JOURNAL OF
Research on Adolescence

Society for Research on Adolescence
www.s-r-a.org

JOURNAL OF RESEARCH ON ADOLESCENCE, 25(1), 1–13

Using Experimental Designs to Understand the Development of Peer Relations

Bram Orobio de Castro, Sander Thomaes, and Albert Reijntjes
Utrecht University

In the past decades, tremendous advances have been made in our understanding of peer relations. A description is emerging of *how* peer relations develop. Surprisingly, though, we know less about *why* peer relations develop as they do. Experimental designs provide opportunities to learn about causal processes in peer relations, yet they have not been used frequently in this field. The aim of this paper is to encourage greater use of experimental designs to unravel causal processes that shape the development of peer relations. To this end, limitations of causal inferences from longitudinal studies are presented, it is demonstrated how experimental designs and experiments nested in longitudinal studies can be used to test causal mechanisms, and intervention trials as real-life experiments are discussed.

> If you want to understand something, Urie, try to change it.
> (Dearborn, quoted in Bronfenbrenner, 1979, p. 407)

Imagine you are a Martian, on a mission to study earth. By chance, your spaceship drops you on a large train station, so you start your observations there. After a while, you notice a clear pattern: Time after time, when groups of people gather on platforms, after a few minutes a train arrives. Being a Martian with the very human-like inclination to see causality in temporal patterns, you immediately call your headquarters to inform them of your discovery: “On planet earth, crowds cause trains!”

The captain of your spaceship is not quite convinced and beams you a statistician (Mars is known for its excellent statisticians). Together you do further analyses and discover that you can model distinct developmental trajectories, with large crowds apparently causing large trains and small crowds causing small trains. Multilevel analyses reveal that these trajectories even seem to depend on the specific station where crowds gather. In a final groundbreaking study, you even establish mediation: the relation between crowds and trains is fully mediated by a cracking voice from a sound system blurting out the mysterious text “mind the gap” just before the train arrives.

The predicament of the Martian in this example is not so different from the task developmental scientists face. The mainstay of developmental science in the past decades seems to provide adequate descriptions of changes over time and to use covariation between time points as a basis to draw inferences about causality. In the train example, these inferences are clearly wrong. But how can we tell? Only because we know that crowds cannot cause trains to emerge. What if we did not know? Would we not have considered this prediction of trains by crowds a plausible indicator of causality, as a risk factor to be addressed by policy? How many of our colleagues call for “longitudinal studies to demonstrate causality” in discussion sections of articles? Would we not be impressed by full mediation? Apparently, we have a very strong inclination to see covariation over time as indications of causality (Kant, 1999/1781).

Experimental designs provide unique opportunities to learn about causal processes in peer relations. After all, the clearest test of causality is obtained when a hypothesized cause is manipulated and the hypothesized effect is measured. The Martian in our example could only have falsified his causal theory by creating (or removing) a crowd and noting that this did not affect the arrival of trains. Similarly, we can only falsify our theories about the causal processes in peer relations through experimental manipulation. Surprisingly, though, experiments have not been used frequently in this field. A simple search in Web of Science

This article is adapted from a keynote address delivered at the SRA Peer Preconference in Vancouver, Canada, March 7, 2012. We would like to thank Willem Koops and Brad Bushman for their thought-provoking discussions of causation and development.

Requests for reprints should be sent to Bram Orobio de Castro, Utrecht University, P.O. Box 80.140, 3508 TC, Utrecht, The Netherlands. E-mail: b.castro@fss.uu.nl

DOI: 10.1111/jora.12103

indicates that only 84 of 4.739 publications on peer relations in the past years mention an experimental manipulation in their abstracts. Experiments nested within longitudinal studies are even rarer. We believe experimental designs provide far more opportunities for developmental sciences than have been realized so far.

The aim of this article is to inspire colleagues to use experimental designs to discover the causal processes that shape the development of peer relations. To this end, we first discuss limitations of inferences about causality based on longitudinal data, including mediation analyses. Next, we discuss the iterative nature of causal effects in current developmental theorizing. We then propose that experimental designs and experiments nested in longitudinal (intervention) studies can be used to test such causal mechanisms and provide illustrations of this approach. Finally, we discuss the potential benefits and limitations of experimental designs for research on peer relations.

