Journal of Person-Oriented Research

2016, 2(1–2)
Published by the *Scandinavian Society for Person-Oriented Research*
Freely available at http://www.person-research.org
DOI: 10.17505/jpor.2016.04

Directional Dependence in the Analysis of Single Subjects

Wolfgang Wiedermann[1], Alexander von Eye[2]

[1] University of Missouri, Columbia
[2] Michigan State University

Contact
wiedermannw@missouri.edu

How to cite this article
Wiedermann, W. & von Eye, A. (2016). Directional Dependence in the Analysis of Single Subjects. *Journal of Person-Oriented Research*, *2*(1–2), 20–33. DOI: 10.17505/jpor.2016.04

Abstract: Many statistical methods applied in person-oriented research make use of theoretical principles originally derived in a variable-oriented context. From this perspective, it naturally follows that advances originated in variable-oriented methodology may potentially contribute to the development of methods suitable for person-oriented perspectives. Direction Dependence Analysis (DDA) constitutes one of these recent advances and provides a framework to statistically evaluate asymmetric properties of observed variable relations. These asymmetric properties enable researchers to make statements whether a model of the form $x \rightarrow y$ or a model assuming $y \rightarrow x$ is more likely to approximate the underlying data-generating process in non-experimental settings. The present article introduces DDA to the context of person-oriented research and extends the DDA principle to (linear) vector autoregressive models (VAR) which can be used to describe individual development. We show that DDA can be used to empirically evaluate directional theories of (potentially multivariate) intraindividual development (e.g., which of two longitudinally observed variables is more likely to be the explanatory variable and which one is more likely to reflect the outcome). An illustrative example is provided from a study on the development of experienced mood and alcohol consumption behavior. It is demonstrated that VAR-DDA resolves the issue of identifying the direction of contemporaneous effects in longitudinal data. Temporality issues of directional theories used to explain intraindividual development, guidelines to achieve acceptable power, methodological requirements, and potential further extensions of DDA for person-oriented research are discussed.

Keywords: Direction dependence, vector autoregressive model, single-subject data, intensive longitudinal data, non-normality

Introduction

The person-oriented approach – deeply rooted in the holistic-interactionistic research paradigm (see, e.g., Magnusson, 2001) – emphasizes the individual as a functioning entity and focuses on individual characteristics of persons, their dynamic development over time, and their variation across contexts. Modern conceptualizations of the person-oriented approach can be summarized by seven tenets which acknowledge 1) individual specificity of development, 2) process complexity, 3) interindividual differences in intraindividual change, 4) the characterization of processes in terms of patterns, 5) holism, 6) pattern parsimony, and 7) dimensional identity (cf. Bergman & Magnusson, 1997; Bergman, Magnusson, & El Khouri, 2003; von Eye & Bergman, 2003; von Eye, Bergman, & Hsieh, 2015). This person-oriented perspective is opposed to a variable-oriented approach which focuses on variation of characteristics across individuals to establish statements about the correlational structure of variables, their constancy and change over time, and systematic differences across contexts on a population level.

Various statistical methods have been identified as being well-suited to study person-oriented research questions (cf. Bergman & Magnusson, 1997; Sterba & Bauer, 2010; von Eye et al., 2015). These include (among others) cross-sectional mixture models, latent growth mixture models, item response theory models, configural frequency and log-linear models, hierarchical linear models, and methods for the analysis of single subjects (such as time series and dynamic factor models). One important feature of many person-oriented statistical methods is that their underlying theoretical principles have originally been derived in variable-oriented contexts. For example, all statistical methods listed above have in common that they essentially rely on the generalized linear model (GLM; McCullagh & Nelder, 1989). The development of the linear regression model itself was, however, driven by the idea of establishing lawful statements about characteristics of defined *populations*. For example, the first regression lines ever drawn can be traced back to Sir Francis Galton while studying heredity in sweet peas and in humans (Galton, 1886; Hanley, 2004). Here, the systematic analysis of variation between individuals led to the statistical conceptualization of the linear regression model and to the first insights into basic principles of genetics. From this perspective, one may conclude that, although variable- and person-oriented approaches tend be seen as being complementary in nature, methodological advances in one of the two domains may also lead to advances in the other domain. More specifically, we propose that advances in variable-oriented methods, such as the linear regression model, contribute to the development of the person-oriented methodology.

In the present article, we focus on recent advances in the linear model which concern asymmetric properties of variable relations (cf. Dodge & Rousson, 2000, 2001; Dodge & Yadegari, 2010; Sungur, 2005; Wiedermann, Hagmann, & Eye, 2015; Wiedermann & von Eye, 2015a, 2015c). These asymmetric properties, which have recently been summarized as Direction Dependence Analysis (DDA; Wiedermann & von Eye, 2015a), allow empirical statements concerning the status of a variable as being either the cause of variation or the outcome of a process even when data were obtained in purely observational settings. In other words, DDA evaluates one key element necessary to establish causal statements, i.e., the directionality of effects.

Systematic variable manipulation within a randomized controlled research design is deemed as being the gold standard to establish causality. However, experimental approaches to the study of causation may often be questionable within in a person-oriented research paradigm (Bergman & Lundh, 2015), because manipulating one “target component” of a complex dynamic network may unintendedly manipulate other interconnected components which hampers statements about the unique contribution of the “target component” on an outcome. Here, correlational analyses are identified to better capture the complexity of dynamic developmental phenomena (e.g., Geldhof et al., 2014). In the person-oriented domain, Bergman (2009) proposes to replace the standard conceptualization of causality (rooted in the experimental paradigm) with the notion of *individual causality* (rooted in Russell’s, 1913, philosophical account of causation) which focuses on functional relations of phenomena. Thus, DDA may constitute a valuable tool to arrive at directional statements concerning intraindividual development and the functional relation of variables observed over time.

The present article is structured as follows: We start with introducing the statistical underpinnings of the direction dependence principle within a variable-oriented context. Then, we briefly review issues of competing directional theories in person-oriented research. More specifically, we focus on intensive longitudinal data settings which are commonly applied to derive conclusions about (multivariate) change processes for single subjects (Walls & Schafer, 2006). We then present extensions of DDA to (linear) vector autoregressive models (VAR) and demonstrate that these extensions can be used to evaluate which of two longitudinally observed variables is more likely to be the explanatory variable (the cause) and which one is more likely to be the outcome. An empirical example is given from a study on the development of alcohol consumption behavior of self-identified alcoholics. The article closes with discussing conceptual and methodological requirements of the proposed direction dependence approach together with an outline of future extensions relevant for person-oriented research.

Direction Dependence in Linear Regression: The Variable-Oriented Perspective

It is well-known that standard correlational and linear regression techniques are of limited use for answering questions concerning the direction of observed effects in observational studies. Given that one observes a statistically meaningful association between two variables, x and y, one has to consider four possible explanations: 1) a directional relation of the form $x \rightarrow y$, 2) a directional relation of the form $y \rightarrow x$, 3) a reciprocal relation $x \leftrightarrow y$, and 4) a spurious association due to a (potentially unobserved) third variable. Both, the Pearson correlation coefficient (as a measure of the linear association between two variables) and the ordinary least square (OLS) regression slope, do not carry any information to make empirically grounded decisions about which one of the four possible explanations holds for the underlying data-generating process (cf. von Eye & DeShon, 2012). Decisions on the nature of variable associations must be based on a priori theory. These methodological limitations can be explained by the fact that conventional correlation and regression approaches do only consider variation up to the second order moments of variables (i.e., variances and covariances). The key element of the direction dependence principle is to consider information beyond second order moments (i.e., skewness and kurtosis) to gain deeper insight in the underlying mechanism that generates the observed variable association. In other words, DDA requires and makes use of non-normality of variables to derive statements about the direction of effect.

In the following section, we review asymmetric properties of the ordinary linear regression model which emerge from non-normality of observed data. These asymmetric properties, which constitute the main components of DDA, concern 1) distributional characteristics of observed variables (x and y), 2) distributional characteristics of the error terms obtained from competing regression models ($x \rightarrow y$ versus $y \rightarrow x$), and 3) independence properties of error terms and predictors of competing models. Table 1 summarizes these DDA components together with corresponding significance tests and decision guidelines. For simplicity, we restrict the presentation of DDA components to the bivariate case and refer to multiple variable extensions whenever possible.

