistical analyses to study developmental trajectories of behavior (Nagin & Tremblay, 1999), identify their risk factors (Nagin & Tremblay, 2001) and compare the effectiveness of a preventive intervention on developmental trajectories of behavior (Lacourse et al., 2002). Developmental trajectories graphically describe statistically different courses of development over a given developmental period based on repeated assessments of intermediate behaviour problem states (Muthén, 2004; Nagin, 2005: Tremblay, 2010). Developmental trajectory analyses with random samples of the population provide information on age-normative and atypical development. Studies of developmental trajectories usually take into account yearly or biennial assessments within and between global age periods, such as early childhood, middle childhood, adolescence and adulthood (e.g., Barker et al., 2010; Côté, Vaillancourt, LeBlanc, Nagin, & Tremblay, 2006; Nagin & Tremblay, 1999). It is important to note that his statistical approach to the study of behavior development is very different from the analyses used to create the so called "developmental taxonomies" generally used by clinicians dealing with disruptive and delinquent individuals: 'early–late onset' Conduct Disorder and 'life-course-persistent, adolescent-limited' antisocial behavior. These taxonomies were created before the advent of statistically-based developmental trajectory analyses (Tremblay, 2010). They are the product of ad hoc dichotomies of assessments at two points in time: before vs. after age 10 for 'early-late onset' Conduct Disorder (American Psychiatric Association, 1994) and childhood vs. adolescence in the case of 'life-course-persistent, adolescent-limited' antisocial behaviour (Moffitt, Caspi, Dickson, Silva, & Stanton, 1996).

References

Barker, E.D., Vitaro, F., Lacourse, E., Fontaine, N. M. G., Carbonneau, R., & Tremblay, R. E. (2010). Testing the developmental distinctiveness of male proactive and reactive aggression with a nested longitudinal experimental intervention. Aggressive Behavior. EPUB.

Camp, B. W., Blom, G. E., Hebert, F., & Van Doorminck, W. J. (1977). Think Aloud: A program for developing self-control in young aggressive boys. *Journal of Abnormal Child Psychology, 5*, 157-169.

Cartledge, G., & Milburn, J. F. (1980). *Teaching social skills to children. Innovative approaches* (pp. 92-109). New York, NY: Pergamon Press.

Côté, S.M., Vaillancourt, T., LeBlanc, J.C., Nagin, D.S., & Tremblay, R.E. (2006). The development of physical aggression from toddlerhood to pre-adolescence: A nation-wide longitudinal study of Canadian children. Journal of Abnormal Child Psychology, *34*, 71–85.

Goldstein, A. P., Sprafkin, R. P., Gershaw, N. J., & Klein, P. (1980). The adolescent: Social skills training through structured learning. In G. Cartledge & J. F. Milburn (Eds.), *Teaching social skills to children: Innovative approaches* (pp. 249-277). New York, NY: Pergamon Press.

Kettlewell, P. W., & Kaush, D. F. (1983). The generalization of the effects of a cognitive behavioral treatment program for aggressive children. *Journal of Abnormal Child Psychology, 11*, 101-114.

Lacourse, E., Côté, S., Nagin, D.S., Vitaro, F., Brendgen, M., & Tremblay, R.E. (2002). A longitudinal-experimental approach to testing theories of antisocial behavior development. Development and Psychopathology, 14, 909–924.

Meichenbaum, D. (1977). *Cognitive-behavior modification: An integrative approach*. New York, NY: Plenum Press.

Michelson, L., Sugai, D., Wood, R., & Kazdin, A. E. (1983). *Social skills assessment and training with children*. New York, NY: Plenum Press.

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (ed.), Handbook of quantitative methodology for the social sciences (pp. 345–368). Newbury Park, CA: SAGE.

Nagin, D., & Tremblay, R.E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. Child Development, *70*, 1181–1196.

Nagin, D., & Tremblay, R.E. (2001). Parental and Early Childhood Predictors of Persistent Physical Aggression in Boys From Kindergarten to High School. Archives of General Psychiatry, *58*, 389-394.

Nagin, D.S. (2005). Group-based modeling of development over the life course. Cambridge, MA: Harvard University Press.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia Press.

Schneider, B. H., & Byrne, B. M. (1987). Individualizing social skills training for behavior-disordered children. *Journal of Consulting and Clinical Psychology, 55*, 444-455.

Tremblay, R. E. (2008). Understanding development and prevention of chronic physical aggression: Towards experimental epigenetic studies. *Philosophical Transactions of the Royal Society of London, 363*(1503), 2613-2622.

The Seattle Longitudinal Study of Adult Cognitive Development

K. Warner Schaie
University of Washington, USA
E-mail: Schaie@u.washington.edu

and

Sherry L. Willis
University of Washington, USA
E-mail: Schaie@u.washington.edu

The Seattle Longitudinal Study (SLS; Hertzog, 2010; Schaie, 1996a, b, 2000, 2005a) began as Schaie's doctoral dissertation at the University of Washington (Seattle, WA) in 1956. In an effort to resolve the discrepancies between cross-sectional and longitudinal findings in the study of adult intellectual development, Schaie designed a follow-up study, put into the field in 1963, that provided some answers, but also raised enough methodological and substantive questions to demand a continuing program of studies that is still in progress and has included eight major and several collateral data collections as well as three cognitive training studies. An important aspect of the study has been the investigation of cohort differences and their implications for the study of adult cognition. In the more recent phases of the study, a number of contextual, health, and personality variables have been identified that offer explanations for differential change and that provide a basis for possible interventions. Within the context of our monitoring of individual change, the authors therefore designed cognitive interventions that have been successful in reversing carefully documented declines in ability performance and in improving the cognitive functions of older persons who have remained stable.

Many reasonable arguments have been made for the proposition that motivational and other personality variables might have greater potency in predicting adjustment and competence in midlife than does intelligence, but the empirical evidence for this proposition is less than convincing. Questions such as who should be retired for cause (read incompetence) in the absence of mandatory retirement at relatively early ages; whether there is sufficient remaining competence for independent living; or whether persons can continue to conserve and dispose of their property all involve the assessment of intellectual functioning (see Schaie, 1988a). We need to examine at what age developmental peaks do occur and assess generational differences as well as within-generation age changes. Most important, we must determine the reasons why some individuals show intellectual decline in early adulthood, while others maintain or increase their level of intellectual functioning well into advanced old age.

Initial studies in 1949 (PMA 11–14; L. L. Thurstone & T. G. Thurstone, 1949) and 1953 (Schaie, Rosenthal, & Perlman, 1953), explored whether the factorial independence of the five abilities measured in the most advanced form of the PMA test would be retained in adulthood. In a 1963 follow-up study, in addition to tracking down and
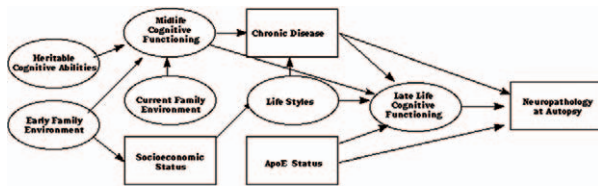
retesting as many of the individuals studied in 1956 as possible, we decided to draw a new random sample from the original population frame to provide the necessary controls for examining retest effects and began addressing the possibility that socio-cultural variation affects intellectual performance. The latter concern was stimulated by the thoughtful admonitions previously voiced by Raymond Kuhlen (1963). Our new sample therefore was extended over the original age range (22–70 years) plus an additional 7-year interval to match the age range now attained by the original sample.

The second cross-sectional study essentially replicated the findings of the base study. The short-term longitudinal study, however, disclosed substantially different information about peak levels and rate of decline. Publication of findings was therefore delayed until a theoretical model could be built that accounted for the discrepancy between the longitudinal and cross-sectional data (Schaie, 1965, 1977, 2007). These analyses suggested that comparisons of age group means needed to be conducted for the repeatedly measured samples as well as for successive independent samples drawn from the same cohort. Results were reported that called attention to substantial cohort differences and that questioned the universality and significance of intellectual decrement with advancing age in community-dwelling persons (Nesselroade, Baltes, & Schaie, 1972; Schaie & Strother, 1968).

The results from the third data collection seemed rather definitive in replicating the short-term longitudinal findings, but discrepancies between findings in the repeated-measurement and independent-sampling studies suggested the need for a replication of the 14-year longitudinal sequences, and it further seemed useful to follow the original sample for as long as 21 years, so a fourth data collection was conducted in 1977, again retesting the previous samples and adding a new random sample, this time from an expanded population frame (Schaie & Hertzog, 1983), adding collateral questions of the effects of monetary incentives on participant characteristics (Gribbin & Schaie, 1976); an examination of the aging of tests; and the beginning of causal analyses of health and environmental factors on change or maintenance of adult intellectual performance (Gribbin, Schaie, & Parham, 1980; Hertzog, Schaie, & Gribbin, 1978).

Influences from Neighboring Sciences

The early introduction to the issues of cohort differences and secular trends led to serious questions as to what the meaning of these effects might be beyond their role as control variables or as bothersome design confounds. Increased attention was therefore given to the impact of social structures and microenvironments on cognitive change (Schaie, 1974). This work was influenced early on by the writing of

Matilda Riley (Riley, Johnson, & Foner, 1972) and later by the work of Carmi Schooler (1987; Schooler & Mulatu, 2005), as well as many other sociologists, anthropologists, and epidemiologists who have contributed to the Penn State social structure conference series (e. g., Bengtson, Schaie, & Burton, 1995; Kertzer & Schaie, 1989; Schaie & Abeles, 2008; Schaie & Elder, 2005; Schaie & Hendricks, 2000; Schaie, Krause, & Booth, 2004; Schaie & Schooler, 1998; Schaie & Uhlenberg, 2007).

The Study of Latent Constructs

Until the fourth (1977) cycle of the SLS, we followed the then-conventional wisdom of assessing each primary ability with the observable marker variable deemed to be the most reliable and valid measure of the latent construct to be estimated. With the widespread introduction of modern methods of confirmatory (restricted) factor analysis into the behavioral and social sciences, it became obvious that we needed to extend our concern with changes in level of intellectual functioning in adulthood to the assessment of structural relationships within the ability domain. This concern argued for collecting further data with a much expanded battery in which each ability would be multiply marked (Schaie, Dutta, & Willis, 1991; Schaie, Maitland, Willis, & Intrieri, 1998). We took this opportunity also to add marker tests for the ability domains of Perceptual Speed and Verbal Memory. The original battery of five mental ability tests was therefore expanded to 20 tests beginning with the 5th study wave.