INFERRING CAUSALITY FROM LONGITUDINAL DATA

Scholars have questioned how to establish causality for millennia (e.g., Aristotle, see Hankinson, 1998; Bryant, 1990; Kant, 1999/1781). Longitudinal designs provide important information regarding the development of relations over time, but cannot test whether these relations are causal. Yet the advent of sophisticated statistical techniques to analyze longitudinal data may perhaps obscure to some extent how our developmental sciences are still predominantly relying on (longitudinal) correlational data that cannot prove causality, no matter what sophisticated statistical technique is used. What some techniques (e.g., cross-lagged analysis, mediation analyses) can do is establish temporal ordering. But, like the train example illustrates, that does not equal causality. Causal inferences made from longitudinal studies are rarely warranted, even when statistical procedures are used that take temporal ordering of measurements into account.

Testing mediation has become a popular means to test—or suggest—causal processes in longitudinal studies. But mediation is no substitute for causation. The rationale for mediation analyses is that if a mediator is measured after the hypothesized cause and before the hypothesized effect, and if the mediator explains covariation between these, the mediator is likely to be the causal mechanism underlying the relation between hypothesized cause and effect. As the Martian example shows, this line of reasoning is problematic, because it is based on the assumption that covariance and temporal ordering suffice to infer causality.

In addition, mediation is often tested on the wrong time scale. Causal effects in social relations (e.g., mutual conditioning or social information processing) typically occur fast. Consider this example: Child A bumps into child B, child B infers hostile intent by child A and complains, child A sees this remark made in public as a threat to his status and he bumps into child B again, the rest of the classroom laughs at child B, and someone calls him a loser. Here, in a couple of seconds, a dominance hierarchy is reestablished, child A’s behavior is reinforced, and child B’s expectations of hostile intent are strengthened. When mediation analyses are used to test hypotheses concerning these processes, mediation should clearly be assessed in timeframes of seconds rather than years. Yet measurement occasions separated by at least several months are the norm in longitudinal studies. When measurements of hypothesized cause, mediation, and hypothesized effect are so far apart in time that mediation tests actually investigate whether child A bumping into child B causes child B to make a negative attribution 1 year later, that would in turn cause him to speak out to child A after 2 years. These are clearly not the causal processes the concerning theories refer to.

Consequently, many studies demonstrate relations between broad constructs over time, but discuss them in terms of causal mechanisms that were not studied. For example, predictive relations between internalizing and externalizing problems have frequently been studied with cross-lagged analyses of only these two constructs, but discussed in terms of mechanisms that were not included in those studies, such as rejection and decreasing self-efficacy. This limitation may become less evident when longitudinal studies include more direct measures of mediating mechanisms on more measurement occasions (e.g., in microgenetic designs). But even with such extensive longitudinal designs, the fundamental issue that assessment of temporal ordering does not demonstrate causality remains. Apparently, to find out how to study causation, we first need a clear picture of the nature of causal processes in current developmental theory.

CAUSAL MECHANISMS IN DEVELOPMENT

Development of social relations is shaped by transactional processes (e.g., Van Geert & Steenbeek, 2005). When two or more people interact, the behavior of each person influences the others and vice versa. After each interaction, all the persons involved have changed. A next interaction between these persons will be an interaction between these "changed" persons, not between the persons they were before the interaction occurred. Such transactional processes are formally called iterative processes. In an iterative process, the state of the system at time T determines the system's behavior at T_{+1}, which in turn determines the system at T_{+2}, and so on. Thus, the state of a system at T_i is not a function of the system at T_0, but of the system at T_{i-1}. In such systems, it may be possible to explain variance at T_i from variance at T_0, but understanding the actual causes of changes in the system would require understanding the function that changes T into T_{+1} at each iteration (Van Geert & Steenbeek, 2005).