Distributional Properties of Observed Variables

Asymmetric properties of observed variable distributions directly emerge from the additive nature of the linear regression model, i.e., an outcome variable is defined as the sum of two elements, a (non-normally distributed) explanatory variable and a (normally distributed) error term. Let

$$y = b_{yx}x + \varepsilon_{yx} \tag{1}$$

be the true model describing the underlying data-generating process (for simplicity, but without loss of generality, we assume that the intercept is fixed at zero). Here, x is assumed to be the cause of y with b_{yx} being the OLS regression slope and ε_{yx} describing the error component which is assumed to be normally distributed, serially independent, and independent of the predictor x. Further, let

$$x = b_{xy}y + \varepsilon_{xy} \tag{2}$$

constitute the mis-specified model, i.e., the model that erroneously treats y as the cause of variation in x.

Dodge and Rousson (2000, 2001) as well as Dodge and Yadegari (2010) showed that the Pearson correlation $\rho_{xy} = cov(x,y)/(\sigma_x\sigma_y)$ – with $cov(x,y)$ being the covariance and σ_x and σ_y denoting the standard deviations of x and y – has asymmetric properties when considering higher moments of x and y. Specifically, these authors show that the cube of the Pearson correlation can be expressed as the ratio of third moments of outcome and predictor (i.e., the skewness of outcome, γ_y, and the skewness of predictor, γ_x),

$$\rho_{xy}^3 = \frac{\gamma_y}{\gamma_x}. \tag{3}$$

Similarly, the fourth power of the Pearson correlation can be written as the ratio of the fourth moments of outcome and predictor (i.e., the excess kurtosis of y, κ_y, and the excess kurtosis of x, κ_x),

$$\rho_{xy}^4 = \frac{\kappa_y}{\kappa_x}. \tag{4}$$

Because the Pearson correlation is bounded on the interval -1 and 1, it follows that the (absolute) skewness and excess kurtosis[1] of the response variable y will always be smaller than the (absolute) skewness and excess kurtosis of the explanatory variable x. In other words, given that model (1) is capable of describing the data-generating process, the outcome variable y will be closer to the normal distribution than a non-normal predictor x. This fundamental distributional property opens the door to evaluate the directional plausibility of a regression model through evaluating skewness and excess kurtosis of a tentative outcome and a tentative predictor. Note that this first DDA component is currently restricted to simple (bivariate) linear regression settings which may hamper application in practice. The following two components, however, can be used to overcome this limitation.

Distributional Properties of Error Terms

Again, let model (1) be the true model and let model (2) be the directionally mis-specified model. The second DDA component focuses on the distributional shape of the error terms associated with the two competing models, ε_{yx} and ε_{xy}. Wiedermann, Hagmann, Kossmeier, and von Eye (2013), Wiedermann et al. (2015) as well as Wiedermann (2015) showed that the skewness and the excess kurtosis of the error term obtained from the mis-specified model (ε_{xy}) can be written as

$$\gamma_{\varepsilon_{xy}} = (1-\rho_{xy}^2)^{3/2}\gamma_x \tag{5}$$

and

$$\kappa_{\varepsilon_{xy}} = (1-\rho_{xy}^2)^2\kappa_x \tag{6}$$

In other words, both, the third and the fourth moments of the false error term, can be expressed as functions of the third and fourth moments of the true predictor. From Equations (5) and (6) one can conclude that the skewness and excess kurtosis of ε_{xy} systematically increase with the skewness and excess kurtosis of x. This relation has the natural interpretation that the amount of non-normality of x is reflected in the "false" error term whenever the true predictor is erroneously used as the outcome variable. Because the true error term associated with the true model, ε_{yx}, is assumed to follow a normal distribution ($\gamma_{\varepsilon_{yx}} = \kappa_{\varepsilon_{xy}} = 0$), systematic differences in higher moments of ε_{yx} and ε_{xy} may, again, inform researchers about the plausibility of the model in terms of directionality of an effect. Note that this DDA component can straightforwardly be extended to multiple-variable settings (see Wiedermann & von Eye, 2015b, 2015c), i.e., to multiple linear regression models to draw conclusions about the directionality of effects in variable pairs (x, y) while adjusting for potential covariates (z_j, $j = 1,\ldots,k$).

[1]Note that γ, κ, and ρ refer to the skewness, excess kurtosis, and the Pearson correlation coefficient on the *population level*. Thus, while Equations (3) and (4) will exactly hold in the population, one cannot expect the relations to hold exactly based on sample estimates. However, given that $y = b_{yx}x + \varepsilon_{yx}$ constitutes the true underlying mechanism, γ_y/γ_x and κ_y/κ_x approximate the third and the fourth power of the correlation coefficient with increasing sample size.

Table 1. Summary of DDA components.

Criteria	True Model: $y = b_{yx}x + \varepsilon_{yx}$	False Model: $x = b_{xy}y + \varepsilon_{xy}$	Significance Tests	Decision Guidelines
Distribution of observed variables	The true outcome will always be closer to the normal distribution than the true predictor.		Separately evaluating the normality of x and y using skewness-, kurtosis-, and omnibus normality tests (cf. von Eye & DeShon, 2012). Bootstrap confidence intervals for the difference in higher moments have been proposed by Pornprasertmanit and Little (2012).	If y is closer to the normal distribution than x then $x \rightarrow y$ is more likely to reflect the true data-generating process.
Normality of error terms	The error term (ε_{yx}) is assumed to be normally distributed (i.e., zero skewness and zero excess-kurtosis).	The error term (ε_{xy}) is a function of the true predictor. The skewness and excess-kurtosis of ε_{xy} increase with non-normality of the true predictor x.	Separately applying skewness-, kurtosis-, and omnibus normality tests. Asymptotic skewness-, and kurtosis-difference tests were proposed by Wiedermann et al. (2015). Bootstrap difference tests are discussed by Wiedermann and von Eye (2015c).	If ε_{yx} is closer to a normal distribution than ε_{xy} then $x \rightarrow y$ is more likely to reflect the true data-generating process.
Independence of predictor and error term	The error term is assumed to be independent of the predictor.	The error term and the predictor will be non-independent whenever the true predictor (x) deviates from normality. Note that heteroscedasticity constitutes a special case of non-independence.	Separately applying non-linear correlation tests (cf. Wiedermann & von Eye, 2015a). Gretton et al. (2008) proposed the HSIC-test which detects any form of dependence in the large sample limit. Wiedermann, Artner, and von Eye (2016) discuss the application of homoscedasticity tests.	If ε_{yx} and x are independent and ε_{xy} and y are dependent then $x \rightarrow y$ is more likely to reflect the true data-generating process.

Independence of Predictors and Error Term

The third DDA component focuses on the independence assumption of predictors and the corresponding error term. In essence, the independence assumption implies that the magnitude of error made in predicting scores of the outcome variable does not depend on the values of the predictors. One essential feature of OLS estimation is that the estimated regression residuals will be uncorrelated with the predictors used in the model. It is important to note that uncorrelatedness will hold regardless of correctness of hypothesized path directions, i.e., both, the Pearson correlation of x and ε_{yx} and the Pearson correlation of y and ε_{xy} will be zero by definition. However, uncorrelatedness and (stochastic) independence are, in fact, two different concepts. Uncorrelatedness implies independence when all considered variables are normally distributed. However, in the non-normal case, uncorrelatedness does not necessarily imply independence. It can be shown that the independence assumption will be violated whenever the true predictor is erroneously used as the outcome of a regression model. Because independence is assumed in the true model, decisions concerning the direction of effect are possible through separately evaluating the independence properties of competing regression models (cf. Wiedermann & von Eye, 2015a). In the present article, we focus on extending this asymmetric property of regression models to methods suited to answer research questions in the person-oriented context. Technical details together with a discussion of how to statistically evaluate stochastic independence will be given below.

Competing Directional Theories in Person-Oriented Research

In the following discussion, we take an intensive longitudinal data perspective, i.e., sufficiently frequently repeated measurements which allow conclusions about the separate developmental processes of single subjects, dyads, or pre-defined groups (Walls & Schafer, 2006; Bolger & Laurenceau, 2013). Intensive longitudinal designs are typically applied to characterize change processes of subjects in their natural settings. As Bolger and Laurenceau (2013) note, “By characterize we mean not only functional form of change but also its causes and consequences" (p. 1). In other words, the intensive longitudinal data perspective is ideally-suited to empirically evaluate causal theories about complex (multi-variable) developmental processes. However, many processes do not lend themselves to experimentation as the gold-standard for establishing causation (e.g., for ethical reasons). Further, even if randomization is feasible, one still has to deal with the drawback that laboratory effects may not necessarily translate to real world settings. This situation clearly calls for methods that allow statistical statements concerning the directionality of effects, i.e., a falsification-procedure for directional developmental theories.

