In this comment, I first highlight the contributions of Robinson-Cimpian, Lubienski, Ganley, and Copur-Gencturk (2014) in particular and a more interdisciplinary approach in general for the subdiscipline of developmental psychology. Second, I identify some historic methodological foci of psychology and encourage Robinson-Cimpian et al. to consider these possibilities. Finally, I end with my perception that the emerging subdiscipline of developmental methodology can provide a common context for interdisciplinary dialogue in developmental science.

Keywords: developmental methodology, causal analysis, measurement, longitudinal design

It is a pleasure to comment on this important article by Robinson-Cimpian, Lubienski, Ganley, and Copur-Gencturk (2014). I focus my comments on the methodological aspects of this article, highlighting disciplinary differences between economics (the discipline in which Robinson-Cimpian was trained) and the discipline of psychology. In my comment, I first highlight the contributions of this work in particular and a more interdisciplinary approach in general for the subdiscipline of developmental psychology. Second, I identify some historic methodological foci of psychology and encourage Robinson et al. to consider these possibilities. Finally, I end with my perception that the emerging subdiscipline of developmental methodology can provide a common context for interdisciplinary dialogue in developmental science.

Interdisciplinary Contributions to Developmental Psychology

In addition to the substantive contributions of Robinson-Cimpian et al. (2014), some of the methodological features of this study serve as important models for developmental psychology. At the risk of misattributing these strengths to the authors’ disciplinary backgrounds or homogenizing my portrayal of those disciplines, I describe two key strengths of this article that should serve as valuable models for developmental psychology.

Robinson-Cimpian et al. (2014) aimed to use instrumental variable analyses and propensity score matching to draw causal conclusions from nonexperimental data. This effort to make causal conclusions in real-world (i.e., necessarily nonexperimental) situations is an important endeavor that psychology seems to have abandoned. Most of us trained in psychology are very familiar with the diversity of threats to conclusions of causality in various nonexperimental designs and very sensitive to the incessant third-variable threat in naturalistic longitudinal designs. Although recognition of threats to causality is important, it seems that psychology has taken this too far; now cause is viewed as a forbidden word, which is sure to lead to immediate and catastrophic ridicule if one dares to use this c-word in the context of nonexperimental data. Psychology, including developmental psychology, has been introduced to the techniques of instrumental variable analysis (e.g., Gennetian, Magnuson, & Morris, 2008) and propensity score analysis (e.g., Hong & Yu, 2008; Stuart & Green, 2008). However, to date, these analytic techniques and the more general effort to draw causal conclusions from nonexperimental data have been limited. Robinson-Cimpian et al.’s article serves as an example that using nonexperimental data to study causal processes is a valuable goal for developmental science.

A second key strength of Robinson-Cimpian et al. (2014) is that the sample of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–1999 (ECLS-K), is nationally representative. This attention to sampling is not unique to economics, as criminology, education, public health, and sociology all seem attentive to obtaining representative samples. In contrast, psychology, as well as developmental science in general, has too often neglected to prioritize efforts to obtain representative samples. In a survey of

Editor’s Note. Jacquelynne Eccles served as the action editor for this article.—JSE
psychology and interdisciplinary developmental journals, Bornstein, Jager, and Putnick (2013) found that an extremely small percentage of studies attempted to draw samples representative of larger populations. This shortcoming is concerning, as nonrepresentative samples can lead to several biases (see Jaeger, Bornstein, & Putnick, 2013). Given how much attention psychology places on precise inferential tests (e.g., p values; see, e.g., Cohen, 1994; Harlow, Mulaik, & Steiger, 1997), it would be productive to better sample those populations psychologists wish to make inferences about.

Potential Contributions of Developmental Psychology to Robinson-Cimpian et al. (2014)

Without detracting from the many strengths of Robinson-Cimpian et al. (2014), there are also aspects of it that could be improved by drawing from methodological traditions that have been foci within psychology. Next, I describe considerations regarding measurement and time that would have benefited this article.

Measurement

Psychology has historically given substantial attention to measurement through the subdiscipline of psychometrics. There are two lessons from psychometrics that could further strengthen this article. First, a fundamental realization of psychometricians is that responses on any item are a function of the construct that the item is intended to assess (e.g., teacher rating of math proficiency), aspects that are specific to that item (e.g., an item assessing ordering versus other aspects of math performance), and aspects that are essentially random. Robinson-Cimpian et al. (2014) used variables that are simple manifest variable composites of items; therefore, between-child variability in the variables used included not only individual differences in math but also individual differences in item-specific components and random variability. The impact of the unreliability of manifest variables on simple two-variable associations (e.g., gender differences in teacher ratings of math proficiency) is straightforward: Unreliability will always attenuate (minimize) these effect sizes. However, in multivariate models, the impact of unreliability becomes difficult to predict. Unreliability of the predictor or dependent variable will tend to attenuate predictive paths, but unreliability of the predictors, covariates, and/or instrumental variables will attenuate controls and therefore artificially inflate these same predictive paths. Given the complexity of the models presented in this article, the impact of this unreliability is difficult to predict.