Many theories of social behavior, including social behavior between peers, essentially concern such iterative processes, describing how an organism and the environment change each other continuously over the course of development, and giving clear hypotheses about the causal mechanisms involved at each iteration. Some examples in the realm of peer relations are mutual operant conditioning in coercive interactions (Patterson, Dishion, & Bank, 1984), escalating conflicts resulting from hostile social information processing (Dodge, 2006), and the development of participant roles in classrooms (Salmivalli, 2010). For instance, participant roles theory proposes that bullying does not result from some children having characteristics that turn them into bullies and others being disposed to become victims. Bullying is rather seen as the consequence of a series of iterations in the group. The process may start with one unkind remark from child A to child B that is reinforced by laughter from children C and D. In a next iteration, child A may be more likely to repeat this reinforced behavior; children C and D may be more likely to laugh again because they have been reinforced by attention from child A; and child B may have become less likely to defend himself successfully due to decreased social status. The dynamics of the participant roles in this example are not explained by a long-term prediction from start to outcome, but rather by repetitive iterative processes.

Surprisingly, though, the iterative causal processes underlying these theories are rarely tested directly in the longitudinal studies on peer relations that dominate the field to date. Longitudinal studies tend to consist of measurement occasions separated by intervals of several months or even years. So what is actually measured are a few time points in a long series of iterations, not the iterative process itself. By taking snapshots with long intervals between them, we may miss the very process of development itself. If we seek to understand the causes of social development and if our theories concern iterative processes, we need to test whether these processes actually occur at each iteration and whether they are causal. A promising approach to part of this problem is provided by *microgenetic* designs with much more frequent measurements, tested with iterative models of development. Good examples include cascade models (e.g., Lansford, Malone, Dodge, Pettit, & Bates, 2010) and the modeling of iterative development (Steenbeek & van Geert, 2008). Yet, even such intensive longitudinal designs do not directly address the issue of causality. In iterative processes, the state of a system may vary tremendously over short intervals, which may result in different findings when measurements happened to be conducted at T_i than when they would have occurred at T_{i+1}.

AN EXPERIMENTAL APPROACH

To test causal mechanisms in the development of peer relations, it may be interesting to look at the specific process involved in a single iteration. To establish causality, we should then manipulate the input for this iteration and assess whether the outcome changes as we predicted. In other words, we should conduct a social experiment. The results of this experiment may then be extrapolated to a series of iterations over time.

However, experimenting is only useful when there are specific hypotheses about the causal relation to be tested. To test causality, only one independent variable at a time can be manipulated. Therefore, a proper analysis of causation requires a theory that is explicit about causal processes. Such a theory does not have to be a simplistic unicausal one-directional affair. It can be complex, involving multiple interdependent causes, transactional influences, and mediated moderation. Yet—in contrast to many current psychological theories—each hypothesized causal link needs to be specified. Testing for causality then requires a series of experiments, each involving a unique manipulation of one causal link with specified moderators and mediators.

DIRECT MANIPULATION OR TRIGGERED MEDIATION AND MODERATION

Once hypotheses about causal processes have been specified, experimental designs provide the possibility to test these hypotheses by manipulating each hypothesized cause and testing the effects on the hypothesized consequences. Ideally, such experimental tests consist of direct systematic manipulation of the hypothesized cause. In the domain of peer relations, one may, for example, manipulate social stimuli, peer group composition, contextual variables, and so on. We provide numerous examples below.

However, not every hypothesized cause can be manipulated. Important constructs such as genes, temperament, or the culture children live in, cannot be manipulated, making it hard to establish their causal role in the development of peer relations. Some constructs that cannot be manipulated due to their inherent stability (e.g., genetic makeup, personality traits) may be considered distal causes that moderate the proximal causal processes in everyday interactions. Perhaps the best way to handle these constructs experimentally is to include them as potential moderators in experimentally triggered mediation and moderation designs.

Triggered mediation and moderation designs do not involve direct manipulation of a putative cause, but experimental manipulation of a trigger hypothesized to activate the putative cause. Evidence for causality is then provided if the manipulation triggered the effect of the putative cause on the dependent variable, preferably through a priori specified moderated mediation. Suppose the hypothesis is that A causes C through mediator B when event D occurs. And suppose direct manipulation of A is not possible. A triggered mediation and moderation experiment may then test the plausibility of the hypothesis by examining whether experimental manipulation of event D causes changes in C through mediator B specifically for high A. This design is preferable over nonexperimental moderated mediation designs because it does not depend on mere covariation between variables, but on changes in covariation induced by the manipulated trigger.