Person- and variable-oriented approaches may share that often competing directional theories about phenomena of interest exist. While directional theories in the variable-oriented setting commonly concern mechanisms assumed to hold on a population level, person-oriented directional theories address variable relations and their complex development within the same subject. For example, various studies analyzed the relation (and causal ordering) of alcohol consumption and intimate partner violence (IPV). From a variable-oriented perspective, it has been shown that IPV is indeed linked to alcohol consumption (see, e.g., Luthra & Gidycz, 2006; Williams & Smith, 1994). A person-oriented perspective is, for example, taken by Katerndahl, Burge, Ferrer, Becho, and Wood (2010). These authors used data of 16 women who were victims of domestic violence. Study participants provided daily ratings on type and severity of violence together with estimates of the husband's daily alcohol intake for 60 consecutive days. When analyzing the causal ordering of the two variables (alcohol intake and IPV) the following four possible (conceptual) models may be entertained: 1) alcohol intake causes subsequent IPV, 2) IPV causes subsequent alcohol intake, 3) alcohol intake and IPV are related in feedback loops, and 4) IPV and AC are both caused by another, third factor (e.g., dysfunctional social interaction and relationship skills; cf. Downs, Smyth, & Miller, 1996). Acknowledging that developmental processes may be specific to the individual (as reflected in the first tenet of the person-oriented approach; Bergman & Magnusson, 1997), we can expect that each of the four models will hold true in at least a subset of individuals. In other words, individual specificity is, of course, not restricted to idiosyncrasies in developmental trajectories of single characteristics. Individual specificity also implies variations in how several characteristics are causally related to each other. Further, causal relations within the same person may change over the life course. Returning to the example of IPV and alcohol consumption behavior, it may be true for some individuals that alcohol intake is responsible for subsequent IPV. However, this may be followed by a later developmental stage were alcohol consumption of couples is used as a self-medicated coping strategy. Thus, both models, alcohol → IPV and IPV → alcohol, may hold true depending on the developmental stage under consideration. In the following section, we propose extensions of DDA (originally proposed in the variable-oriented setting) to test competing directional theories of multiple (longitudinally observed) variables.

Directional Dependence in Vector Autoregressive Models

An important difference between variable- and person-oriented analyses concerns the definition of the *study sample*. While variable-oriented analyses typically focus on generalizations of model parameters obtained from a sample of n participants to an a priori defined population, the person-oriented analysis considered here focuses on generalizations of an individual model. Measurement occasions of the same subject take the place of number of participants. Thus, generalizations concern the time dimension.

VAR modeling (e.g., Lütkepohl, 2007; Rovine & Lo, 2012) constitutes a straightforward approach to introduce DDA principles in the person-oriented domain. Let x_t and y_t $(t = 1, \ldots, T)$ denote two stationary series of scores repeatedly obtained from the same individual. Further, assume x scores contribute to the development of y, i.e., x is the (longitudinally observed) cause and y represents the (longitudinally observed) outcome. Figure 1a shows the conceptual path diagram of a (first-order) VAR model underlying x and y for three measurement occasions (note that the proposed approach is also valid for higher-order VARs, however, for notational simplicity, we restrict the presentation to first-order VARs).

Here two different types of model parameters can be distinguished, (first-order) autoregressive effects and cross-lagged effects. Autoregressive effects describe the expected change from one occasion $t-1$ to the subsequent occasion t within the same variable (e.g., capturing the effect of past events of alcohol consumption on present and future drinking behavior). The parameter a_{xt-1} represents the autoregressive effect of regressing x_t on x_{t-1} and a_{yt-1} denotes the autoregressive effect of regressing y_t on y_{t-1}. Cross-lagged parameters (b_{yt} and b_{yt-1}) describe the contribution of the time series x in generating the time series y, i.e., the causal effects of x on y (e.g., the influence of alcohol consumption on IPV). Note that two different types of causal effect are considered in the present model, first-order lagged effects which represent the influence of x_{t-1} on y_t and zero-order lagged effects (known as *contemporaneous effects*) describing the influence of x_t on y_t (reasons for including contemporaneous effects will be discussed in detail below). The equations for this model can be written as

$$
\begin{aligned}
y_t &= a_{yt-1}y_{t-1} + b_{yt}x_t + b_{yt-1}x_{t-1} + \varepsilon_{yt} \\
x_t &= a_{xt-1}x_{t-1} + \varepsilon_{xt}
\end{aligned}
\quad (7)
$$

where ε_{yt} and ε_{xt} denote the error terms of the model which are assumed to be serially uncorrelated, independent of the corresponding predictors, and independent of each other. The conceptualization of error components in time series models differs slightly from the definition of error terms in ordinary cross-sectional regression models. In a cross-sectional linear model such as the model given

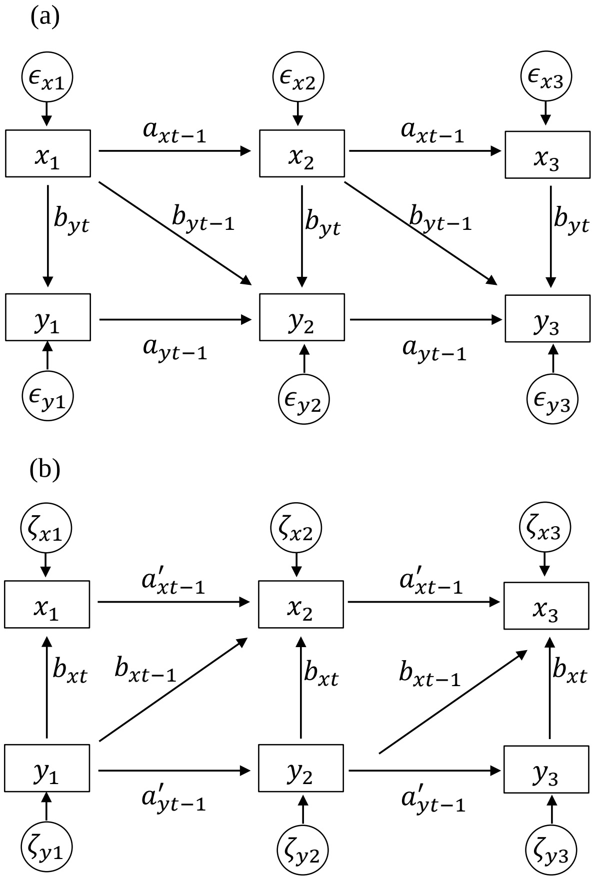

Figure 1. Path diagrams of competing VAR models.

in Equation (1), the error component usually captures measurement imprecision and variables outside the model. In contrast, in time series modeling the error component (sometimes referred to as "innovations"; cf. Lütkepohl, 2007) represents new information at a given measurement occasion t which influences the future development of a process. Thus, to evaluate hypotheses compatible with direction dependence, we assume that the error components in (7) are non-normally distributed.

Figure 1b shows the path diagram for the mis-specified model. Here, again, x is erroneously treated as the (repeatedly observed) outcome and y serves as the (repeatedly observed) cause of variation. This model takes the form

$$\begin{aligned} x_t &= a'_{xt-1}x_{t-1} + b_{xt}y_t + b_{xt-1}y_{t-1} + \zeta_{xt} \\ y_t &= a'_{yt-1}y_{t-1} + \zeta_{yt}, \end{aligned} \tag{8}$$

with a'_{xt-1} and a'_{yt-1} are the first-order autoregressive effects, b_{xt} and b_{xt-1} describe the contemporaneous and the first-order lagged effects of y on x, and ζ_{xt} and ζ_{yt} denote the error terms of the mis-specified model. In the following section, we show that DDA can be used to identify the correct VAR model (of course, assuming that it exists). We focus on the third DDA component which concerns the independence of predictors and error terms.

Independence Properties of Predictors and Errors in VAR Models

As discussed above, the decision of selecting between two directionally competing models can be based on separately evaluating whether the independence assumption is violated. Assuming that the assumption of independent predictor and error components is fulfilled in the "true" VAR model (e.g., $x \rightarrow y$), we now show that the independence assumption will systematically be violated in a VAR model which erroneously considers a reversed causal flow ($y \rightarrow x$).