Introducing Cognitive Interventions

The fifth (1984) SLS cycle also marked the assumption of a major role in the SLS by Sherry L. Willis, who brought to this project her skills in designing and implementing cognitive training paradigms. Thus, a major part of the fifth cycle included the implementation of a cognitive training study with our long-term participants aged 64 years or older, designed to assess whether cognitive training in the elderly serves to remediate cognitive decrement or increase levels of skill beyond those attained at earlier ages (Schaie & Willis, 1986; Willis, 1987, 1989, 1990; Willis & Nesselroade, 1990; Willis & Schaie, 1986b, 1988).

The database available through the fifth cycle also made it possible to update the normative data on age changes and cohort differences (Schaie, 1990a, 1990b, 1990c; Schaie & Willis, 1993) and to apply sequential analysis designs controlled for the effects of experimental mortality and practice (Cooney, Schaie, & Willis, 1988). Finally, this cycle saw the introduction of measures of practical intelligence (Willis & Schaie, 1986a), analyses of marital assortativity using data on married couples followed over as long as 21 years (Gruber-Baldini, Schaie, & Willis, 1995), and the application of event history methods to hazard analysis of cognitive change with age (Schaie, 1989).

Excursions into Behavior Genetics and Chronic Diseases

The sixth (1989–1991) SLS cycle included a set of four related studies. First, with the collaboration of Robert

Plomin, a noted developmental behavior geneticist, we took advantage of the longitudinal database to collect data to implement a study of cognitive family resemblance in adulthood. We did this by recruiting the participation of a large number of adult offspring and siblings of our longitudinal panel members (Schaie, Plomin, Willis, Gruber-Baldini, & Dutta, 1992; Schaie & Willis, 1995). Second, we abstracted the health histories of our panel members and conducted more detailed investigations of the relationship between health and maintenance of intellectual functioning (Bosworth & Schaie, 1997). Third, we conducted a 7-year follow-up of the cognitive training study and replicated the study with a more recent cohort of older persons (Willis, 2001). Fourth, we were able to conduct longitudinal analyses of cognitive ability structures and further update our normative database with the collection of a sixth (1991) wave using the standard approach of retesting and drawing a sixth new independent sample (Schaie, 1994, 1996a). Finally a seventh cycle conducted in 2005 resulted in the retesting of 1207 previous study participants.

Explorations of Neuropsychology, the APOE Gene, and the NEO

From 1997 to 1999, we conducted a follow-up of all previous participants who could be retrieved as well as a new seventh (1998) wave. In addition, we began to collect blood for genotyping on the *APOE* gene and administered a neuropsychological test battery to participants aged 60 years or older (cf. Schaie, 2005a, chapter 17; Schaie, Caskie, Revell, Willis, Kaszniak, & Teri, 2005). This battery continues to be administered in a 3-year follow-up cycle. We also conducted a 14-year follow-up for members of the family study and recruited additional eligible participants.

Between study waves, in 1993, we conducted a mail survey of health behaviors for those persons who had participated in the 1989 family study and the 1991 longitudinal and sixth-wave studies. This survey was used to develop a set of latent dimensions for the study of health behaviors (Maier, 1995; Maitland, 1997). Another mail survey, collecting data on the NEO scales (Costa & McCrae, 1992) was conducted in 2001 (cf. Schaie, 2005a, chapter12; Schaie, Willis, & Caskie, 2004). Finally, we recruited several hundred third-generation members (those with at least one parent and one grandparent in the study) to expand the family analyses. Currently in progress are structural MRI studies of a sub-sample of 165 individuals on whom we had previously obtained cognitive data in midlife.

Objectives of the Seattle Longitudinal Study

Throughout the history of the SLS, an effort now extending over 47 years, our focus has been on five major questions, which we have attempted to ask with greater clarity and increasingly more sophisticated methodologies at each successive stage of the study. These are elaborated next.

Does Intelligence Change Uniformly Through Adulthood, or Are There Different Life Course Ability Patterns?

Our studies have shown that there is no uniform pattern of age-related changes across all intellectual abilities,

and that studies of an overall Index of Intellectual Ability (IQ) therefore do not suffice to monitor age changes and age differences in intellectual functioning for either individuals or groups. Our data do lend some support to the notion that fluid abilities tend to decline earlier than crystallized abilities. There are, however, important ability-by-age, ability-by-gender, and ability-by-cohort interactions that complicate matters. Moreover, whereas fluid abilities begin to decline earlier, crystallized abilities appear to show steeper decrement once the late 70s are reached.

Although cohort-related differences in the rate and magnitude of age changes in intelligence remained fairly linear for cohorts who entered old age during the first three cycles of our study, these differences have since shown substantial shifts. For example, rates of decremental age change have abated somewhat, and at the same time modestly negative cohort trends are beginning to appear as we begin to study members of the baby boom generation. Also, patterns of socialization unique to a given gender role in a specific historical period may be a major determinant of the pattern of change in abilities. More fine-grained analyses suggested, moreover, that there may be substantial gender differences as well as differential changes for those who decline and those who remain sturdy when age changes are decomposed into accuracy and speed.

With multiple markers of abilities, we have conducted both cross-sectional and longitudinal analyses of the invariance of ability structure over a wide age range. In cross-sectional analyses, it is possible to demonstrate configural but not metric factor invariance across wide age/cohort ranges. In longitudinal analyses, metric invariance obtains within cohorts over most of adulthood, except for the youngest and oldest cohorts. Finally, we examined the relationship of everyday tasks to the framework of practical intelligence and perceptions of competence in everyday situations facing older persons.

At What Age Is There a Reliably Detectable Decrement in Ability, and What Is Its Magnitude?

We have generally shown that reliably replicable average age decrements in psychometric abilities do not occur prior to age 60, but that such reliable decrement can be found for all abilities by 74 years of age. Analyses from the most recent phases of the SLS, however, suggested that small but statistically significant average decrement can be found for some, but not all, cohorts beginning in the sixth decade. However, more detailed analyses of individual differences in intellectual change demonstrated that even at age 81, fewer than half of all observed individuals have shown reliable decremental change over the preceding 7 years. In addition, average decrement below age 60 amounts to less than 0.2 of a standard deviation; by 81 years of age, average decrement rises to approximately 1 population standard deviation for most variables.

As data from the SLS cover more cohorts and wider age ranges within individuals, they attain increasing importance in providing a normative base to determine at what ages declines reach practically significant levels of importance for public policy issues. Thus, our data have become relevant to issues such as mandatory retirement, age discrimination in employment, and prediction of proportions of the population that can be expected to live independently in the community. These bases will shift over time because we have demonstrated in the SLS that both level of

performance and rate of decline show significant age-by-cohort interactions.

What Are the Patterns of Generational Differences, and What Is Their Magnitude?

Results from the SLS have conclusively demonstrated the prevalence of substantial generational (cohort) differences in psychometric abilities. These cohort trends differ in magnitude and direction by ability and therefore cannot be determined from composite IQ indices. As a consequence of these findings, it was concluded that cross-sectional studies used to model age change would overestimate age changes prior to the 60s for those variables that show negative cohort gradients and underestimate age changes for those variables with positive cohort gradients.

Our studies of generational shifts in abilities have in the past been conducted with random samples from arbitrarily defined birth cohorts. As a supplement and an even more powerful demonstration, we have also conducted family studies that compared performance levels for individuals and their adult children. By following the family members longitudinally, we are also able to provide data on differential rates of aging across generations. In addition, we have also recruited siblings of our longitudinal participants to obtain data that allow extending the knowledge base in the developmental behavior genetics of cognition to the adult level by providing data on parent-offspring and sibling correlations in adulthood.

What Accounts for Individual Differences in Age-Related Change in Adulthood?

The most powerful and unique contribution of a longitudinal study of adult development arises from the fact that only longitudinal data permit the investigation of individual differences in antecedent variables that lead to early decrement for some persons and maintenance of high levels of functioning for others into very advanced age. A number of factors that account for these individual differences have been implicated; some of these have been amenable to experimental intervention. The variables that have been implicated in reducing risk of cognitive decline in old age have included (a) absence of cardiovascular and other chronic diseases; (b) a favorable environment mediated by high socioeconomic status; (c) involvement in a complex and intellectually stimulating environment; (d) flexible personality style at midlife; (e) high cognitive status of spouse; and (f) maintenance of high levels of perceptual processing speed.

Can Intellectual Decline with Increasing Age Be Reversed by Educational Intervention?

Because longitudinal studies permit tracking stability or decline on an individual level, it has also been feasible to carry out interventions designed to remediate known intellectual decline as well as to reduce cohort differences in individuals who have remained stable in their own performance over time but who have become disadvantaged when compared with younger peers. Findings from the cognitive training studies conducted with our longitudinal subjects (under the primary direction of Sherry L. Willis) suggested that observed decline in many community-dwelling older people might well be a function of disuse and is clearly reversible for many. Indeed, cognitive training resulted in approximately two-thirds of the experimental subjects showing significant improvement; and about 40% of those who had declined significantly over 14 years were returned to their predecline level. In addition, we were able to show that we did not simply "train to the test" but rather trained at the ability (latent construct) level, and that the training did not disturb the ability structure. We have now extended these studies to include both a 7-year and a 14-year follow-up that suggest the long-term advantage of cognitive interventions.

Methodological Issues

The dialectical process between data collection and model building that has been part of the SLS has made possible substantial methodological advances in the design and analysis of studies of human development and aging. In addition, the study has provided baselines for clinical assessment and has made contributions relevant to education, basic instruction in psychological aging, and a variety of public policy issues (Schaie, 2008; Schaie, Willis, & Panek, 2005; Schaie & Zanjani, 2006). Within the context of the Seattle Longitudinal Study there have been a significant number of methodological advances, particularly with respect to the design of sequential studies and the relationship between cross-sectional age differences and longitudinal age changes (Schaie, 1965, 1974, 1977, 2005a, b, 2006; Schaie & Caskie, 2005; Schaie & Willis, 2001, chapter 5;). We have also repeatedly addressed the issues of participant attrition (Baltes, Schaie, & Nardi, 1971; Cooney, Schaie, & Willis, 1988; Gribbin & Schaie, 1979; Schaie, Labouvie, & Barrett, 1973; Schaie, 2005a), and that of the problem of practice effects (Schaie, 1977, 1988b).

Acknowledgments

References

Baltes, P. B., Schaie, K. W., & Nardi, A. H. (1971). Age and experimental mortality in a seven-year longitudinal study of cognitive behavior. *Developmental Psychology, 5*, 18-26.

Bengtson, V. L., Schaie, K. W., & Burton, L. (Eds.) (1995). *Intergenerational issues in aging*. New York: Springer Publishing Co.