Good examples of this approach are experimental studies on differential susceptibility in the parenting literature. Here, effects of parenting are hypothesized to depend on genetic differences in child susceptibility to environments. Obviously, this genetic susceptibility cannot be manipulated. Therefore, in experiments, the environment is manipulated through parenting intervention (D), genetic susceptibility is treated as a moderator (A), child emotional responses to changes in parenting are treated as mediators (B), and child behavior as the outcome (C) (Bakermans-Kranenburg, Van IJzendoorn, Pijlman, Mesman, & Juffer, 2008).

FEASIBILITY

One reason why experimental studies of the development of peer relations are relatively rare could be that experimentation with social relations may not be considered feasible. Social relations are central to children's well-being, and peer relations depend on intricate relations between many individual and interpersonal variables. The idea of experimentally manipulating this complex private world may perhaps seem bewildering.

We argue that experimental designs in research on peer relations are much more feasible than is often assumed. The feasibility of such experiments is clear from successful examples in the literature. Experimental designs are the norm in most disciplines of psychology. In developmental psychology, strong lines of experimental research have laid the foundations for our understanding of causal processes involved in such diverse fields as parenting and social cognition. Not coincidentally, the most well-established interventions for social problems have resulted from these lines of research (e.g., mediation therapy and cognitive behavioral treatment).

A good example of the benefits of an experimental approach to studying development is provided by social information processing (SIP) theory (Crick & Dodge, 1994; Dodge, 1986; Lemerise & Arsenio, 2000). This theory includes detailed hypotheses concerning causal processes in transactions between children. According to SIP theory, interaction partners process information about each other's behaviors in a specified series of processing steps, resulting in a behavioral response that is in turn processed by the other person. The theory describes in detail how this iterative process may result in stable deviant behavior patterns when information processing patterns of a child are reinforced by the responses their behavior evokes (Dodge, 2006; Dodge & Pettit, 2003). Interestingly, this theory has from the onset been studied experimentally. As early as 1980 (Dodge, 1980), Dodge and others started an influential experimental program in which causal relations between social stimuli, social information processing, and behavior were systematically studied. Social information was manipulated to test effects on social information processing, for example, by testing the effects of manipulating the ambiguity of social cues about intentions of other children on encoding, interpretation, and aggressiveness of responses (e.g., Dodge, Pettit, McClaskey, & Brown, 1986). Relations between different aspects of social information processing were studied experimentally, for example, by testing the effects of facilitating correct interpretations of social cues on consequent selection of responses (e.g., De Castro, Bosch, Veerman, & Koops, 2003; De Castro, Slot, Bosch, Koops, & Veerman, 2003; Dodge & Frame, 1982). Social information processing was manipulated to test effects of information processing on actual aggressive behavior in play sessions (Dodge, 1980; Lochman & Dodge, 1998). A host of contextual variables were manipulated to examine their effects on social information processing, including peer behavior (Dodge & Somberg, 1987), parent behavior (Barrett, Rapee, Dadds, & Ryan, 1996), and mood (De Castro, Slot et al., 2003). These experiments proper were embedded in a series of longitudinal studies concerning longer term precursors and consequences of social information processing, such as parenting (Lansford et al., 2006) and transactions between SIP and peer relations (Lansford et al., 2010). Findings from this line of research provided the basis for many interventions, some of which were used in longitudinal-experimental tests of social information processing theory (e.g., Lochman & Wells, 2002).

Experimental research on peer relations has also proven to be feasible in other lines of research. Groundbreaking studies on formation of social relations were conducted with experimentally composed play groups (Coie, Dodge, Terry, & Wright, 1991). In other studies, coercive interaction patterns were experimentally manipulated (Patterson et al., 1984), as were anger in peer conflicts (Smith, Hubbard, & Laurenceau, 2011), effects of peer feedback on retaliation (Sandstrom & Herlan, 2007), covert antisocial behavior (Hinshaw, 2005), and social exclusion (Underwood, Scott, Galperin, Bjornstad, & Sexton, 2004). Even aggression and reconciliation among boys with conduct disorders have been evoked experimentally (Kempes, De Castro, & Sterck, 2008; Van Nieuwenhuijzen et al., 2005).