The key element which allows to identify directional model-misspecifications is the so-called Darmois-Skitovich theorem (Darmois, 1953; Skitovich, 1953). These authors analyzed consequences of stochastic independence of linear functions of variables. They showed that if two stochastically independent variables exist that are defined as the weighted sum of the same set of random variables (z), i.e.,

$$\begin{aligned} u &= \lambda_1 z_1 + \lambda_2 z_2 + \ldots + \lambda_k z_k \\ v &= \xi_1 z_1 + \xi_2 z_2 + \ldots + \xi_k z_k \end{aligned} \tag{9}$$

with λ and ξ being non-zero values, then all variables z for which $\lambda\xi \neq 0$ follow a normal distribution (note that Mamai, 1963, extended this result to the case of infinite sums). From this, it straightforwardly follows that when at least one *non-normal variable* z exists for which $\lambda\xi \neq 0$ is satisfied, u and v are *stochastically dependent*. Reconsidering the two competing VAR models given in Equations (7) and (8) reveals that the true outcome y_t [which is erroneously treated as the predictor in Equation (8)] and the false error term ζ_{xt} are both linear combinations of the *same* variables and, thus, fulfill the requirements to apply the Darmois-Skitovich theorem. This can be shown by, first, re-writing the "true" model (7) as

$$y_t = a_{yt-1}y_{t-1} + (b_{yt}a_{xt-1} + b_{yt-1})x_{t-1} + b_{yt}\varepsilon_{xt} + \varepsilon_{yt} \tag{10}$$

(which is obtained through inserting the autoregressive component of x_t into the equation describing the development of y_t) and, second, solving the mis-specified model given in (8) for the error term ζ_{xt} after inserting the true into the mis-specified model which results in

$$\zeta_{xt} = (1 - b_{xt}b_{yt})\varepsilon_{xt} - \theta_{xt-1}x_{t-1} - \theta_{yt-1}y_{t-1} - b_{xt}\varepsilon_{yt} \tag{11}$$

with $\theta_{xt-1} = (a_{xt-1} - a_{xt-1}b_{xt}b_{yt} - b_{xt}b_{yt-1} - a'_{xt-1})$ and $\theta_{yt-1} = (a_{yt-1}b_{xt} + b_{xt-1})$. From (10) and (11) we conclude that both, y_t and ζ_{xt}, are linear combinations of 1) past y values, y_{t-1}, 2) past x value, x_{t-1}, 3) the "true" error term ε_{xt}, i.e., innovations of x_t, and 4) the "true" error term ε_{yt}, i.e., innovations of y_t. Thus, it follows from the Darmois-Skitovich theorem that y_t and ζ_{xt} will be stochastically dependent if 1) the error component ε_{xt} is non-normally distributed and $b_{yt}(1 - b_{xt}b_{yt}) \neq 0$ and/or 2) the error component ε_{yt} is non-normal and $b_{xt} \neq 0$. Finally, assuming that the independence assumption holds for the true model, we arrive at the following simple guidelines to select between directionally competing models:

- If the independence assumption holds for x_t and ε_{yt} and, at the same time, the independence assumption is violated for y_t and ζ_{xt} then the model $x \rightarrow y$ is more likely to approximate the true data-generating process.
- If the independence assumption holds for y_t and ζ_{xt} and, at the same time, the independence assumption is violated for x_t and ε_{yt} then the model $y \rightarrow x$ is more likely to approximate the true data-generating process.
- If the independence assumption is violated in both models then unmeasured confounders may be present and no distinct decision on the direction of effect can be made. If the independence assumption is fulfilled in both models then no decision can be made. Note that this scenario points at violated distributional requirements and/or power issues due to small sample size.

Testing Independence

The selection procedure described above relies on testing independence of predictors and error terms. In practice, residuals will be used to approximate the unknown error components. Because regression residuals and the predictor will be uncorrelated by definition, evaluating independence of the two components becomes more complex. We focus on two promising approaches: 1) Non-linear correlation approaches and 2) a kernel-based test for statistical independence.

Non-Linear Correlation Approaches. These statistical approaches emerge from the fact that (stochastic) independence is a much stronger assumption than uncorrelatedness. While uncorrelatedness refers to the situation of zero-covariances of pairwise variables, independence implies that zero-covariances exist for any non-linear functions of the pairwise variables (cf. Hyvärinen, Karhunen, & Oja, 2001). In practice, this implies that correlation tests can be applied on (non-linearly) transformed variables to evaluate stochastic non-independence. In other words, instead of testing the null hypothesis of zero correlation of y_t and ζ_{xt} (which will be zero by definition), one evaluates whether zero-correlations hold for, e.g., $\{y_t^2, \zeta_{xt}^2\}$, $\{\exp(y_t), \sin(\zeta_{xt})\}$, $\{\tanh(y_t), \zeta_{xt}\}$, etc. Of course, testing all possible non-linear transformations is not feasible in practice, which raises the important question which non-linear transformation should be applied to make statements about potential non-independence of predictor and residuals. Wiedermann and von Eye (2015a, 2016) as well as Wiedermann et al. (2016) showed that squaring the variables, i.e., $\{y_t^2, \zeta_{xt}\}$ and $\{y_t, \zeta_{xt}^2\}$, is particularly useful to detect non-independencies because covariances of $\{y_t^2, \zeta_{xt}\}$ and $\{y_t, \zeta_{xt}^2\}$ increase with the skewness of the "true" predictor. Assuming that both estimated models (7) and (8) show adequate model fit (i.e., to be considered eligible as competing candidates, both models should fit the data well in terms of second order moments) and making use of the proposed decision guidelines results in the following model selection procedure:

- If the null hypotheses $H_0 : cor(x_t^2, \varepsilon_{yt}) = 0$ and $H_0 : cor(x_t, \varepsilon_{yt}^2) = 0$ are both retained and, at the same time, at least one of the null hypotheses $H_0 : cor(y_t^2, \zeta_{xt}) = 0$ and $H_0 : cor(y_t, \zeta_{xt}^2) = 0$ is rejected then the model $x \rightarrow y$ is more likely to approximate the true data-generating process.
- If the null hypotheses $H_0 : cor(y_t^2, \zeta_{xt}) = 0$ and $H_0 : cor(y_t, \zeta_{xt}^2) = 0$ are both retained and, at the same time, at least one of the null hypotheses $H_0 : cor(x_t^2, \varepsilon_{yt}) = 0$ and $H_0 : cor(x_t, \varepsilon_{yt}^2) = 0$ is rejected then the model $y \rightarrow x$ is more likely to approximate the true data-generating process.
- If null hypotheses of zero (nonlinear) correlations are rejected in both models then unmeasured confounders are likely to be present. Hence, no decision can be made.
- If null hypotheses of zero (nonlinear) correlations are retained in both models, no distinct decision can be made (this again points at low power or violated data requirements).

This independence approach has the advantage of computational simplicity. In fact, given that the test statistics rely on the ordinary Pearson correlation test, this approach is readily available in virtually all statistical software programs. However, applying this approach introduces additional Type II error risks (aside of Type II error due to small sample sizes), because retaining null hypotheses of zero correlation for selected non-linear functions does not guarantee that no other non-linear functions may exists for which non-independence hold. The following kernel-based test can be used to overcome these limitations.

Kernel-Based Independence Test. Gretton et al. (2008) introduced a kernel-based approach for testing independence, the so-called Hilbert-Schmidt Independence Criterion (HSIC). Instead of analyzing independence properties of random variables, the HSIC deals with testing the independence of functions of random variables. Thus, the HSIC approach is provably universal in detecting any dependence between two random variables. Due to space restrictions, we do not present technical details of the test statistic (for details see Gretton et al., 2008). This approach may have the disadvantage of high computational complexity. Further, the test is not readily available in standard statistical software programs (Matlab implementations of the test can be found at `http://www.gatsby.ucl.ac.uk/~gretton/indepTestFiles/indep.htm`). Again, making use of the proposed decision guidelines, we arrive at the following model selection procedure (of course, again, both candidate models should fit the data well in terms of second order moments):

- If the null hypothesis of the HSIC-test $H_0 : x_t \perp \varepsilon_{yt}$ (i.e., x_t and ε_{yt} are stochastically independent) is retained and, at the same time, at the null hypothesis $H_0 : y_t \perp \zeta_{xt}$ is rejected, then the model $x \rightarrow y$ is more likely to approximate the true data-generating process.
- If the null hypothesis $H_0 : y_t \perp \zeta_{xt}$ is retained and, at the same time, at the null hypothesis $H_0 : x_t \perp \varepsilon_{yt}$ is rejected, then the model $y \rightarrow x$ is more likely to approximate the true data-generating process.
- If both null hypotheses are rejected then unmeasured confounders are likely to be present.
- If both null hypotheses are retained, no distinct decision concerning the direction of effect can be made.