Bosworth, H. B., & Schaie, K. W. (1997). The relationship of social environment, social networks, and health outcomes in the Seattle Longitudinal Study: Two analytical approaches. *Journal of Gerontology: Psychological Sciences, 52B*, 197-205.

Cooney, T. M., Schaie, K. W., & Willis, S. L. (1988). The relationship between prior functioning on cognitive and personality variables and subject attrition in longitudinal research. Journal of Gerontology, *43*, 12-17.

Gribbin, K., & Schaie, K. W. (1976). Monetary incentive, age, and cognition. *Experimental Aging Research, 2*, 461-468.

Gribbin, K., & Schaie, K. W. (1979). Selective attrition in longitudinal studies: A cohort sequential approach. In H. Orino, K. Shimada, M. Iriki & D. Maeda (Eds.), *Recent advances in gerontology* (pp. 549-551). Amsterdam, The Netherlands: Excerpta Medica.

Gribbin, K., Schaie, K. W., & Parham, I. A. (1980). Complexity of life style and maintenance of intellectual abilities. *Journal of Social Issues, 36*, 47-61.

Gruber-Baldini, A. L., Schaie, K. W., & Willis, S. L. (1995). Similarity in married couples: A longitudinal study of mental abilities and rigidity-flexibility. *Journal of Personality and Social Psychology, 69*, 191-203.

Hertzog, C. (2010). Regarding methods for studying behavioral development: The contributions and influence of K. Warner Schaie. *Research on Human Development*.

Hertzog, C., Schaie, K. W., & Gribbin, K. (1978). Cardiovascular disease and changes in intellectual functioning from middle to old age. *Journal of Gerontology, 33*, 872-883.

Kertzer, D., & Schaie, K. W. (Eds.) (1989). *Age structuring in comparative perspectives*. Hillsdale, NJ: Erlbaum.

Kuhlen, R. G. (1963). Age and intelligence: The significance of cultural change in longitudinal vs. cross-sectional findings. *Vita Humana, 6*, 113–124.

Maier, H. (1995). *Health behaviors in adults: interrelationships and correlates*. Unpubl. doctoral dissertation. University Park, PA: the Pennsylvania State University.

Maitland S. B (1997). *Factorial invariance and concordance of health behaviors and health status: A study of individual differences in familial context*. Unpubl. doctoral dissertation. University Park, PA: The Pennsylvania State University.

Nesselroade, J. R., Baltes, P. B., & Schaie, K. W. (1972). Ontogenetic and generational components of structural and quantitative change in adult cognitive behavior. *Journal of Gerontology, 27*, 222-228.

Riley, M. W., Johnson, M. J., & Foner, A. (1972). *Aging and society: Vol. 3: A sociology of age stratification*. New York: Russell Sage.

Saczinski, J. S., Willis, S. L., & Schaie, K. W. (2002). Strategy use in reasoning training with older adults. *Aging, Neuropsychology and Cognition, 9*. 48-60.

Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin, 64*, 92-107.

Schaie, K. W. (1974). Translations in gerontology - From lab to life: Intellectual functioning. *American Psychologist, 29*, 802-807.

Schaie, K. W. (1977). Quasi-experimental designs in the psychology of aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 39-58). New York: Van Nostrand Reinhold.

Schaie, K. W. (1988a). The delicate balance: Technology, intellectual competence, and normal aging. In G. Lesnoff-Caravaglia (Ed.), *Aging in a technological society* (Vol. 6, pp. 155- 166). New York: Human Sciences Press.

Schaie, K. W. (1988b). Internal validity threats in studies of adult cognitive development. In M. L. Howe & C. J. Brainerd (Eds.), *Cognitive development in adulthood: Progress in cognitive development research* (pp. 241-272). New York: Springer-Verlag.

Schaie, K. W.(1989). The hazards of cognitive aging. *The Gerontologist, 29(4)*, 484-493.

Schaie, K. W. (1990a). Intellectual development in adulthood. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 291-309). New York: Academic Press.

Schaie, K. W. (1990b). Late life potential and cohort differences in mental abilities. In M. Perlmutter (Ed.), *Late life potentials* (pp. 43-61). Washington, DC: Gerontological Society of America.

Schaie, K. W. (1990c). The optimization of cognitive functioning in old age: Predictions based on cohort-sequential and longitudinal data. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 94-117). Cambridge, UK: Cambridge University Press.

Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist, 49*, 304-313.

Schaie, K. W. (1996a). *Intellectual development in adulthood: The Seattle Longitudinal Study*. New York: Cambridge University Press.

Schaie, K. W. (1996b). The natural history of a longitudinal study. In M. Merrens & G. Brannigan (Eds.), *The developmental psychologists*, (pp. 233-249). New York: McGraw-Hill.

Schaie, K. W. (2000). Living with gerontology. In J. E. Birren & J. J. F. Schroots (Eds.), *A history of geropsychology in autobiography* (pp. 233-248). Washington, DC: American Psychological Association.

Schaie, K. W. (2005a). *Developmental influences on adult intellectual development: The Seattle Longitudinal Study*. New York: Oxford University Press.

Schaie, K. W. (2005b). What can we learn from longitudinal studies of adult intellectual development? *Research in Human Development, 2*, 133-158.

Schaie, K. W. (2006). The concept of event time in the study of aging, In J. Baars & H. Visser (Eds.), *Concepts of time and aging* (pp. 121-135). Amityville, NY: Baywood Publishing Co.

Schaie, K. W. (2007). Generational differences; the age-cohort-period model. In J. E. Birren (Ed.), *Encyclopedia of gerontology: Age, aging, and the aged* (2nd edit., pp. 601-609). Oxford, UK: Elsevier.

Schaie, K. W. (2008). Historical processes and patterns of cognitive aging, In S. M Hofer & D. F. Alwin (Eds.), *Handbook on cognitive aging: Interdisciplinary perspective* (pp. 368-383). Thousand Oaks, CA: Sage.

Schaie, K. W., & Abeles, R. A. (Eds.) (2008). *Social structures and aging individuals: Continuing challenges*. New York: Springer Publishing Co.

Schaie, K. W., & Caskie, G. I. L. (2005). Methodological issues in aging research. In D. M. Teti (Ed.), *Handbook of research methods in developmental psychology* (pp. 21-69). Cambridge, UK: Blackwell.

Schaie, K. W., Caskie, G. I. L., Revell, A. J., Willis, S. L., Kaszniak, A. W., & Teri, L. (2005). Extending neuropsychological assessment into the Primary Mental Ability space. *Aging, Neuropsychology and Cognition, 12*, 245-277.

Schaie, K. W., Dutta, R., & Willis, S. L. (1991). The relationship between rigidity-flexibility and cognitive abilities in adulthood. *Psychology and Aging, 6*, 371-383.

Schaie, K. W., & Elder, G. H. Jr. (EDS.) (2005). *Historical influences on lives and aging*. New York: Springer Publishing Co.

Schaie, K. W., & Hendricks, J. (Eds.) (2000). *Evolution of the aging self: Societal impacts*. New York: Springer Publishing Co.

Schaie, K. W., & Hertzog, C. (1983). Fourteen-year cohort-sequential studies of adult intelligence. *Developmental Psychology, 19*, 531-543.

Schaie, K. W., Krause, N., & Booth, A. (Eds.) (2004). *Religious influences on health and wellbeing in the elderly*. New York: Springer Publishing Co.

Schaie, K. W., Labouvie, G. V., & Barrett, T. J. (1973). Selective attrition effects in a fourteen-year study of adult intelligence. *Journal of Gerontology, 28*, 328-334.

Schaie, K. W., Maitland, S. B., Willis, S. L., & Intrieri, R. C. (1998). Longitudinal invariance of adult psychometric ability factor structures across seven years. *Psychology and Aging, 13*, 8-20.

Schaie, K. W., Plomin, R., Willis, S. L., Gruber-Baldini, A., & Dutta, R. (1992). Natural cohorts: Family similarity in adult cognition. In T. Sonderegger (Ed.), *Psychology and aging: Nebraska symposium on motivation, 1991* (pp. 205-243). Lincoln, NE: University of Nebraska Press.

Schaie, K. W., Rosenthal, F., & Perlman, R. M. (1953). Differential deterioration of factorially "pure" mental abilities. *Journal of Gerontology, 8*, 191-196.

Schaie, K. W., & Schooler, C. (Eds.). (1998). *Impact of work on older adults*. New York: Springer Publishing Co.

Schaie, K. W., & Strother, C. R. (1968). A cross-sequential study of age changes in cognitive behavior. *Psychological Bulletin, 70*, 671-680.

Schaie, K. W., & Uhlenberg, P. (Eds.) (2007). *Demographic influences on health and wellbeing of the elderly*. New York: Springer Publishing Co.

Schaie, K. W., & Willis, S. L. (1986). Can decline in adult intellectual functioning be reversed? *Developmental Psychology, 22*, 223-232.

Schaie, K. W., & Willis, S. L. (1993). Age difference patterns of psychometric intelligence in adulthood: Generalizability within and across ability domains. *Psychology and Aging, 8*, 44-55.

Schaie, K. W., & Willis, S. L. (1995). Perceived family environments across generations. In V. L. Bengtson, K. W. Schaie, & L. Burton (Eds.), *Adult intergenerational issues in aging* (pp. 174-209). New York: Springer Publishing Co.

Schaie, K. W., Willis, S. L., & Caskie, G. I. L. (2004). The Seattle Longitudinal Study: Relation between personality and cognition. *Aging, Neuropsychology and Cognition.11*, 204-234.

Schaie, K. W., Willis, S. L., & Pennak, S. (2005). A historical framework for cohort differences in intelligence. *Research in Human Development, 2*, 43-67.

Schaie, K. W. & Zanjani, F. A. K. (2006). Intellectual development across adulthood. In C. Hoare (Ed.), *Handbook of adult development and learning* (pp. 99-122). New York: Oxford University Press.

Schooler, C. (1987). Psychological effects of complex environments during the life span: A review and theory. In C. Schooler & K. W. Schaie (Eds.), *Cognitive functioning and social structure over the life course* (pp. 24–49). Norwood, NJ: Ablex.

Schooler, C., Mulatu, M. S., & Oates, G. (1999). The continuing effects of substantively complex work on the intellectual functioning of older workers. *Psychology and Aging, 14*, 483–506.

Thurstone, L. L., & Thurstone, T. G. (1949). *Examiner manual for the SRA Primary Mental Abilities Test* (Form 10-14). Chicago: Science Research Associates.

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore: Williams and Wilkins.

Willis, S. L. (1987). Cognitive training and everyday competence. In K. W. Schaie (Ed.), *Annual Review of Gerontology and Geriatrics* (Vol. 7, pp. 159-188). New York: Springer Publishing Co.