Whether and how aspects of peer relations can be manipulated is not set in stone. The above examples illustrate the rich variety in procedures that can be used. Manipulations have included such diverse methods as instructing child actors to perform specific behaviors (Underwood et al., 2004), rigging games to evoke behaviors from a peer or parent toward the target child (Brummelman et al., 2013; Van Nieuwenhuijzen et al., 2005), suggesting the presence of fictitious peers using rigged computers (Thomaes et al., 2010), intercoms (Dodge & Somberg, 1987), or video sets (Snoek, Van Goozen, Matthys, Buitelaar, & Van Engeland, 2004), excluding a child from participation in a PC game (Cyberball: Williams & Jarvis, 2006), morphing facial expressions (Montagne, Kessels, De Haan, & Perrett, 2007), presenting hypothetical peer information (numerous studies), staging actual interactions (e.g., Kempes et al., 2008; Lochman & Dodge, 1998), composing new peer groups (Coie et al., 1991), providing specific instructions (De Castro, Bosch et al., 2003), psychoeducation (Yeager, Trzesniewski, Tim, Nokelainen, & Dweck, 2011), raising different expectations about one's social future (Baumeister, Twenge, & Nuss, 2002), or even providing designated interventions (e.g., Lochman & Wells, 2002; Van Lier, Vuijk, & Crijnen, 2005). Dependent variables do not only include behavior toward peers and indices of peer relations, but also cognitions, emotions, physiological indices of emotion and arousal, visual information seeking (Horsley, De Castro, & Van der Schoot, 2010), and information seeking on (rigged) social media (Reijntjes et al., 2011).

A WISHING LIST

Our understanding of developmental processes can thus be furthered by experimental studies. Yet, to fully understand development, experimental investigation of transactional iterations and longitudinal studies of series of iterations may best be combined. A pragmatic order to test for causality may be the following.

First, theory needs to be specified to such an extent that the presumed causal mechanism is clear and unambiguous. This may require considerable theoretical work, as many psychological theories are rather global in their specification of causal mechanisms (Bryant, 1990). For example, a theory predicting that low self-esteem would lead to rejection by peers due to socially incompetent behavior should be specified to include specific hypotheses about exactly how low self-esteem would cause incompetent behavior and so on.

Second, as a first test of the plausibility of the hypotheses under study, cross-sectional or longitudinal designs may be used to test whether the predicted relations exist. Structured observations of the exact temporal order of theoretically relevant processes in a controlled setting may be useful at this point. Specific analyses of these data may help test whether the pattern of relations is in line with theoretical predictions, for example, by testing for a specific structure (such as mediation or cascades) in the data. Such studies have the great advantage over experimental designs that multiple independent variables can be assessed over a longer period of time, even though causal effects cannot be established.

Third, if the predicted relations over time have been found, experiments may be used to test whether the established relations are indeed causal. The longitudinal data from the previous step help select the limited number of variables that can be manipulated in these experiments. The most straightforward test of causality is an experiment proper, where causality is examined by manipulating an independent variable and testing the effect on the dependent variable. The effect of the manipulation can either be tested by comparing participants who did and who did not receive the manipulation (between-subjects design) or by comparing responses to different conditions by the same participants (within-subjects design). More information about the causal process can be obtained by testing potential mediators and moderators. Many variations on this design are possible, including—but not limited to—multiple consecutive manipulations, multiple dependent variables, and effects measured at multiple time points (sometimes called microgenetic studies).

Fourth, even when experimental evidence has been found for the specified causal relation, we still need to demonstrate that this causal relation has an impact on actual child development for extended time periods. At this point, longitudinal-experimental designs may be used to follow the consequences of an experimental manipulation over a longer period of time (Lacourse et al., 2002). Importantly, in many cases, a single experimental manipulation cannot be expected to have long-term consequences. In these cases, the manipulation may be repeated at regular intervals to increase or sustain its impact. For example, one may test whether shared success experiences increase friendship quality in children with an isolated social position by randomly assigning these children to either weekly cooperative assignments or control assignments.