Empirical Example: Subjective Mood and Alcohol Consumption Behavior

In the following section, we illustrate how to apply the proposed methodology in practice. We use data from a study on the development of alcoholism of adults who self-identified as alcoholics (Perrine, Mundt, Searles, & Lester, 1995). Using automated interviews, daily alcohol consumption (i.e., number of alcoholic beverages), together with daily subjective mood, stress, and health ratings were obtained over a study period of three years. Various previous studies analyzed this dataset from a person-oriented perspective (von Eye & Bergman, 2003; von Eye & Mun, 2012; von Eye et al., 2015). For example, von Eye and Wiedermann (2016, this issue) use daily recorded alcohol consumption to illustrate configural models to answer whether there exist interindividual differences in intraindividual change. In the following application, we ask questions concerning the direction of effect of subjective mood and alcohol consumption behavior from a person-oriented perspective, i.e., whether individual changes in mood are the cause of changes in alcohol consumption (i.e., *mood* → *alcohol*) or whether alcohol consumption patterns cause changes in perceived mood (i.e., *alcohol* → *mood*) for a single subject. Specifically, we analyze time series data of one participant (respondent 3032) based on observed weekly aggregates. Note that both competing VARs may have theoretical support: The well-established "tension reduction hypothesis" (Conger, 1956; Young, Oei, & Knight, 1990), where (negative) mood is assumed to prompt alcohol use as a self-medication approach, poses that a model of the form *mood* → *alcohol* explains the observed development over time. In contrast, the "hedonic motive hypothesis" (Gendolla, 2000) states that alcohol may reinforce the mood which, in turn, supports a model of the form *alcohol* → *mood*.

Figure 2 (upper panel) shows the standardized scores of observed (weekly averaged) mood ratings together with the (weekly averaged) number of alcoholic beverages for 105 consecutive weeks. To remediate potential issues of non-stationarity of both time series, first-order differences between observations ($y'_t = y_t - y_{t-1}$ and $x'_t = x_t - x_{t-1}$) were computed. Thus, time series used in VAR-DDA reflect the observed change between each observation in the original series. Figure 2 (lower panel) shows the differenced series for perceived mood and alcohol intake. The KPSS-test (Kwiatkowski, Phillips, Schmidt, & Shin, 1992) indicated stationarity of both differenced series (both p's $> .1$).

Next, for both directionally competing models, *mood* → *alcohol* (Model I) and *alcohol* → *mood* (Model II), the Bayesian Information Criterion (BIC) was used to select proper lag length (cf. Lütkepohl, 1985). For both possible directions (*mood* → *alcohol* and *alcohol* → *mood*) a series of VAR model were estimated and lagged values up to the fifth order were considered for the corresponding predictor variables. For Model I, including first- and second-order lags for alcohol intake and first-order lags for mood ratings showed the lowest BIC. For Model II, the lowest BIC was observed when considering first-, second-, and third-order lags for subjective mood ratings and first-order lags for alcohol consumption (the Durbin-Watson test confirmed the absence of autocorrelation of disturbances for both models; both p's $> .7$). Before we can interpret the parameter estimates of the models, it is important to ensure that both candidate models describes the data well in terms of second moments. Model I showed a multiple R^2 of 0.45 [$F(4,97) = 19.86, p < .001$], for Model II we obtained a multiple R^2 of 0.38 [$F(5,95) = 11.75, p < .001$]. Further, the model goodness of fit test suggested that both models are able to describe the data well [Model I: $\chi^2(2) = 1.89, p = .389$; Model II: $\chi^2(3) = 3.60, p = .308$]. Figure 3 shows the path diagrams together with the estimated coefficients of both models. In general, interpretations of parameter estimates for both models are complementary in nature. For Model I, we observe that past increases in alcohol consumption lead to a decrease in present alcohol intake (the effects of past alcohol intake decrease with lag length) while present alcohol consumption increases with past and present mood (the impact of mood also decreases with lag length). For Model II, we obtain a complementary picture, i.e., past increases in mood lead to decreases in present mood while present mood increases with past and present alcohol consumption (again, all effects decrease with lag length).

In the next step, we ask questions concerning the direction of effect. Note that the presented direction dependence approach relies on the assumption of non-normally distributed error terms. Thus, we, first, computed the regression residuals of both models and checked whether we find empirical support for this distributional assumption in at least one of the two models. Residuals of Model I showed a skewness of 0.24 and an excess kurtosis of –0.54. The Kolmogorov-Smirnov test rejected the null hypothesis of normally distributed residuals ($D = 0.13, p = .047$). For residuals of Model II, we obtained a skewness of –1.09 and an excess kurtosis of 2.07. Here, the Kolmogorov-Smirnov test retained the null hypothesis of normality according to the 5% nominal significance level ($D = 0.12, p = .121$). However, overall, we can conclude that the residuals are sufficiently non-normal to proceed with testing direction dependency.

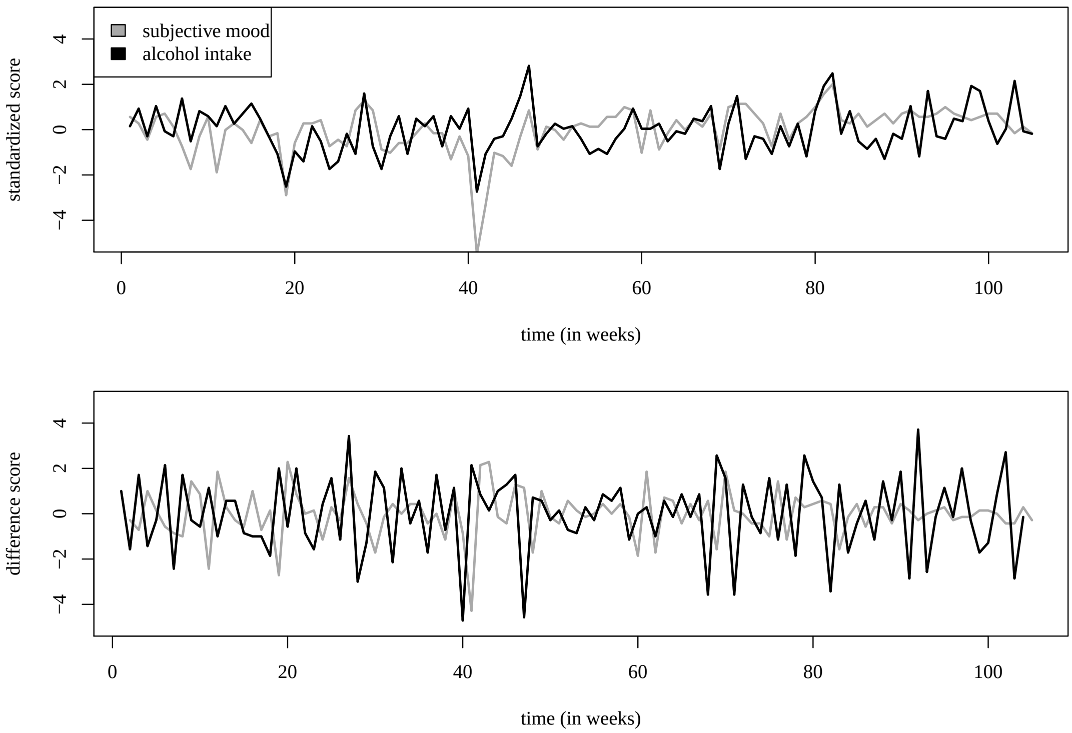

Figure 2. Observed time series for respondent 3032 (upper panel: Observed weekly averaged mood ratings and weekly averaged number of alcoholic beverages; lower panel: First-order differenced time series).