Willis, S. L. (1989). Improvement with cognitive training: Which dogs learn what tricks. In L. W. Poon, D. C. Rubin & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 545-589). Cambridge, UK: Cambridge University Press.

Willis, S. L. (1990). Current issues in cognitive training research. In E. A. Lovelace (Ed.), *Aging and cognition: Mental processes, self-awareness, and interventions* (pp. 263-280). Amsterdam, The Netherlands: Elsevier.

Willis, S. L. (2001). Methodological issues in behavioral intervention research with the elderly. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (5th ed., pp. 78-108). San Diego, CA: Academic Press.

Willis, S. L., & Nesselroade, C. S. (1990). Long-term effects of fluid ability training in old age. *Developmental Psychology, 25*, 905-910.

Willis, S. L., & Schaie, K. W. (1986a). Practical intelligence in later adulthood. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Origins of competence in the everyday world* (pp. 236-268). New York: Cambridge University Press.

Willis, S. L., & Schaie, K. W. (1986b). Training the elderly on the ability factors of spatial orientation and inductive reasoning. *Psychology and Aging, 1*, 239-247.

Willis, S. L., & Schaie, K. W. (1988). Gender differences in spatial ability in old age: Longitudinal and intervention findings. *Sex Roles, 13*, 189-203.

Author Note

Experience Sampling Method: Profiling Indian Families

Deepali Sharma
Panjab University, Chandigarh, India
E-mail: deepali1sharma1@gmail.com

and

Suman Verma
Panjab University, Chandigarh, India
E-mail: deepali1sharma1@gmail.com

Time use studies have been widely carried out in the past few decades to examine the amounts of time spent by people in varied categories of activities such as schoolwork, household work, office work, and leisure (see Larson &

Verma, 1999 for a detailed review of time budget studies). Along with providing snapshots of daily activities, time budgets also give insights into the amount of time spent with different companions and facilitate evaluation related to individual profiling across different days of the week. Assessment of the time spent by people in various contexts provides researchers with an opportunity to examine subjective experiences associated with each setting. The data also enhance our understanding of the lifestyles and experiences of individuals and provides a meaningful indication of the pattern of social behavior (Larson & Richards, 1994). Thus, time use studies afford scope for studying a framework that outlines the socialization and developmental opportunities available for the target population (Czikszentmihalyi & Larson, 1984, 1987). This information has implications for almost all branches of social planning and policy making.

The present paper reviews studies carried out by the authors utilizing the Experience Sampling Method (ESM) as an approach to study time use among a select sample of middle-class Indian adolescents and their families. The paper also elaborates the advantages and data collection particularities related to ESM. The ESM data on the Indian sample is part of a research project carried out by the second author in collaboration with Reed Larson of the University of Illinois at Urbana-Champaign, USA. The focus of the project was to examine how cultural change is shaping the lives of Indian adolescents and their families. In order to do so, we recruited members of 100 families from the city of Chandigarh, an urban area of 1.2 million inhabitants, in northern India, for in the study.

The respondents (mothers, fathers, and 8th graders) wore alarm watches for a period of one week and were beeped eight times in a day, across their waking hours. After each signal, the participants filled out self-report forms and provided information related to their daily contexts including *who* they spent their time with, *where* they spent their time, and *what* they did during that time. Along with these indicators, subjective experiences were reported by the participants after each signal which provided an insight into their emotional, cognitive, and motivational states. Fathers in the study responded to an average of 78.2% signals, the mothers to 83.1% signals and the adolescents responded to 85.9% of the signals. Data comprising over 13,600 ESM self reports were collected and on a questionnaire completed at the end of the week, participants reported that the missed reports occurred most often during sleep, bathing, job related activities, and other activities at home. The ESM reports provided a representative sample of the typical activities and states experienced by the family members, particularly because the signals occurred randomly during the day. While individuals were the main unit of analysis, the ESM data also recorded the interrelationships between the lives of the family members and the points of intersection.

Following the guidelines of Larson and Delespaul (1992), the analysis of the ESM data involved converting measures of relative positions of subjective states to *z*-scores in order to make them comparable across respondents. These were obtained by subtracting the respondents' mean across all of his or her self-reports from each value and then dividing by the respondents' overall standard deviation. Z-scores have the advantage of representing respondents'

experience in a given context relative to the rest of their experience and also to eliminate variance due to personality differences and response tendencies. On the z-scored scales, a value of 0.0 corresponds to each person's mean and an interval of 1.0 corresponds to each person's standard deviation for that scale.

Summary of Findings

A brief summary of the ESM findings related to school work, adolescents' emotional profile, women's experience of household labor, fathers' involvement in families, and family patterns of TV viewing is now presented. The first set of findings is related to the school experience of the adolescents and their associated subjective states (Verma, Sharma, & Larson, 2002).

Influence of School Demands on Daily Time Use and Subjective States. The adolescents in the study (mean age = 13.25 years) were found to spend large portions of their waking time in school work and school-related activities, with girls spending more such time than the boys. After school, the adolescents spent additional time being tutored and finishing homework or doing extra studies. Schoolwork was seen to generate negative states and the respondents very often reported feeling bored, anxious, or unhappy, and having the perception of low choice with the work that they were doing. These negative states were most frequent during homework. The adolescents who spent more time doing homework experienced more negative emotions while those who spent more time in leisure experienced more positive affect; however, the latter also reported higher academic anxiety and lower academic achievement.

Adolescents' Emotional Profile. In order to test the hypothesis that adolescence is a period of greater emotionality, we analyzed the ESM data to construct emotional profiles of the adolescents and their parents (Verma & Larson, 1999). The results indicated that teenagers displayed wider emotional swings than their parents, thus, implying that not only did they feel more negative emotions such as being sad or angry but also reported positive emotions such as feeling happy, great, and excited more often than their parents. Also, the negative emotional states of adolescents were related to school work and were inversely related to family and peer variables, thus suggesting the crucial roles that family and peers play as emotional buffers.

Women's Experience of Household Labor. ESM findings on mothers' involvement in the daily household chores suggest that the mothers spent four to seven times more of their time on household labor and child care than their husbands did (Verma & Larson, 2001). Furthermore, the contribution of the fathers in household chores increased only marginally in cases where the mothers were employed. The employed mothers were thus involved in a "double shift" between the office and the home. However, the mothers did not report markedly negative emotions while carrying out the housework and felt it was an activity of their own volition along with the feeling of being in control.

Fathers' daily work and family life. The time use profiles (Larson, Verma, & Dworkin, 2001) indicated that although the fathers spent less time in the family sphere, they were still very involved in the everyday interpersonal lives of their families. They spent substantial amounts of time with

their children, were not perceived as authoritarian, and felt relaxed with the family. Further, while the fathers reported a greater focus of attention on work than on home life, there was interdependence between the fathers' work and family lives suggesting a functional relationship of strain and recuperation. Finally, the fathers did not consider themselves as the main authority figure and felt that everyone was equally involved when it came to decision making.

Patterns of television viewing. The ESM data also provided information related to the television viewing patterns of the adolescents since this is such an integral part of the family experience in India. The findings suggest that the amount of time that the adolescents spent with their families was more than they spent with their friends and that this family time was also a positive experience for them, with television viewing a major family activity (Verma & Larson, 2002). The self reports indicated that TV viewing occupied about 12 hours per week for the children, with the majority of the television viewing occurring at home, along with the parents. The rates of TV viewing were also correlated with the mothers' employment such that more television was viewed by the adolescents when the mothers were not working. During TV viewing, adolescents reported lower anxiety levels which could be seen as an antidote to an otherwise stressful day comprising school work, being tutored, and additional independent study tasks.

Continuity versus Change within the Indian Family Setup

The findings of the ESM data from the studies reported above suggest middle-class Indian families to be in a state of transition. However, the results of the ESM studies also indicate many ways in which the Indian families and the upbringing of the adolescents continue to follow the traditional model. Thus, even though the adolescent respondents spent a lot of their time engaged in school work and after school classes, family time remained a high priority. The time spent with the family (such as watching television together, meal time, and general conversations) was also seen to buffer somewhat the academic anxiety and work overload that the eighth-grade respondents associated with school. The amount of time that the adolescents spent with their families was even higher than that spent with their peers. This set of findings is different from those of the ESM studies carried out in western societies which suggest that peer group activities, especially during adolescence, take precedence over family time and activities carried out therein (Csikszentmihalyi & Larson, 1984; Larson, Richards, Sims, & Dworkin, 2001).

Regarding the mothers in the study, even though a substantial percentage of the mothers were employed, they were still involved in the traditional role of assuming complete responsibility for the household tasks after getting home from work. While doing housework, they felt in charge and in control, thus suggesting a clear demarcation of the roles and responsibilities within the household. Similarly, the ESM data suggest that even though the fathers' roles are apparently undergoing change, with fathers being more and more involved in the lives of the family members and reporting an egalitarian attitude toward family decision-making, their limited involvement in the household chores suggests the continuity of a traditional model within the middle class Indian family setup.

Studying Indian Families under the ESM Microscope

The ESM studies carried out within the Indian context allowed the authors the opportunity to examine family dynamics and interrelationships at close quarters. While this was by no means an easy task, given the lifestyle-related time pressures of the parents and their anxiety that the daunting eight-signals-a-day regimen, over a period of one week, would interfere with their child's study schedule, the data available at the end of the study period were worth the time and effort of all involved with the project. Herein lie the unique strengths and advantages of the ESM approach. Not only does the ESM avoid dependency on recall, which is typically used in the diary method, it also reduces the intrusive observation by the investigator as is typically employed in the observation method. This approach thus provided us with a representative sample of daily life experiences in terms of activities, thoughts, location, and subjective feelings of the respondents along with giving us insights into the creative world of the adolescents in their free-to-doodle space.

Along with the advantages of the study methodology and its uniqueness, it being the first of its kind in India, the data allowed us to observe Indian families closely and try to understand them like a mystery being unraveled. For example, why was the mother, who had returned from the office, with no domestic help, and no support from her husband in doing household chores, still feeling cheerful, very much in control, and checking the box "high volition" on the self report form? The answer was suggested in the statement that one mother gave immediately following the options that she had checked: "at least this is one place where I am my own boss." For the researchers, that explanation is indeed very helpful.

For the Indian sample, the limitations of the study included not being able to test a causal model due to the cross-sectional nature of the data; having a small sample of 100 families which did not necessarily represent the varied socio-economic and regional disparities in India; and having a volunteer sample as participants.

Will we repeat the ESM approach again in order to understand the ever-changing patterns in the Indian scenario? The answer hinges on certain issues, namely, being able to manage researchers' own subjective states when respondents drop out of the study mid-week; becoming more adept at responding to statements from the subjects such as, "we did not know what we were getting into"; and realizing that human errors can be made during any stage of data collection and that these can be rectified to a large extent. Having resolved these issues, the ESM will continue to offer meaningful insights into current research topics.