Fifth, in some cases, to provide even stronger evidence for prolonged causal effects of an experimental manipulation, the manipulation may be scaled up to an actual intervention to prevent or treat social problems. A randomized trial of this intervention may then serve as a large-scale experiment (e.g., Christakis et al., 2013; Lacourse et al., 2002; Lochman & Wells, 2002; Peterson, 2013; Project Multisite Violence Prevention, 2013). Thus, randomized intervention studies can serve as longitudinal-experimental studies with an intervention as the manipulation. Using interventions as experimental studies is most informative when they are not limited to mere measurement of intervention effects. Even in strictly protocolled intervention trials, the actual intervention dosage and intervention quality children receive vary considerably, as does the specific response pattern of participants (Chorpita et al., 2011). A comprehensive approach to intervention trials for experimental purposes would therefore include assessments of five Ws: What (the actual intervention delivered to each individual participant), Works (intervention effects, including side effects), When (the conditions of intervention delivery), for Whom (moderators), and Why (mediators). Randomized intervention trials including assessment of these five Ws have recently provided stringent tests of causal factors in, for example, bullying (Karna et al., 2011) and disruptive behavior (Van Lier et al., 2005).

It must be noted that the actual impact interventions have on development often comes at a price: In practice, interventions tend to be so large and complex that they do not serve as manipulations of one specific independent variable, but rather influence multiple potential causes. If a multicomponent intervention shows substantial effects on development, specific causal processes at work may be detected by breaking down the intervention into distinct components, to be tested separately (Chorpita et al., 2011).

A SOCIAL MEDIA EXAMPLE

The advent of social media for children is providing a whole new world of possibilities for experimental designs. To illustrate the potential of social media for experimental research and the feasibility of our pragmatic “wishing list” for research designs, it may be useful to use our recent line of research on self-views and aggressive behavior as an example.

There has been considerable debate about the relations between self-views and aggressive behavior. We aimed to determine whether aggressive behavior somehow results from low self-esteem or rather results from fragile inflated self-views being threatened by negative feedback. Following the above wishing list, we first set out to specify potential causal processes from relevant theories. Although this was no empirical exercise, it proved to be instructive, as we could not succeed in finding or making up a plausible theoretical causal link between low self-esteem and aggressive behavior (see Boxtel, De Castro, & Goossens, 2004 for details). In contrast, a detailed theory about threatened inflated self-esteem has been formulated by Baumeister and colleagues (Baumeister, Bushman, & Campbell, 2000). According to their theory, it is not relevant to aggressive behavior whether self-views are high or low in an absolute sense. What matters is how realistic self-views are and how much people are concerned about holding favorable self-regard. People with inflated self-esteem are hypothesized to feel threatened when they are confronted with negative information about their social functioning and to aggress toward the source of negative social information to thwart this threat.

Next, we turned to cross-sectional and longitudinal data showing covariance between the variables discussed above. Cross-sectionally, we and others found that fragile inflated views of own social acceptance were related to aggressive behavior, specifically for children who receive much threatening social feedback because they are chronically rejected by their peers (Boxtel et al., 2004; Brendgen, Vitaro, Turgeon, Poulin, & Wanner, 2004). Longitudinally, we basically found the same pattern (De Castro, Brendgen, Van Boxtel, Vitaro, & Schaepers, 2007).

At this point, experimental designs were called for to examine causality. Direct manipulation of fragile self-views appeared to be difficult, so a triggered mediation and moderation approach was taken. We hypothesized that threatening social feedback would affect state self-esteem, mood, and behavior through hostile intent attribution (mediator), specifically in children with fragile self-views (moderator). We used Survivor, a virtual Internet game, to study effects of feedback from peers. It is introduced to participants as an Internet popularity contest. After filling out baseline measures (e.g., mood, state self-esteem), participants start by making an online personal profile on a computer, including a picture and information about hobbies, music preferences, and so on. Next, participants are presented with the personal profiles of the (alleged) other participants in the game. In reality, these profiles have been created for the experiment and can be systematically manipulated on many dimensions, such as gender, ethnicity, popularity, hobbies, preferences, and attractiveness of faces on the pictures. While participants view these profiles, their computer keeps a log of the pages they view and the time they spend viewing specific items, allowing analysis of specific preferences, such as a tendency to look at popular versus less popular children. After viewing the profiles, participants are informed that the other contestants will now cast their votes. The computer takes some time to make (bogus) calculations and then presents the feedback from the judges. Obviously, this feedback is experimentally manipulated to be negative (voted out), neutral, or positive. Again, the nature of the feedback can be manipulated in many ways, such as a mere decision (being voted in or out), specific remarks made anonymously, or remarks made by specific contestants. Finally, responses to the social feedback serve as dependent variables, with many possibilities, including—but not limited to—opportunities for behavioral responses toward the other contestants (or even a specific contestant) such as giving or taking away rewards from a contestant, blasting loud noise through contestants' headphones, or posting a positive or negative comment about the contestant on the (bogus) site. It is also possible to administer measures of, for example, mood and state self-esteem throughout the game. Moreover, participants may be given the opportunity to view information about the other contestants again, enabling analyses of individual tendencies to focus on specific information, such as the personal profiles of children who gave negative or positive feedback. Finally, participants are thoroughly debriefed. Obviously, all studies using the Survivor Internet game have been approved by an ethical review board. None of the participants or their parents have ever expressed serious concerns or shown signs of lasting distress after the experiment (see also Thomaes et al., 2010).