A second potential contribution of psychometrics Robinson-Cimpian et al. (2014) comes from the literature on measurement invariance (Little, 1997; Meredith, 1993; Widaman, Ferrer, & Conger, 2010). Briefly, measurement invariance is established using latent variable models (confirmatory factory analyses or item response theory models) in which measurement properties (factor loadings, indicator intercepts) are constrained to be equal. These constraints ensure that constructs are defined equivalently across different groups or measurement occasions. Robinson-Cimpian et al. did not perform such comparisons and did not use latent variable models in which the constraints of measurement equivalence were held in place. Therefore, it is unclear whether any gender differences that were found were due to gender differences in the constructs or simply due to gender differences in the measurement of the constructs. Conversely, instances in which gender differences were not found could have occurred when there were actual gender differences in the constructs but differences in the measurement properties obscured these differences. Finally, tests of longitudinal invariance (i.e., equality of the measurement across time) would have strengthened conclusions of developmental differences. It is difficult to accept the conclusion that gender differences in a construct change over time when there is no evidence that one is assessing the same construct across time.

Time

Another issue that has received much attention from psychologists, especially developmental psychologists and clinical psychologists, is time. Any causal process—whether change across development, improvement across the course of treatment, or teacher perceptions influencing later math achievement test scores—must unfold over some finite period of time. Robinson-Cimpian et al. (2014) used data collected in the fall and spring of kindergarten and in the springs of Grades 1, 3, and 5, so the lags between successive waves of data collection were roughly 6 months, 1 year, and then 2 years. These decisions about data collection timing were made for the larger ECLS-K study and presumably not determined by any theoretical expectations for the processes investigated here. Cole and Maxwell (2003) described the biasing impact of misspecifying time lags in longitudinal designs (i.e., choosing a time span that is longer or shorter than the time over which the causal process occurs). The impact of misspecification of time lags in longitudinal predictions is difficult to predict: Estimates may be inflated if the process occurs many times between successive occasions, estimates may be deflated if the time lag studied does not allow for the process to completely unfold, and estimates may also be deflated if the causal impact has occurred but subsequently dissipated. Caution should be used when making statements of causal processes when there is a lack of clarity about whether the times spans studied are truly those across which the causal process unfolds.

There might also be a second concern about time lags. Selig, Preacher, and Little (2012) studied another widely used, nationally representative data set, the Early Head Start Research and Evaluation Study. They found that there was sizable between-family variability in the time lags between measurement occasions; for two data collection waves that were intended to be 10 months apart, families were actually assessed at various time lags ranging from under 6 months to close to 17 months between measurement occasions. Given the presence of such large variability in lag in that data set, it would be beneficial to know the range of time lags between measurement occasions in the ECLS-K used in the target article.

If the ECLS-K contains between-child variability in time lag, this by itself might suggest a weakness. However, this could

---

2 I acknowledge the difficulty in establishing longitudinal invariance when it is also necessary to change items across time. Here, it would be best to establish partial measurement invariance (Millsap & Kwok, 2004) by imposing equalities of factor loadings and indicator intercepts on those items that were the same across successive measurement occasions.
instead be considered a strength if one follows the analytic recommendations of Selig et al. (2012). Use of the lag as moderator approach, in which time lag is treated as an individual difference variable and then evaluated as a moderator of predictive or causal effects, would allow Robinson-Cimpian et al. (2014) to identify empirically the time lag over which the causal processes that they identified operate.

Developmental Methodology as a Context for Interdisciplinary Dialogue in Developmental Science

In this comment, I have identified strengths of Robinson-Cimpian et al. (2014) derived from the Robinson-Cimpian’s disciplinary rooting in economics, as well as points that could potentially further strengthen the work offered from the discipline of psychology. This sharing of disciplinary strengths is one of the main motivations of taking an interdisciplinary approach to understanding human development. My focus on methodological features, as opposed to other aspects (e.g., theoretical groundings, implications), may represent one of the most readily available routes to interdisciplinary dialogue. As illustrated by the issues discussed here, different disciplines have emphasized different methodological aspects, but these aspects are not in opposition; I expect that few psychologists would disagree with the importance of evaluating causal processes or using representative samples, nor would members of other disciplines likely disagree with the importance of measurement and attention to time in longitudinal studies. The emerging subdiscipline of developmental methodology thus provides an optimal context for engaging in interdisciplinary discussions that can immediately be understood and applied to positively impact the quality of developmental science. I thank Robinson-Cimpian, Lubieski, Ganley, and Copur-Gencturk for their contribution and for sharing interdisciplinary insights with the readers of Developmental Psychology.

References


Received August 23, 2013
Revision received October 30, 2013
Accepted November 11, 2013