References

Csikszentmihalyi, M., Larson, R. (1984). *Being adolescent: Conflict and growth in the teenage years*. New York: Basic Books.

Csikszentmihalyi, M., Larson, R. (1987). Validity and reliability of the experience sampling method. *Journal of Nervous and Mental Disease, 175*, 526-536.

Larson, R., & Delespaul, P. (1992). Analyzing experience sampling data: A guidebook for the perplexed. In M. DeVries (Ed.), *The experience of psychopathology* (pp. 58-78). Cambridge: Cambridge University Press.

Larson, R., & Richards, M. (1994). *Divergent realities: The emotional lives of mothers, fathers, and adolescents*. New York: Basic Books.

Larson, R., Richards, M.H., Sims, B., & Dworkin, J. (2001). How urban African American young adolescents spend their time: Time budgets for locations, activities, and companionship. *American Journal of Community Psychology, 29*, 565-597.

Country Focus

Broadening the Scope of Research Possibilities in Guatemala

Francisco Ureta
Universidad del Valle de Guatemala, Guatemala
E-mail: fureta@uvg.edu.gt

and

Ruby Batz
Universidad del Valle de Guatemala, Guatemala
E-mail: fureta@uvg.edu.gt

and

María del Pilar Grazioso
Universidad del Valle de Guatemala, Guatemala
E-mail: fureta@uvg.edu.gt

Educational research in Guatemala and Central America is scarce; the limited funds that are available in the educational system and universities for this purpose reinforce this weak development. This article offers a summary of the research and publications by the Center for Educational Research and a proposed model for implementing effective services in Guatemala in an effort to enhance the investment in child development. These studies were conducted by the Center for Educational Research, a member of the Research Institute of the Universidad del Valle de Guatemala (UVG). This center was founded in 1986 and has since conducted evaluations and research of education in Guatemala; it is the only center in the country with solid longitudinal data related to children's development.

Our concern with the educational thematic approach focuses on student achievement, specifically in the official primary schools in Guatemala, in first, third and sixth grades, considering some variables that influence the achievement. The center has focused attention on the process of learning the English language and the verbal and mathematical reasoning of high school students and aspirants to universities in Guatemala. Among results observed are evident differences between urban and rural students, between ladino (a Spanish-speaking ethnic group in Guatemala) and indigenous children, and between girls and boys. The clearest results were obtained among urban male ladino students. In high school, women perform better than men in English as a second language, but this difference disappears during university studies. The verbal and numerical reasoning of male high school students is stronger than in females, a difference that persists up to the time of application for admission to universities.

Academic Skills in Middle-level Students and Those Aspiring to Enter Universities in Guatemala

This study arises from a concern about the quality of verbal and mathematical reasoning of young Guatemalans (Ureta, F. 2009). One can obtain data of this kind through the Prueba de Aptitud Académica (PAA) designed by the College Board, which has been administered in Guatemala since 2001. This study includes data from 2001 through 2008, during which time 15,768 students were assessed– 2,353 in middle school, of whom 995 are males and 1,358 are females; 12,804 freshmen in universities in Guatemala, comprising 5,917 males and 6,887 females. There is a remainder of 611 that cannot be located in any of the previous two levels, due to the lack of information regarding the institution where they were assessed. The objective of this study was to explore questions of performance differential by comparing test results between males and females in the total test and in the mathematical and verbal scales assessed in the PAA test, and the effect of these differences.

Males perform better than females in the total scale, as well as in verbal and mathematical reasoning at university level; similar results appear in the high schools, where males do better in the total scale and in the mathematical reasoning scale, but not in the verbal reasoning scale, where they have the same performance. The effect size of the differences found in the diversified level is small but, when students are applying for college entrance the effect size becomes moderate, especially in mathematical reasoning and total PAA, results that confirm the findings made with independent samples (*t* test differences of mean).

These results confirm previous studies conducted in Guatemala from the elementary to middle level, where boys consistently performed better than girls in reading and math, for first, third, sixth and ninth grades. They also agree with studies conducted with college students in Chile, where men perform better than women, especially in mathematics. However, these results are not consistent with other studies conducted at universities in the U.S., Australia and other developed countries where such differences have disappeared. The findings have educational and teaching implications for mathematics and language teachers from those schools and universities where men and women have the same performance, as well for other educational establishments where differences remain between men and women.

ELASH in Guatemala, an Evolutionary and Gender Analysis

How is the learning of English as a second language among secondary and university students in Guatemala conducted? To answer this question, information available through the English Language Assessment System for Hispanics (ELASH) was examined. ELASH is an integrated program that assesses English as a second language, and is designed especially for the Hispanic population by the Office of Puerto Rico and Latin America by the College Board (Ureta, F. 2008). In its development, the linguistic interference between English and Spanish was taken into account. Since it was implemented by the UVG in 2000, 21,803 people have been assessed. Currently it is applied in 26 private schools, which

use the results as a parameter of the effectiveness of their English programs. It is also used in various universities in Guatemala (Universidad del Istmo and UVG), with specific functions in each: entrance exams and diagnosis of language proficiency, placement, or exemption from courses in English requested by each university or graduation requirement.

In conducting the analysis of the results obtained during its 8 years of implementation, it appears that in 2001 and 2002, men and women had the same performance in the English language, but in 2000, 2003, 2004, 2005, 2006 and 2007 women had a better performance than men. The analysis of the areas that ELASH evaluates indicated that the difference becomes more noticeable in grammar. Fewer differences are seen in forms 2A, 2B and 2C which include the upper grades of high school. In primary and basic grades the differences in language skills are more apparent in the female gender, and eventually the differences are reduced and no longer are significant. It appears that females make a better start in English language proficiency, but there are no differences between genders as the level of difficulty increases. It is necessary to pay special attention to the listening and reading scales, since the lowest results obtained by the persons assessed in Guatemala were in these scales.

Correlation between Micronutrients and Achievement in Reading and Math in Sixth Graders in Rural Primary Schools—a Nationwide Sample Study

Is the appropriate intake of vitamin A and iodine among sixth graders in official primary schools related to their performance in reading and math tests? This question guided this portion of the study (Ureta, Recinos, & Martínez, 2008). In September 2005 the Center for Research in Education (abbreviated CIE in Spanish) of the UVG, at the express request of the Guatemalan Ministry of Education, conducted a study of the correlation between micronutrients and performance in reading and math in sixth graders in rural primary schools, with the financial and technical support of the United Nations Fund for Children and the Institute of Nutrition of Central America and Panama. The aim of this study was to correlate the level of attainment and achievement in reading and math with the iodine levels present in urine and salt samples, and vitamin A in sugar samples. The sample was formed by 450 rural schools in 22 departments of the country; salt and sugar samples were taken from 2,450 students, of which 1,183 were girls (48.3%) and 1,267 were boys (51.7%). The children who had valid samples of salt and sugar consumed in their home was 1,708, from which 818 were girls (47.8%) and 889 (52.2%) were boys. We obtained 167 urine samples, 75 from girls (44.9%) and 92 from boys (55.1%). The selection students for the administration of the tests, the taking of urine samples, and the sampling of salt and sugar samples from students' homes was done randomly.

The results verified that the level of iodine in urine and salt samples correlated with the level of attainment and achievement in reading, but there is no correlation between these levels and mathematics achievement. We can say with a confidence level of 95% that the consumption of an appropriate level of iodine supports the students' verbal learning process by supporting neonatal and postnatal neural development. In the case of vitamin A in samples of sugar consumption, the evidence indicates that it has no correlation with the level of achievement in reading or mathematics. Vitamin A is related to the retina of the eyes which less directly related to the learning process. These findings have implications in improving the manufacture of salt in Guatemala, as well as food support programs that are conducted by the Ministry of Education, by routinely adding iodine to salt.

Performance in Reading and Math in First and Third Graders, a Nationwide Sample Study

What is the performance level in reading and math among first and third graders in primary official schools in Guatemala? This question was submitted by the Ministry of Education, and it was answered in this study (Ureta, Fortin, & Molina, 2006). The evaluation was conducted by the Center for Research in Education, from the Research Institute of the UVG, based on the National Assessment of Educational Achievement (abbreviated PRONERE in Spanish), at the request of the Ministry of Education of Guatemala. During 2004, PRONERE conducted an evaluation in 823 public schools regarding the reading and math performance of students in first and third grades. In first grade instruments were used referring to criterion, in third grade and used instruments referred to the norm. Interviews and questionnaires were administered to students, teachers and principals. In first grade, tests were conducted in Spanish, kaqchikel and k'iche'; in the third grade, tests were given in mam and q'eqchi' in addition to those already mentioned.

The sample, which was made in panels and for multiannual applications, is nationally and departmentally representative. It considered the variables of geographic location (urban-rural) and school size (number of students attending). In estimating the results we considered the weight that each layer has on people around the country. First grade (19,530) and third grade (16,770) students were assessed.

In first grade, 48% of the students met the reading criterion and 27.5% met the mathematics criterion. Urban students, boys and schools with mostly non-indigenous students performed better. The third primary scores were standardized to an average of 100 points as an arbitrary national average, and students of urban and-non indigenous schools performed better, with no gender differences. The level of achievement in the two grades is generally poor. Associated variables that influence the performance are: if the student had breakfast the day of evaluation, levels of insecurity nationwide, the affinity for the school, literate environment at school and home, the teacher's level of experience, and school infrastructure (piped water and electricity).

Broadening the Scope of Research Possibilities in Guatemala

Social investment in Guatemala continues to be one of the lowest in Latin America. As a country where children and adolescents compose more than half of its population, 50 percent of them living in poverty (UNICEF, 2010), Guatemala needs to serve these youth by investing in programs

These pictures refer to the test applications to first and third graders who participated in the Achievement research mentioned in this article

that are empirically validated. The implementation of evidence-based practices (EBP's), comprising careful observation and analysis, is far behind in Guatemala due to constraints on human and financial resources, weak systems, lack of infrastructure, lack of explicit guidelines and policies to inform and reinforce good practices, inadequate knowledge of research skills, straining the already limited resources and resulting in further inequalities. Improved nutrition, careful observation, and increased resources can provide a start.