Survivor was used to test whether children with fragile inflated self-views would respond more aggressively to threatening negative social feedback than their peers and whether this relation was mediated by hostile intent attribution. The alternative hypothesis that children with low self-esteem would respond most aggressively was also tested. In a series of experiments, effects of negative social feedback on mood, state self-esteem, and aggressive responses were studied. As predicted, aggressive responses were strongest by children with fragile inflated self-views who received negative feedback (Thomaes, Bushman, Stegge, & Olthof, 2008). These children also showed the greatest declines in mood and in state self-esteem (Thomaes, Reijntjes, De Castro, & Bushman, 2009; Thomaes et al., 2010). The effects of negative social feedback on aggression were mediated by hostile intent attribution (Reijntjes, Thomaes, Kamphuis, De Castro, & Telch, 2010; Reijntjes et al., 2011). No effect of low (or high) self-esteem on any dependent variable was found. Thus, the experiments provided clear support for the threatened inflated and fragile self-views hypothesis.

Following the wishing list above, experimental evidence should be supplemented with longitudinal-experimental evidence that causal effects established in simple experiments have an actual, and possibly enduring, impact on development in daily life. Here, we used a direct manipulation of fragility of self-views. Making self-views more fragile for a longer period of time obviously seemed unethical. We therefore decided to seek a positive manipulation as an ethically acceptable alternative. Using an experimental manipulation to the positive to test hypotheses concerning negative outcomes may at times be necessary for ethical reasons, but depends on the tenability of the assumption that the independent variable can be considered a continuum, where changes for the better have opposite causal effects from changes to the worse. In this case, we assumed that if experimentally *decreasing* fragility of self-views would decrease sensitivity to threat and aggressive behavior, this also speaks to opposite effects for increasing fragile self-views. Thus, instead of threatening fragile inflated self-views, we aimed to reduce the fragility of self-views. To this end, we gratefully used the brief self-affirmation procedure developed by Cohen, Garcia, Apfel, and Master (2006). In this brief (15 min) writing exercise, participants write about their most important personal values. This procedure has been found to make children less sensitive to stereotype threats and to have a marked positive effect on academic performance by children from ethnic minorities who face stereotype threats to their academic performance.

We predicted that the self-affirmation exercise would diminish sensitivity to negative social feedback by providing a focus on personal values children hold, regardless of how others may think about them. In a brief longitudinal-experimental field study (Thomaes, Bushman, De Castro, Cohen, & Denissen, 2009), participants were randomly assigned to the self-affirmation writing exercise or a general writing assignment of the same length. The self-affirmation writing exercise reduced aggressive behavior by children with inflated self-views for a week, as rated by peers who were blind to intervention status. A repetition of the writing session in the same study gave the same effect. Recently, these findings were supplemented with a prolonged positive effect of the self-affirmation exercise on children's prosocial behavior (Thomaes, Bushman, De Castro, & Reijntjes, 2012) and self-concept (Facchin, Margola, Molgora, & Revenson, 2013). Obviously, a brief writing exercise is no complete intervention by itself. These studies were not aimed to provide the largest possible effect, but to demonstrate causal impact. Having demonstrated this impact makes this specific manipulation a candidate to become part of more comprehensive interventions.