Three non-linear correlation tests using the squaring function and the HSIC-test were used to evaluate the independence assumption for both models. For Model I, all non-linear correlation tests were non-significant: $cor(mood_t, \varepsilon^2_{alc}) = -0.12, p = .214$; $cor(mood^2_t, \varepsilon_{alc}) = -0.13, p = .183$; $cor(mood^2_t, \varepsilon^2_{alc}) = -0.11, p = .294$. Further, the HSIC-test confirmed independence of predictor and error term (HSIC = 0.25, p = .650). In contrast, all non-linear correlation tests rejected the independence assumption for Model II: $cor(alc_t, \varepsilon^2_{mood}) = -0.22, p = .026$; $cor(alc^2_t, \varepsilon_{mood}) = -0.23, p = .022$; $cor(alc^2_t, \varepsilon^2_{mood}) = 0.40, p < .001$. The HSIC-test failed to reject the null hypothesis (HSIC = 0.44, p = 0.188). Applying the decision guidelines proposed above, we conclude that a model of the form (Model I) is more likely to describe the data-generating process based on non-linear correlation tests. The HSIC-test does not allow a distinct decision because null hypotheses for both models were retained. One reason for this may be the rather small sample size (i.e., short time series) which reduces the power of the test. However, we observe a larger HSIC value for Model II indicating a larger magnitude of non-independence compared to Model I (0.44 versus 0.25) which again, points toward the model *mood* → *alcohol*. Thus, overall, DDA suggests that (for respondent 3023) experienced mood is more likely to cause alcohol intake than vice versa which is in line with the "tension-reduction hypothesis".

Discussion

The present study introduced basic principles of DDA to person-oriented researchers and presented extensions of DDA to VAR models which can be used to study change in single subjects. The presented approach allows one to study directionality of functional relations observed over time and thus matches the notion of *individual causality* (cf. Bergman, 2009). In the following paragraphs, we start with briefly summarizing the results of extensive Monte-Carlo simulation experiments (assessing the Type I error and power performance of the proposed approach) to provide guidelines for practical applications. Further, it is important to reiterate that the VAR model considered here differs in two important aspects from "traditional" models: First, to apply direction dependence principles, we assume that the error terms of the model deviate from the normal distribution. Second, we consider contemporaneous effects (i.e., effects which occur at the same measurement occasion t) in addition to lagged effects of the two series. We discuss the plausibility of these requirements from both, practical and methodological perspectives. Finally, we close the article with outlining further applications of DDA in the person-oriented domain.

Type I Error and Power Considerations

We have conducted extensive Monte-Carlo simulation experiments to assess the Type I error and power performance of the proposed model selection procedure. Here, we summarize simulation results focusing on the length of the time

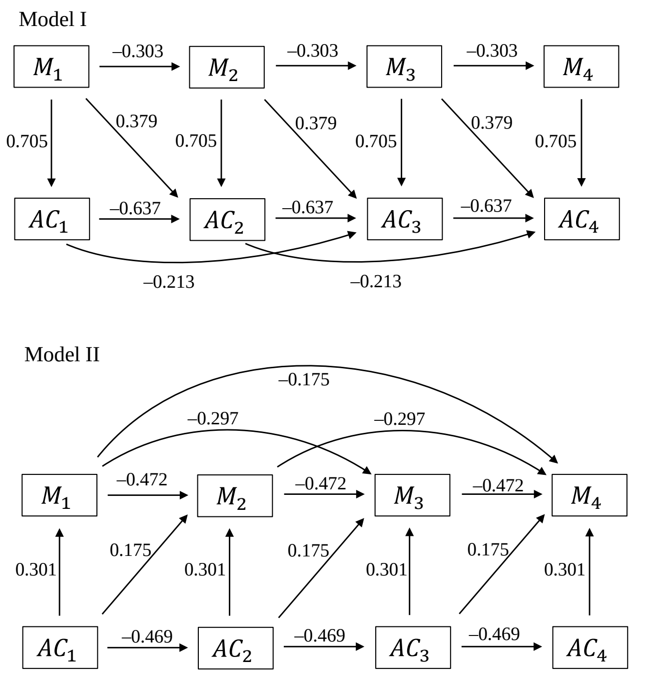

Figure 3. Results of competing VAR models for mood and alcohol consumption (all coefficients in both models are statistically significant using a nominal significance level of 5%; AC = alcohol consumption, M = perceived mood).

series, the degree of asymmetry of error term distributions, and effect sizes of autoregressive, lagged, and contemporaneous effects.

As expected, nonlinear correlation approaches of the form $cor(y_t^2, \zeta_{xt})$ and $cor(y_t, \zeta_{xt}^2)$ were able to protect the nominal significance level of 5% when error terms were randomly sampled from a normal distribution. Empirical Type I error rates of the HSIC were systematically lower than 5%. In general, empirical power to select the correct VAR model increases with the skewness of the error terms, the length of the time series, and the magnitude of the contemporaneous effect. In contrast, the power slightly decreases with the magnitude of autoregressive and cross-lagged effects. With respect to Cohen's (1988) definition of small, medium, and large effects, the following rough guidelines can be used in order to achieve a power of 80%: First, for large contemporaneous effects together with skewness values of the error terms larger than 1.5, $T \geq 100$ observations are necessary for medium autoregressive effects ($T \geq 50$ may be sufficient for small autoregressive effects). Second, for medium-sized contemporaneous effects, $T \geq 200$ observations are necessary to achieve a power of 80%. Third, the nonlinear correlation test of the form $cor(y_t, \zeta_{xt}^2)$ is more powerful than the HSIC-test. Correlation approaches of the form $cor(y_t^2, \zeta_{xt})$ showed the lowest power.

Non-Normal Errors

Various studies repeatedly demonstrated that the normality assumption is very likely to be violated in empirically observed data. For example, Micceri (1989) analyzed over 440 empirical datasets regarding their distributional properties and found that only 4.3% were reasonably normal while the Kolmogorov-Smirnov test indicated non-normality for all 440 distributions at a 1% significance level. Quite similar results were obtained by Blanca, Arnau, López-Montiel, Bono, and Bendayan (2013) who studied 693 empirically observed distributions (measures of cognitive ability and other psychological variables) and found that only 5.5% were approximately normally distributed. One theoretical explanation why variables (and error terms) are likely to deviate from the normal distribution is, for example, given by Beale and Mallows (1959). The error term of a model usually captures measurement error and unconsidered background variables. Now, assume that the error is a mixture of several unobserved independent and normally distributed variables (z_j). In this case, the resulting error term will show elevated excess kurtosis values whenever the variances of the unconsidered variables, z_j, differ from each other. Because unequal variances are likely to occur in practice, the error terms are also likely to deviate from normality (note that the same argument applies to the predictor variable of a model). From these results, we may conclude that distributional requirements of DDA are very likely to hold in practical applications.

However, the assumption of non-normal error terms deviates from the classic VAR set-up. In the classic VAR model, testing for non-normality of error terms is commonly recommended as a model-checking procedure (Lütkepohl, 2007) where non-normality of residuals may indicate that additional variables or lag structures improve the model. However, it is important to note that non-normality may still exist after including additional variables and/or additional higher order lags. The VAR-DDA approach presented here differs from the classical set-up in the sense that non-normality may not only be used to indicate potential model-misspecification of the predictor side of the model equation. Considering non-normality as a valuable source of information enables researchers to identify directional model-misspecifications.

Note that this is not the first study which makes use of higher than second moment information of time series data. For a discussion of non-normality in the time series domain see, for example, Granger and Newbold (1976), Swift and Janacek (1991) and Sim (1994). Further, Peters, Janzing, Gretton, and Schölkopf (2009) evaluate the reversibility of autoregressive moving average (ARMA) models through testing the independence of the error terms and preceding values of the time series. More recently, Hernández-Lobato, Morales-Mombiela, and Suárez (2011) showed that time-reversed residuals of linear ARMA models will always be closer to the normal distribution than non-normal true errors of the chronologically ordered time series (i.e., a so-called Gaussianization effect). Hyvärinen, Zhang, Shimizu, and Hoyer (2010) proposed an extension of the so-called linear non-Gaussian acyclic model (LiNGAM; cf. Shimizu, Hoyer, Hyvärinen, & Kerminen, 2006) – in essence, a causal discovery algorithm used to recover the underlying structure of directed acyclic graphs (DAG) – to VAR models. Conceptual differences of DDA and causal discovery algorithms are discussed by Wiedermann and von Eye (2015a).