A new Proposal to Invest in Children: Effective Services in Guatemala

How can we ensure that the programs intended to alter the developmental trajectories in disadvantaged children and protect them against risk factors have the features to succeed and be maintained over time? A thorough and comprehensive Systems Change approach becomes an essential prerequisite for understanding and changing

practices, assuring the efficient use of resources (Batz, 2009). Early Human Development programs, as described by Fixsen, Blase, Naoom, and Wallace (2009) and Fixsen, Naoom, Blase, Friedman, and Wallace (2005), can provide all children the opportunity of being successful while working with the available resources. The responsibility of serving the youngest, with the possibility of changing the future, is enough reason to consider a systematic process of implementation, which our center is now pursuing in Guatemala.

Author Note

References

Batz, R. (2009). *Investing in Children: A proposed model for Implementing Effective Services in Guatemala*. Unpublished manuscript.

Fixsen, D., Blase, K., Naoom, S., & Wallace, F. (2009). *Core Implementation Components. Research on Social Work Practice.* Vol.000, p.1-10.

Fixsen, D., Naoom, S., Blase, K. Friedman, R. & Wallace, F. (2005). *Implementation Research: A Synthesis of the Literature.* Tampa, FL. University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication # 231).

Ureta, F. (2009). Academic skills in middle-level students and applicants to universities in Guatemala. *Revista de la Universidad del Valle de Guatemala, 19*, 66-73. Guatemala. Available online at: http://www.uvg.edu.gt/ revista/Revista%2019.pdf

Ureta, F. (2008). ELASH in Guatemala, and evolutionary analysis by gender. *Revista de la Universidad del Valle de Guatemala, 18*, 131- 143. Guatemala. Available online at: http://www.uvg.edu.gt/revista/Revista%2018.pdf

Ureta, F., Fortin, A., & Molina, R. (2006). Reading and Mathematics Performance of Students in first and third primary. *Revista de la Universidad del Valle de Guatemala, 15, 9-37*. Guatemala. Available online at: http://www.uvg.edu.gt/revista/Revista%2015.pdf

Ureta, F., Recinos, S., & Martínez, A. (2008). Correlation of Micronutrients and reading and math performance of students from the sixth primary rural schools. *Revista de la Universidad del Valle de Guatemala, 17*, 18-33. Guatemala. Available online at: http://www.uvg.edu.gt/ revista/Revista%2017.pdf

UNICEF Guatemala: perfil de país (2010). Retrieved on 02.02.2010 from http://www.unicef.org/guatemala/ spanish/overview.html

Notes from The President

What Should ISSBD Accomplish?

The realization that this is my last article for the *Bulletin*, (previously *Newsletter*) was startling. It seems like just yesterday that we were starting out on this journey together building ISSBD. Now I recognize that a presidency is more like a relay race; my turn is ending and I will pass the baton at the end of ISSBD2010 into the capable hands of Wolfgang Schneider, President–Elect.

We have accomplished much together! The purpose I articulated in my first message to all of you in 2006 is for ISSBD to advance developmental science through theory, methods, and research, and to serve society and humankind in doing so. The two most important contributions to our field are surely our biennial meeting and our flagship journal. Both are doing better than ever! IJBD Editor Marcel van Aken and his outstanding team of Associate Editors, are bringing to ISSBD members important research that advances understanding of human development in global context. Karina Weichold and Bonnie Barber have continued to strengthen the ISSBD *Newsletter*/now *Bulletin*. Thanks also to Andy Collins, and his Publications Committee, for their sage advice on publication issues.

The ISSBD 2010 meeting will be the third iteration of our new approach to conference sponsorship. Building from the ground-breaking work of Ann Sanson and colleagues in 2006, and Wolfgang Schneider and colleagues in 2008, Robert Serpell and colleagues are creating an ISSBD2010 meeting that bring the issues and scholarship of Africa more strongly into the rich mix of strengths that ISSBD represents. More on ISSBD2010 later.

We have also worked hard to increase the effectiveness of ISSBD, both scientifically and administratively. Suman Verma and Catherine Cooper conducted an important study (reported in the Fall 2009 *Bulletin*) on ISSBD's regional workshops, to enhance the effectiveness of these workshops for ISSBD members in the region. Building from that report, the 2009 ISSBD Regional Workshop in Kenya led by Paul Oburu was outstanding, led from the region and initiating two important experiments in capacity building. Paul also enrolled a record number of new ISSBD members. (I was unable to attend the regional workshops in China and AustralAsia but am sure that these were also effective.) I believe that these efforts will more meaningfully engage ISSBD members who may not be able to attend the biennial meeting, and build human development research capacity more broadly.

With ISSBD's growth, we have also attended to issues of membership and financial management. Ann Sanson, with Membership Secretary Xinyin Chen, led a committee to examine membership distribution, fees and subsidies, among other issues; the recommendations of their excellent work have already been partially implemented. Finances turned out to be much more difficult, despite the outstanding efforts of Treasurer Ingrid Schoon and Finance Committee Chair Liz Susman, because of the changes in banking that limit how transactions can be conducted. ISSBD has now approved a small contract for professional financial management, still under the direction of the Treasurer and President (and the entire Executive Committee, when appropriate). This issue has taken much more time than it deserves and we are hopeful that going forward the Treasurer and the President can focus on their primary purposes. I know that ISSBD Secretary General Katharina Salmela-Aro will welcome this additional assistance for her great work as well.

Support for ISSBD management has also come from SAGE Publications, and specifically through the good efforts of our liaison, Kerry Barner. Kerry tirelessly works on our behalf, not only for all of our publications but also in support of membership, elections, our website, and many other functions. ISSBD remains an organization run by volunteers but without Kerry's and SAGE's support, we would be much less effective.

The ISSBD Young Scholars are another very bright spot on ISSBD's landscape. Under the leadership of the first Young Scholar Executive Committee member, Zena Mello, the Young Scholars are a very lively and engaged part of ISSBD. Major highlights include a new Young Scholar elected position on the Executive Committee, young scholar representation on every ISSBD committee, as well as Young Scholar leadership with some major efforts. This group assures us that ISSBD and the field have a strong future with their talent!

We hope that most of you are able to join us for a spectacular meeting in Lusaka, Zambia in July. This is our first ISSBD meeting on the continent of Africa. Robert and his colleagues have been working hard to make this an outstanding meeting scientifically, professionally, and culturally.

I will say much more about ISSBD's status and future challenges for our field at the meeting in Zambia, both in my Presidential address, and in my comments at the business meeting. I was just moving to Stanford at the beginning of my Presidency and have now moved to the University of Michigan. Please let me know if you have any suggestions to share. Thanks for all your good efforts!

Anne C Petersen, President
University of Michigan, USA
E-mail: annepete@umich.edu

Workshop Report

Report of the 8th African Regional Workshop on Building Young Scholars' Capacity in Human Development Research: Maseno University, Kenya, November 30th to 2nd December 2009

Paul Oburu
Maseno University, Kenya
E-mail: poburu@yahoo.com

The 8th African ISSBD Regional Workshop was held at Maseno University, Kisumu, Kenya between 30th November and 2nd December, 2009. The workshop was attended by over 100 early career, middle-level, and senior scholars from Europe, USA, and East, West, Central and Southern Africa. The meeting was officially opened by Professor Dominic Makawiti, the Deputy Vice-Chancellor in charge of Academic Affairs (DVC-AA). The DVC-AA, who read the speech on behalf of the Vice-Chancellor, Prof. Frederick Onyango, reiterated Maseno University's unreserved support for the workshop and also commitment to nurture early career researchers.

One of the main objectives of the 8th African regional workshop was to provide a forum for training junior scholars on a wide range of methodological approaches in human development research and allied areas. The other aim was to equip early career researchers with skills that they might need to complete their already identified research topics and also increase their research and publication capacity. The three-day meeting comprised training sessions addressing career planning and advances in research methodology, scientific writing skills, and publication and dissemination procedures. There were also four keynote discussion forums and poster presentations where eligible PhD candidates made presentations on their ongoing or proposed work.

The senior scientists who were instrumental in the training programs and the young scholars' initiatives included the following:

1. *Anne C. Petersen*: The President outlined the main goals and mission of ISSBD. She stipulated that ISSBD's current focus was on broadening the membership to include a higher representation of junior scholars and also to increase engagement, sustainability and continued financial stabilization of the society. The president also encouraged young scholars to apply for limited travel grants available to those wishing to attend the 2010 Lusaka conference. Anne Petersen also informed participants about the existence of grants for early career researchers from the Portuguese speaking African countries. She urged young scholars to increase their interactions with senior scholars so as to raise the possibility of publishing their work. She also made presentations on Advances and Methodological Approaches in Human Development Research. She challenged emergent African scholars to increase their engagement in longitudinal research processes since development implied assessment of change over long periods of time. She noted that with advances in modern computing it is now possible to model complex systems. She further identified areas in quantitative research that need emphasis by young scholars interested in advanced research approaches. These included errors in data entry, measurement of phenomena, missing data analysis, sample selection bias and interpretation of findings.

2. *Robert Serpell (University of Zambia, Zambia)*: Robert made two presentations on culturally appropriate measurement scales from Zambia. These two non-verbal cognitive scales were Panga Muntu & Home Potential Assessment (HEPA). These scales were developed by Zambians who realized that their children fared badly on imported structured cognitive test items. Robert Serpell also provided participants with relevant information about the 2010 ISSBD Meeting in Lusaka, Zambia.

3. *Marian Bakermans-Kranenburg (University of Leiden, Netherlands)*: Her presentation focused on the promise of Biological & Physiological Research Methods. Marian explained how environment, behavior and molecular genetics could be used to explain attachment processes. In her presentation she highlighted how recent studies show that children who have often been considered 'vulnerable' (difficult temperament, or genetically at risk) may actually be more susceptible to both negative *and* positive environmental influences. She recommended having an open eye for such moderator effects, that is, associations between predictors and outcomes that are significantly different for subgroups. Assessing temperament as a potential moderator may be useful, and even genotyping is not as difficult as it may seem to be: the use of mouth swabs is straightforward, and genotyping labs are available. She stressed that careful assessment of the environment is very important to detect gene-environment interaction effects, and that this cannot be left only to geneticists.

4. *Penny Holding (Kemri-Welcome, Kenya), Arthur Whaley (Jackson State University, USA) & Amina Abu-Bakr (Tilburg University, Netherlands)*: These authors made a joint presentation on scientific writing skills, peer reviews, publication and dissemination procedures. Their main focus was on the planning and decision making processes involved in the manuscript preparation, review, publication and dissemination. They gave a detailed account on different parts of the manuscript and also suggested reasons for acceptance or rejection of manuscripts. They also trained participants on how to effectively review literature relevant to their research areas of focus. The participants were also introduced to popular literature search bases and provided with detailed guidelines on how to plan and critique research articles meant for publication.

5. *Julie Robinson (Flinders University, Australia)*: Julie made two presentations. Her first presentation focused on career planning and orientation to the 8th African regional workshop. Julie helped participants to identify their ideal research future and the skills and achievements necessary to reach the ideal or closest approximation to that which is practical given their resource and skills limitations. Her skills training program was tailored on Russell Ackoff's 'ends-means' career planning approach. Her argument was that if participants were to set their priorities right, and effectively planned their research careers within well developed specific timeframes, it would be possible for them to achieve their preferred research goals. The presenter's other 'take home message' was that all of us had equal time; the difference between the most and least successful researchers was on how they prioritized their most important life needs and managed distractions.

Her second presentation, made jointly with Robert Serpell, provided background information on research mentoring and incubation. They raised the possibility of setting in place a virtual research incubator and mentoring project as viable processes that could be used to support early career researchers in Africa. The outcome of these two novel ideas will be evaluated during the ISSBD Lusaka Conference. Julie also encouraged the participants to view the workshop as an avenue for creating linkages, partnerships and collaborations with other young scholars and also as opportunities for both formal and informal interactions with potential mentors.

6. *Bame Nsamenang (University of Yaounde 1(ENS Bambili) & Human Development Resource Centre, Cameroun)*: Strategies for bringing Africa's developmental knowledge into developmental science. Bame shared his own research expedition and lifetime commitment in making Africa visible in the international research superhighway. The author also identified research niches that African early career researchers could contribute to the research agenda. He pointed out the need to give developmental research an Africentric perspective. The presenter pointed out the need to contextualize Western theories and methods to African indigenous systems. He also emphasized his strong belief that all cultures deserve recognition in the generation of scientific knowledge.

7. *Marinus van IJzendoorn (Leiden University)*: Child maltreatment and its impact on child development. Marinus compared the prevalence of child maltreatment and Institutional or structural neglect in the Netherlands, Athens (Greece), Odessa (Ukraine) and Mumbai (India). He also explored the relevance of institutional care in Africa given the increased prevalence of HIV/AIDS. He also gave a brief comment to Bame Nsamenang's presentation. His main argument was that human developmental science was a universal enterprise that did cut across artificially created borders. There were lots of shared commonalities between ethno-cultural boundaries. The gist of his presentation was that instead of compartmentalizing Africa's contribution to developmental science, efforts aimed at academic visibility should be tailored towards incorporating Africa's contributions into a universal science of cultural divergence and convergence in human development.

8. *Peter Baguma (Makerere University, Uganda)*: Relevance of Developmental research to social policy issues in Africa. Peter explored research-policy interrelations in the African continent. Specifically he provided suggestions on how research could be utilized to influence policy interventions and practices in Africa.

9. *Therese Tchombe (University of Buea & Centre for Research in Child &Family Development and Education, Cameroun)*: Her presentation focused on the need to network in support of addressing Africentric issues in research processes. Therese gave an analysis to the impediments to doing 'quality' research and dissemination of findings in Africa. Her suggestion was that through collaboration, sustained networking and creation of African institutional partnerships, viable solutions to building research capacity in Africa could be enhanced. Some of the other solutions to the endemic research challenges would require Africentric researchers to share resources including personnel and student supervisory responsibilities across the continent. Such an arrangement could be enhanced only if African universities were interconnected and research groups deliberately pooled their resources, and increased their intra-continental mobility.

10. Three young scholars, *Maureen Mweru (Kenyatta University, Kenya), Muma Solomon (Kenyatta University)*

and *Joab Kinzi(Egerton University, Kenya)* were also given opportunities to present their work. Maureen presented a paper on the use of massed education practice in Kenyan primary schools. Joab Kinzi's presentation was on intrinsic drives and achievement of students in the core sciences in Kenyan secondary schools while Solomon focused on how to manage compassion fatigue.

The workshop's other intended aim was to initiate the development of an international mentoring program that will link early and senior career African and international scholars. The mentoring concept was first discussed by an African Research Advisory Panel (ARAP) during their meeting at the University of Buea, Cameroun in August 2009. A meeting convened by Robert Serpell in partnership with ARAP members present during the workshop (Peter Baguma, Paul Oburu, Bame Nsamenang, Therese Tchombe, Ester Akinsola) suggested that a collaborative triangle consisting of early career African researchers, senior African scholars and internationally recognized scholars be created to build African research capacity. Early career scholars were expected to be African graduates either based in an African university or having established links with an African research institution. A senior African scholar was defined as a suitably qualified scholar in the service of academic faculties of an African University or research institute. International scholars will be those based outside the continent willing to mentor or provide technical guidance to early career researchers. Their specific roles varied between being a co-supervisor to a consultant on matters pertaining to the students' research work and institutional capacity building.

In order to implement the mentoring program, eligible candidates were notified beforehand of the existence of the competitive award scheme and the formula that was to be used in the selection process. Posters displayed during the workshop were independently rated on a 10-point scale by a panel consisting of eight senior scientists participating in the workshop. This summary of ratings based on the reviews presented by the senior scholars were used in selecting eight best posters. These eight early

career researchers included Beatrice Matafwali (Zambia), Bestern Kaani (Zambia), Catherine Mbagaya (Kenya), Patricia Kadzo Kitsao-Wekulo (Kenya), Njungwa Martina (Cameroun), Achiri Jacky Lum (Cameroun), Ume Charles (Nigeria), Kinzi Joab (Kenya) and Tande K. Epse (Cameroun). Five of these students will benefit from funds allocated by Johan Jacobs Foundation (JJF) for the support of early career researchers. Anne Petersen and Robert Serpell are still pursuing several ways to get funds that could be used to support the remaining three candidates.

It was noted that the Maseno meeting represented a significant advance over past regional workshops by incorporating some of the norms advanced by the work of Suman Verma and Catherine Cooper and also by attracting and providing capacity building to many young scholars (more than 100 from almost all parts of Africa were in attendance). Furthermore, it was the sincere hope of the workshop organizers that the capacity building process that began in Kenya, will continue to engage more mentors and consequently provide capacity building to many deserving early career researchers in Africa and perhaps to the rest of the world. It was also agreed during the workshop that authors of the best posters during the Maseno workshop be given a chance to present their projects during the Lusaka conference. They were encouraged to submit their applications online before 31st of December, 2009. Each mentor and supervisor of mentored students will be expected to provide a short evaluation report on the progress of the student during the Zambian workshop.

The ISSBD Business Meeting held, on December 2nd 2009 and chaired by Bame Nsamenang resolved as follows:

1. We deeply appreciate ISSBD's continuing support of Africa Region Workshops in particular and discipline development and Young Scholar Training programs on the continent in general.
2. The host country of the 9th Africa Region ISSBD Workshop shall be Nigeria. The host Institution shall be the University of Lagos, Nigeria. The contact person who should submit a proposal to the ISSBD President as soon as possible for consideration by the ISSBD Executive Council (EC) is Prof. Esther Akinsola, a developmental psychologist.
3. In order to sustain ISSBD membership, each country that has hosted the ISSBD regional workshop should elect and propose a representative for this purpose to the ISSBD EC.
4. The meeting also set a new membership target of not less than 100 from the African continent by the end of 2009. The membership dues collected during the workshop were to be submitted to Sage publishers.
5. Continental Africa comprises 53 independent countries and it was felt that it is quite unfair if it is not represented in ISSBD EC by at least one member.

After the end of the workshop, some members visited Kisumu city and the world renowned Maasai Mara National Park on their way home.

News

ISSBD 2010: Zambian Site Visits of Interest to Developmental Scientists

Anitha Menon
University of Zambia, Lusaka, Zambia
E-mail: robertserpell@gmail.com

and

Robert Serpell
University of Zambia, Lusaka, Zambia
E-mail: robertserpell@gmail.com

Many of you attending the 21st Biennial congress in Lusaka, Zambia will be visiting Africa for the first time. Some of you will want to take this opportunity visit tourist attractions, such as the famous Victoria Falls (also known as *Mosi-o-tunya* – "The smoke that thunders") or one of the impressively stocked game parks which are easily accessible by road or by air. Pre- and post-conference tourism packages are advertised on the conference website, which also has links to relevant other sites and contact information for making advance bookings for group or individually packaged tours.

In addition to tourism, the Local Organising Committee is also making arrangements in the days immediately before and after the congress, for guided visits to various Zambian sites of special interest to developmental scientists, including low-income rural and urban residential neighborhoods, service providing institutions such as health, education and care facilities for children, persons with special needs and the elderly, and research facilities. Below are some of the highlights of these sites:

ChildFund Zambia

For those of you who are interested in research and services that relate to children, there are many non-Governmental organizations involved in work with children and young people. One such organization is ChildFund Zambia that has sites all over Zambia and has offered to host groups of ISSBD participants at their sites in and near Lusaka. ChildFund Zambia through the Australian Partnership with African Communities (APAC) program seeks to enhance child participation in community, district and national level programs by creating platforms where children can discuss and suggest solutions to problems affecting them, thereby establishing bridges between children, development actors and policy makers.

The ChildFund Zambia sites to which visits will be available are:

Chainda Children's Committee. This group has a membership of about 20 children and is located in Chongwe, a small town about 30km East of Lusaka. The Committee has a resource center where children come together to meet and plan their work. The center is also involved in community sensitization on child abuse, children's rights, HIV/AIDS, drug abuse and childcare.

Mutamino Children and Youth Groups. These groups' activities are also located in Chongwe and include a curio production unit run by youths. Their membership is about 35, and their objectives are similar to those of the Chainda Committee described above.

Kafue Central Children's Committee. This site is located in Kafue about 45km South of Lusaka and has a membership of about 20 children. The Committee has a resource center where the activities are planned and executed. They provide peer support and identify cases of abuse and report these to authorities. A newsletter is published regularly as an advocacy tool.

Tithandizane Children's Committee. Also located in Kafue, this group provides support to peers, and community sensitization on child abuse, HIV/AIDS, and children's rights and responsibilities. They also hold advocacy meetings with community, district and national leaders to present issues of concern and use PSS tools to build resilience in children.

The Beit CURE Hospital

The development of services for children with disabilities in Zambia faces many challenges. The Beit CURE Hospital is part of the CURE International network of 10 specialty hospitals in the developing world whose mission is to transform the lives of disabled children and their families through medical and spiritual healing by building partnerships and advocating for these children. The hospital offers services in Orthopedics, Neurosurgery, Plastic reconstructive surgery and ENT. Since it was formed in February 2007, the hospital has attended to 5,713 children and performed 3,294 operations, free of charge to children of needy families, who would otherwise not have the possibility of benefiting from such interventions, opening up new prospects of autonomy and dignity.

The University of Zambia Clinic

This visit offers a brief insight into the health situation in Zambia. Located very near the conference center, this site also gives you an opportunity to come into the attractive University of Zambia campus. The University Clinic which is located at the Great East Road Campus of the University of Zambia was established in 1966 to provide health care for the University community.