Considering Contemporaneous Effects

Often, researchers are reluctant to consider contemporaneous effects in VAR models for reasons rooted in properties of the model resulting from second order moments of variables (variances and covariances). Consider, for example, so-called *Granger-causality testing* (Granger, 1969) which was originally introduced in the econometric sciences but also received considerable attention in various life sciences domains such as studying effective connectivity of neurons using fMRI data (see, e.g., Stephan & Roebroeck, 2012) or studying gene regulatory networks (Lozano, Abe, Liu, & Rosset, 2009). In essence, Granger causality testing constitutes a predictor error approach. Here, a time series x is said to "Granger-cause" another time series y, when additionally including past information of x in the prediction of y leads to a better model fit (i.e., a smaller predictor error) than predicting y from its own past alone. Interestingly, although temporal information may be seen as the key ingredient for causal claims (implicitly following Hume's proposition that the cause must precede the effect, Granger, 1969) also considered contemporaneous effects of x and y. However, contemporaneous effects are rarely considered in practice because the direction of contemporaneous effects cannot be derived from standard VARs (see, Hsiao, 1982; Lütkepohl, 2007). From this perspective, including higher order information to identifying the direction of effects, may also resolve an important issue in Granger-causality testing (see also Hyvärinen et al., 2010; von Eye & Wiedermann, 2015).

As noted above, the fact that contemporaneous effects tend to be ignored in time series modeling, may be deeply rooted in *sequential theories of causation* (Hume, 1777; Mill, 1851) where temporality is an essential requirement to establish causation. Of course, from a purely statistical perspective, sequential ordering of variables alone is known to be insufficient to establish causation, i.e., longitudinal information does not rule out spurious correlations over time (Yule, 1921, 1926; Link & Shrout, 1992). Further, considering that even many physical laws are, in fact, formulated in a time-reversible form[2] lends support to *theories of simultaneous causation* which focus on physical mechanisms underlying the process of interest and pose that temporally extended outcomes occur simultaneously with temporally extended causes (Huemer & Kovitz, 2003). Another argument which may further substantiate the inclusion of contemporaneous effects concerns the measurement of time series itself. While change over time is typically conceptualized as being incessant and continuous (Lerner, 1998), time series data (obtained to measure change) are collected using fixed measurement intervals (i.e., the amount of time that elapses between measurement occasions often is considered constant; for a methodological discussion of temporal designs see Timmons & Preacher, 2015). Reconsider the empirical example on subjective mood and alcohol intake: The respondent's mood, which may impact the alcohol consumption behavior, is (consciously or subconsciously) experienced in a continuous manner. Even if respondents are requested to provide information about their prior day's experiences and behaviors, not considering contemporaneous information implies that potential change that occurs in between two consecutive measurement occasions is considered as being unimportant for the process of modeling the relation of the two phenomena of interest. From this perspective, omitting potential contemporaneous effects may be seen as an unnecessary truncation of available information. Further, for the social sciences, it is important to note that concepts of contemporaneous effects may differ from standard conceptualizations of simultaneous causation typically used in philosophical accounts. While philosophical discussions consider contemporaneous causal effects as being truly immediate (i.e., there is no time difference between two causally related events), contemporaneous effects in the social sciences may be better understood as the result of measurement within a short period of time which is commonly assumed to be negligible for the quantification of change. However, even short periods of time may provide enough room for a temporal ordering of variables which, again, lends support to incorporating contemporaneous effects in the analysis of functional relations.

[2]For example, the well-known relation between force (f), mass (m), and acceleration (a), $f = ma$, may entice the view that force is caused by mass and acceleration. However, this fundamental relation does not necessarily imply a specific direction of time, i.e., using $a = f/m$ one may easily conclude that acceleration is caused by force and mass (cf. Huemer & Kovitz, 2003).

Potential Extensions of DDA in the Person-Oriented Domain

Because principles of DDA concern the linear model in its general definition as an additive error model various extensions in both, the variable- and person-oriented domain, are possible. In the variable-oriented setting, for example, extensions to mediation models have been discussed by Wiedermann and von Eye (2015b) and Wiedermann and von Eye (2016). Acknowledging that many person-oriented methods rely on the linear model implies that DDA may also be applied in other person-oriented settings which definitely warrants future research.

For example, von Eye et al. (2015) as well as Asendorpf (2015) discuss hierarchical linear modeling (HLM) as an important methodological tool compatible with the person-oriented perspective. In essence, HLM considers a hierarchically nested data structure (e.g., repeated measures nested within respondents) and allows one to examine level-specific variation (i.e., HLM explicitly models inter- and intraindividual differences). Because the level-specific error terms are assumed to be independent and normally distributed random variates, systematic violations of these assumptions may point at directional model-misspecifications.

Further, dynamic factor models (Molenaar, 1985; Wood, 2012) have been identified as ideally-suited to study the tenets of person-oriented research (cf., Molenaar, 2010). The dynamic factor model can be written as

$$y_t = \Lambda\eta_t + \varepsilon_t$$
$$\eta_t = B_1\eta_{t-1} + B_2\eta_{t-2} + \ldots + B_s\eta_{t-s} + \zeta_t, \quad (12)$$

with y_t being a p-variate manifest time series ($t = 1, \ldots, T$), η_t is a q-variate time series of latent factor scores, Λ denotes $p \times q$ matrix of factor loadings, and ε_t is a p-variate error time series capturing specific and measurement error. Further, η_t is modeled as a function of prior latent states weighted by B_k ($k = 1, \ldots, s$) while ζ_t denotes present-time "innovations" (i.e., the error component representing new information at a given time point). In addition to lagged and cross-lagged effects of latent factors, contemporaneous latent variable effect may be incorporated as well. Note that present "innovations" are assumed to be independent of latent states on the predictor side of the model which may open the door to apply principles of DDA to evaluate directional hypotheses of latent contemporaneous effects in dynamic factor models.

References

Asendorpf, J. B. (2015). Person-oriented approaches within a multi-level perspective. *Journal for Person-Oriented Research*, *1*, 48–55.

Beale, E. M. L., & Mallows, C. L. (1959). Scale mixing of symmetric distributions with zero means. *Annals of Mathematical Statistics*, *30*, 1145–1151.

Bergman, L. R. (2009). Mediation and causality at the individual level. *Integrative Psychological and Behavioral Science*, *43*, 248–252.

Bergman, L. R., & Lundh, L. G. (2015). The person-oriented approach: Roots and roads to the future. *Journal of Person-Oriented Research*, *1*, 1–6.

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*. Mahwah, NJ: Lawrence Erlbaum.

Blanca, M. J., Arnau, J., López-Montiel, D., Bono, R., & Bendayan, R. (2013). Skewness and kurtosis in real data samples. *Methodology*, *9*, 78–84.

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Conger, J. J. (1956). Reinforcement theory and the dynamics of alcoholism. *Quarterly Journal of Studies on Alcohol*, *17*, 296–305.

Darmois, G. (1953). Analyse générale des liaisons stochastique [General analysis of stochastic links]. *Review of the International Statistical Institute*, *21*, 2–8.

Dodge, Y., & Rousson, V. (2000). Direction dependence in a regression line. *Communications in Statistics: Theory and Methods*, *29*, 1957–1972.

Dodge, Y., & Rousson, V. (2001). On asymmetric properties of the correlation coefficient in the regression setting. *The American Statistician*, *55*, 51–54.

Dodge, Y., & Yadegari, I. (2010). On direction of dependence. *Metrika*, *72*, 139–150.

Downs, W. R., Smyth, N. J., & Miller, B. A. (1996). The relationship between childhood violence and alcohol problems among men who batter: An empirical review and synthesis. *Aggression and Violent Behavior*, *1*, 327–344.

Galton, F. (1886). Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute of Great Britain and Ireland*, *15*, 246–263.

Geldhof, G. J., Bowers, E. P., Johnson, S. K., Hershberg, R., Hilliard, L., & Lerner, R. M. (2014). Relational developmental systems theories of positive youth development: methodological issues and implications. In P. C. M. Molenaar, R. M. Lerner, & K. M. Newell (Eds.), *Handbook of developmental systems theory and methodology* (pp. 66–94). New York: Guilford Press.

Gendolla, G. H. E. (2000). On the impact of mood on behavior: An integrative theory and a review. *Review of General Psychology*, *4*, 378–408.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*, 424–438.

Granger, C. W. J., & Newbold, P. (1976). Forecasting transformed series. *Journal of the Royal Statistical Society B*, *38*, 189–203.

Gretton, A., Fukumizu, K., Teo, C. H., Song, L., Schölkopf, B., & Smola, A. J. (2008). A kernel statistical test of independence. *Advances in Neural Information Processing Systems*, *20*, 585–592.

Hanley, J. A. (2004). "Transmuting" women into men: Galton's family data on human stature. *American Statistician*, *58*, 237–243.