Operating on a 24-hour basis, the Clinic offers general out-patient consultations and has a twelve bed facility for short stay admissions. The clinic provides Mother and Child Health (MCH) Services, including services for prevention of mother to child transmission (PMTCT) of HIV.

The Clinic also runs specialist clinics for Psychiatry, Tuberculosis and Ophthalmology. It has a laboratory to support the clinical services offered at the Clinic. In collaboration with the Centre for Infectious Diseases Research in Zambia (CIDRZ), the Clinic provides free counseling and testing for HIV as well as free antiretroviral therapy (ART) to the community.

The University of Zambia HIV and AIDS Response Programme

Another site located on the University of Zambia, Great East road campus hosts the University of Zambia HIV and AIDS Response Programme. The initiative for establishing this unit was born out of recognition of the significance of the impact of HIV and AIDS on the University community such as debilitating illness and death of students, members of staff and their families. The unit works closely with various student and staff groups to put together various awareness, prevention, treatment and care programs. These programs run throughout the year and those visiting the site may have an opportunity to be a part of one of these programs. Since the unit is located next to the University of Zambia clinic, it would be a good idea to combine your visit to these sites.

Other site visits are being negotiated and the information will soon be up on the congress website. If you would like to visit any of the sites, please get in touch with our Professional Conference Organiser (issbd@image.co.zm). Only limited places are available for each site, so please book early. Visiting these sites would involve paying for the cost of the transport, and we are only able to arrange group visits to the sites so as to keep the transport costs low. We hope many of you will make use of this opportunity to get an insight into developmental issues in Zambia.

Young Scholars' Corner

Jochebed Gayles
Penn State University USA
E-mail: jgg137@psu.edu

Greetings and welcome again to the Young Scholars' Corner! We are fast approaching the 21st Biennial meeting of the International Society for the Study of Behavioural Development to be held from July 18th to 22nd, 2010 in Lusaka, Zambia. I hope that you have already made travel plans and applied for travel awards through the ISSBD website. If you are planning to attend please book your flights as soon as possible. We are excited to update you on the wonderful Young Scholars' activities that have been taking place within ISSBD.

Young Scholar Presence at the 21st ISSBD Meeting in 2010

The Young Scholar committee is seeking to increase the presence of young and early career scientists at the biennial meetings. We will hold a young scholar meeting at the Biennial Meeting so please be sure to look for us in the schedule

Newsletter by the ChildFund youth group in Zambia

ChildFund sponsored Youth advocating against child abuse to a Government Minister

ChildFund Children's Committee

and during the conference. If you are interested in becoming a part of the Young Scholars' Committee or joining our community please contact myself (jgg137@psu.edu) or Zena Mello (zmello@uccs.edu) and ask to be added to the young scholar listserve. We are also on Facebook! If you are a Facebook member please see our page (http://www.facebook.com/group.php?gid=109450620290&ref=ts) and request to be a member. This is another opportunity to be kept up to date with young scholar activities and new information that could be useful for your professional development.

Although the deadline for funding awards has been closed, we still have a travel team set up to help with planning travel to the conference. If you have additional questions regarding travel or would like to find out about other scholars traveling to the conference please contact us. The travel team is listed below.

Travel Committee:
Bin-Bin Chen (China; b-bchen@hotmail.com)
Filomena Parada (Portugal; fparada@netcabo.pt)
Joche Gayles (USA; jgg137@psu.edu)
Jonathan Santo (Canada; jonathan.santo@gmail.com)
Dr. Zena Mello (USA; zmello@uccs.edu)
Fun Things to Do in Lusaka, Zambia:

Zambia is a young thriving nation that received its political independence in 1964. Lusaka is the capital of Zambia and the climate of this beautiful region is subtropical. Zambia is known for wonderful landscapes, including Victoria Falls (see photo on right). If you are interested in going to the Falls or taking part in any other excursions while you are in Zambia please contact Jackie Jere-Folotiya (jackie@folotiya.com).

New Methods and Measures Section in the In IJBD

The executive council for ISSBD has passed a motion to initiate a method and measure section in the journal. As a result every issue of the International Journal of Behavioral Development will have a section designated for innovative and cutting-edge papers on advances in developmental methodology. This is a great opportunity for young scholars interested in developmental methodology. The editor for this new section is **Brett Laursen** (laursen@fau.edu).

This section of the Bulletin is continually dedicated to matters concerning young scholars. If you have ideas for topics please do not hesitate to send me an email (Joche Gayles: jgg137@psu.edu).

In Memoriam: Xiaojia Ge

Run Jin
California State University, USA
E-mail: rjin@csustan.edu

Professor Xiaojia Ge from the Institute of Child Development at the University of Minnesota died peacefully in August of 2009 following a battle with lung cancer. He was known for his loyalty, warmth, sense of humor, and brilliant mind. Those of us who knew Ge well often joked that

it would take volumes to contain his biography, given his life story. Ge would simply smile and admit that it was even a mystery to him as to how he became a professor in child development.

Ge was born in Beijing, China, in 1954. He completed just 6 years of grade school before the Cultural Revolution halted his formal education. During the next eleven years, Ge spent a significant amount of time on his own, playing sports and working as a coal miner and a machinist. His life took a dramatic turn after he passed the university entrance exam once it had been re-instated. Ge was part of the first cohort after the Cultural Revolution to attend college. After earning a Bachelor's degree in History and a Master's degree in Sociology, he came to the United States with $50 in his pocket to earn a doctorate in sociology from Iowa State University in 1990. Ge then worked as a research associate at Iowa State's Center for Family Research before accepting a faculty position in Human Development at the University of California, Davis in 1995. In 2007, he accepted a professorship in the Institute of Child Development at the University of Minnesota.

Ge made seminal contributions to the study of child and adolescent development. His work was known for its interdisciplinary approach to the study of developmental psychopathology and puberty, the development of depression, and linkages between genetic factors and family dynamics. Ge collaborated with colleagues around the world, including scholars from the United States, China, the United Kingdom, and New Zealand.

Ge began conducting developmental research in China in 2005 when he was invited as a visiting professor at the Institute of Psychology, Chinese Academy of Science. He also served on the editorial board of *Acta Psychologica Sinica*, one of the most prestigious psychological journals in China. Ge helped to further behavioral genetic research in China by editing a special issue of *Acta Psycholigica Sinica* with articles written by well-known scholars in the field. Ge also initiated and directed a longitudinal study of twins with his colleagues at the Chinese Academy of Science in Beijing. This project is an ongoing effort to investigate the genetic and environmental influences on the development of emotion and behavior in childhood and adolescence. The collaboration has resulted in several important publications on adolescent anxiety, depression, and deviant behavior using a behavioral genetics framework.

Ge was a tireless promoter of the international research community. He was on the editorial board of the International Journal of Behavioral Development (the official journal of ISSBD). He was a founding member of the Asian Caucus within the Society for Research in Child Development (SRCD). In the early 1990s, he spearheaded the biennial gathering of Chinese developmental psychologists at SRCD, now formalized as the Chinese Group for Developmental Science in North America. At the most recent gathering in Denver, Colorado, he gave an inspirational speech on how to succeed in academe to more than 50 Chinese developmental psychologists from Mainland China, Hong Kong, Singapore, and North America. Ge's absence will be deeply felt at the next gathering.

Ge was always very approachable, despite his busy schedule. People who attended the biennial meetings of SRCD and SRA can probably recall that "Ge was everywhere." He was in the lobby, in the hallway, and in the meeting rooms. He was seemingly talking to everyone,

encouraging junior scholars, and making all feel energized and connected to the field. Ge made a concerted effort to make international students feel welcomed at these conferences. He made sure that everyone felt included as members of a wide community of developmental scholars.

Ge set a shining example as both a scholar and a person. He provided expert guidance in the application of sophisticated research methods in the service of advancing theoretical and empirical work. He also strongly emphasized the importance of making a scholarly contribution. All of Ge's former students remember his unwavering commitment to mentoring. It was clear that his students were a top priority and he was unselfish with his time and energy. It was common for him to return comments on drafts of papers within 24 hours. His feedback was critical, but it was always aimed at facilitating improvement. Ge served as an excellent sounding board for ideas, and was concerned about his students' professional development and their personal well-being. Indeed, he encouraged students to focus on their long-term career goals, and not to be defeated by temporary setbacks and frustrations. He emphasized the importance of diligence and perseverance. Ge also offered insights into coping with personal struggles and provided guidance for making the right decisions. As evidence of his successful mentorship, many of his students now hold tenure-track positions.

Despite his incredible intellect and noted accomplishments, Ge was incredibly humble and respectful of everyone. Ge treated students as his friends and hosted extravagant meals at his home. These were fun times marked by good conservation and the occasional bottle of wine.

Upon Ge's passing, his former colleagues, students, friends, employees, and collaborators recounted their memories of him. These individuals hailed from all over the world. The tributes reflect both Ge's character and his widespread influence. His colleague, Ron Dahl, stated that Maya Angelou once said: "I've learned that people will forget what you said, people will forget what you did, but people will never forget how you made them feel." This sentiment is certainly true for all those who had the privilege to interact with Ge. He made everyone feel important and valued. His brilliance as a scholar was also captured well by comments from his colleagues Tytti Solantaus and Raijia-Leena Punamaki of Finland. They wrote "We had the joy of knowing professor Dr. Xiaojia Ge, as our joint interest in the impact of economic recession on children brought us together. Our first meeting is one of our most cherished memories from scientific conferences. Without knowing who shared our breakfast table, we started talking with him about the economic recession in Finland and our problems in handling our data. Finally we ended up drawing models together with him on the paper napkins – the only paper at hand. We Finns then carefully folded the napkins, brought them home and published our paper in Developmental Psychology."

Ge's passing is a tremendous loss to all those who were fortunate enough to have known him. He was a caring mentor, a wonderful colleague, a cherished friend, and a visionary contributor of the scholarly community. Ge was an extraordinary person and is greatly missed.

MAJOR CONFERENCES OF INTEREST

2010 July 11–16
International Congress of Applied Psychology
Location: Melbourne, Australia
Website: www.icap2010.com

2010 July 18–22
21th Biennial Meeting of the International Society for the Study of Behavioural Development
Location: Lusaka, Zambia
Website: www.issbd.org

2011 March 31–April 2
Biennial Meeting of the Society for Research in Child Development (SRCD)
Location: Montreal, Quebec, Canada
Website: www.srcd.org

2011 August 23–27
15th European Conference on Developmental Psychology (ECDP)
Location: Bergen, Norway
Website: www.ecdp2011.com