Hernández-Lobato, J. M., Morales-Mombiela, P., & Suárez, A. (2011). Gaussianity measures for detecting the direction of causal time series. In *Proceedings of the international joint conference on artificial intelligence* (Vol. 22, pp. 1318–1323).

Hsiao, C. (1982). Autoregressive modeling and causal ordering of economic variables. *Journal of Economic Dynamics and Control*, *4*, 243–259.

Huemer, M., & Kovitz, B. (2003). Causation as simultaneous and continuous. *Philosophical Quarterly*, *53*, 556–565.

Hume, D. (1777). *Enquiries concerning human understanding and concerning the principles of morals*. Oxford: Clarendon Press.

Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent component analysis*. New York: Wiley & Sons.

Hyvärinen, A., Zhang, K., Shimizu, S., & Hoyer, P. O. (2010). Estimation of a structural vector autoregression model using non-gaussianity. *Journal of Machine Learning Research*, *11*, 1709–1731.

Katerndahl, D. A., Burge, S. K., Ferrer, R. L., Becho, J., & Wood, R. (2010). Complex dynamics in intimate partner violence: A time series study of 16 women. *Primary Care Companion to the Journal of Clinical Psychiatry*, *12*(4).

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, *54*, 159–178.

Lerner, R. M. (1998). Theories of human development: Contemporary perspectives. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (5th ed., pp. 1–24). New York: Wiley.

Link, B. G., & Shrout, P. E. (1992). Spurious associations in longitudinal research. *Research in Community and Mental Health*, *7*, 301–321.

Lozano, A. C., Abe, N., Liu, Y., & Rosset, S. (2009). Grouped graphical granger modeling for gene expression regulatory networks discovery. *Bioinformatics*, *25*, i110–i118.

Luthra, R., & Gidycz, C. A. (2006). Dating violence among college men and women evaluation of a theoretical model. *Journal of Interpersonal Violence*, *21*, 717–731.

Lütkepohl, H. (1985). Comparison of criteria for estimating the order of a vector autoregressive process. *Journal of Time Series Analysis*, *6*, 35–52.

Lütkepohl, H. (2007). *New introduction to multivariate time series analysis*. Berlin: Springer.

Magnusson, D. (2001). The holistic-interactionistic paradigm: Some directions for empirical developmental research. *European Psychologist*, *6*(3), 153–162.

Mamai, L. V. (1963). On the theory of characteristic functions. *Selected Translations in Mathematical Statistics and Probability*, *4*, 153–170.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models*. London: Chapman Hall.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, *105*, 156–166.

Mill, J. S. (1851). *A system of logic* (3rd ed., Vol. I). London: John W. Parker.

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika*, *50*, 181–202.

Molenaar, P. C. M. (2010). Testing all six person-oriented principles in dynamic factor analysis. *Development and Psychopathology*, *22*, 255–259.

Perrine, M. W., Mundt, J. C., Searles, J. S., & Lester, L. S. (1995). Validation of daily selfreport consumption using interactive voice response (IVR) technology. *Journal of Studies on Alcohol and Drugs*, *56*, 487–490.

Peters, J., Janzing, D., Gretton, A., & Schölkopf, B. (2009). Detecting the direction of causal time series. In *Proceedings of the 26th annual international conference on machine learning* (pp. 801–808)). ACM.

Pornprasertmanit, S., & Little, T. D. (2012). Determining directional dependency in causal associations. *International Journal of Behavioral Development*, *36*, 313–322.

Rovine, M., & Lo, L. (2012). Issues and perspectives in person specific time series models. In B. Laursen, T. Little, & N. Card (Eds.), *Handbook of developmental research methods* (pp. 313–332). New York, NY: Guilford.

Russell, B. (1913). On the notion of cause. *Proceedings of the Aristotelian Society*, *13*, 1–16.

Shimizu, S., Hoyer, P. O., Hyvärinen, A., & Kerminen, A. (2006). A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, *7*, 2003–2030.

Sim, C. H. (1994). Modelling non-normal first-order autoregressive time series. *Journal of Forecasting*, *13*, 369–381.

Skitovich, W. P. (1953). On a property of the normal distribution. *Doklady Akademii Nauk SSSR*, *89*, 217–219.

Stephan, K. E., & Roebroeck, A. (2012). A short history of causal modeling of fMRI data. *Neuroimage*, *62*, 856–863.

Sterba, S. K., & Bauer, D. J. (2010). Matching method with theory in person-oriented developmental psychopathology research. *Development and Psychopathology*, *22*, 239–254.

Sungur, E. A. (2005). A note on directional dependence in regression setting. *Communications in Statistics: Theory and Methods*, *34*, 1957–1965.

Swift, A. L., & Janacek, G. J. (1991). Forecasting non-normal time series. *Journal of Forecasting*, *10*, 501–520.

Timmons, A. C., & Preacher, K. J. (2015). The importance of temporal design: How do measurement intervals affect the accuracy and efficiency of parameter estimates in longitudinal research? *Multivariate Behavioral Research*, *50*, 41–55.

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., Bergman, L. R., & Hsieh, C. A. (2015). Person-oriented methodological approaches. In W. F. Overton & P. C. M. Molenaar (Eds.), *Handbook of child psychology and developmental science: Theory and methods* (pp. 789–841). Hoboken, NJ: Wiley.

von Eye, A., & DeShon, R. P. (2012). Directional dependence in developmental research. *International Journal of Behavioral Development*, *36*, 303–312.

von Eye, A., & Mun, E. Y. (2012). Interindividual differences in intraindividual change in categorical variables. *Psychological Test and Assessment Modeling*, *54*, 151–167.

von Eye, A., & Wiedermann, W. (2015). Manifest variable granger causality models for developmental research: A taxonomy. *Applied Developmental Science*, *19*, 183–195.

von Eye, A., & Wiedermann, W. (2016). Interindividual differences in intraindividual change in predictability: A configural perspective. *Journal of Person-Oriented Research*, 64–77.

Walls, T. A., & Schafer, J. L. (2006). *Models for intensive longitudinal data*. Oxford: Oxford University Press.

Wiedermann, W. (2015). Decisions concerning the direction of effects in linear regression models using fourth central moments. In M. Stemmler, A. von Eye, & W. Wiedermann (Eds.), *Dependent data in social sciences research: Forms, issues, and methods of analysis* (pp. 149–169). Cham: Springer.

Wiedermann, W., Artner, R., & von Eye, A. (2016). *Heteroscedasticity as basis of direction dependence in reversible linear regression models.* (under review)

Wiedermann, W., Hagmann, M., & Eye, A. (2015). Significance tests to determine the direction of effects in linear regression models. *British Journal of Mathematical and Statistical Psychology*, *68*, 116–141.

Wiedermann, W., Hagmann, M., Kossmeier, M., & von Eye, A. (2013). Resampling techniques to determine direction of effects in linear regression models. *InterStat*. Retrieved from http://interstat.statjournals.net/YEAR/2013/articles/1305002.pdf

Wiedermann, W., & von Eye, A. (2015a). Direction-dependence analysis: A confirmatory approach for testing directional theories. *International Journal of Behavioral Development*, *39*, 570–580.

Wiedermann, W., & von Eye, A. (2015b). Direction of effects in mediation analysis. *Psychological Methods*, *20*, 221–244.

Wiedermann, W., & von Eye, A. (2015c). Direction of effects in multiple linear regression models. *Multivariate Behavioral Research*, *50*, 23–40.

Wiedermann, W., & von Eye, A. (2016). Directionality of effects in causal mediation analysis. In W. Wiedermann & A. von Eye (Eds.), *Statistics and causality: Methods for applied empirical research.* Hoboken, NJ: Wiley and Sons. (in press)

Williams, J. G., & Smith, J. P. (1994). Drinking patterns and dating violence among college students. *Psychology of Addictive Behaviors*, *8*, 51–53.

Wood, P. K. (2012). Dynamic factor models for longitudinally intensive data: Description and estimation via parallel factor models of cholesky decomposition. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 571–584). New York: Guilford.

Young, R. M., Oei, T. P., & Knight, R. G. (1990). The tension reduction hypothesis revisited: an alcohol expectancy perspective. *British Journal of Addiction*, *85*, 31–40.

Yule, G. U. (1921). On the time-correlation problem, with especial reference to the variate-difference correlation method. *Journal of the Royal Statistical Society*, *84*, 497–537.

Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series? a study in sampling and the nature of time-series. *Journal of the Royal Statistical Society*, *89*, 1–63.