LIMITATIONS

Despite this richness in possibilities, experimental research on peer relations has practical and ethical limitations. A limitation we already mentioned is that not all variables can actually be manipulated. In these cases, experimentally triggering mediation and moderation may provide stronger indications for causal mechanisms than nonexperimental moderated mediation. Yet this approach may not suffice when neither putative causes nor triggers of mediation can be manipulated. When actual experimental manipulation is impossible, natural experiments may in some cases serve as a less controlled proxy. Good examples of such natural experiments include diverse independent variables such as early childhood deprivation in Romanian and English orphanages (e.g., Colvert et al., 2008), witnessing severe violence and war (e.g., Kuterovac-Jagodic, 2003), gang membership (Tolan, Gorman-Smith, & Henry, 2003), and economic crisis (e.g., Elder, Van Nguyen, & Caspi, 1985). If natural experiments are induced by institutions, it may even be possible to attain a certain degree of experimental control over the timing and assignment of conditions, as was demonstrated in induced mobility of families from high-poverty to low-poverty neighborhoods (Leventhal, Fauth, & Brooks-Gunn, 2005). In fact, many policy-induced changes would lend themselves perfectly to cluster-randomized gradual implementation. Nonetheless, we acknowledge that not all putative causes in developmental theory can be tested through experimental manipulation, moderation analyses, or natural experiments.

A related limitation is that not every experimental manipulation with an effect is the actual cause of that effect in real life. A well-known analogy to illustrate this issue is the effect of aspirin on headaches. Double-blind randomized trials clearly show that aspirin is more effective against headaches than placebo drugs. Yet demonstrating this effect experimentally does not imply that headaches in daily life are actually caused by an aspirin deficit. A proper understanding of this issue requires closer inspection of this example. When using experiments in the manner we suggest here, the aspirin experiment would not be used to demonstrate the effect of aspirin, but rather to test a theory about the causes of headaches. As far as we know (although it is clearly not our field of expertise…), common headaches are mostly caused by lessened blood flow through small veins due to high viscosity of the blood and/or contraction of veins. This theory could be tested experimentally by reducing the viscosity of the blood, which is exactly what aspirin does. Thus, in this analogy, aspirin is merely a tool used to manipulate a potential cause (viscosity of the blood), not a cause in itself. Additional experiments where blood viscosity is manipulated in other ways with the same effect would clarify this. Thus, unclarity about the potential cause under study and the exact nature of a specific experimental manipulation may lead to erroneous conclusions. It is therefore important to manipulate a potential cause in different ways in multiple experiments.

The ecological validity of experimental research into social development has been questioned. It has been argued that artificial situations in controlled laboratories have little to do with the actual dynamics of social relations. It seems to us that two issues may actually have been confounded in this discussion: the actual ecological validity of current studies and the (im)possibility of experimental studies of social development to be ecologically valid. The actual ecological validity of social relations experiments with youth to date seems too divergent for a general evaluation. Certainly, numerous studies with artificial manipulations and poor measures of dependent variables are being conducted in artificial contexts. Just as there are many experimental studies with careful manipulations, studying effects on real-life outcomes in natural contexts. Yet we see no a priori reason why ecological validity of experimental studies would necessarily be lower than the ecological validity of filling out questionnaires, being interviewed, or being rated by an adult informant in nonexperimental studies of development. In contrast, one may argue that external validity may be higher when a putative cause is actually manipulated in an experiment than when it is not. Ultimately, ecological validity of each study (whether experimental or not) is an empirical question. Unfortunately, research directly comparing findings from experimental and nonexperimental studies is scarce. A meta-analysis by Anderson, Lindsay, and Bushman (1999) compared findings from experiments to findings from nonexperimental studies on a number of social psychological topics and found a large degree of consistency in findings between the two approaches. Yet this meta-analysis did not directly assess correspondence between experimental and field findings for individual experiments. Direct empirical study of correspondence between findings with a specific experimental paradigm and corresponding field research (e.g., Van Nieuwenhuijzen et al., 2005) will allow for stronger conclusions about the ecological validity of specific experimental procedures.

Ethical limitations of experimental procedures require careful consideration. Social relations are important for children's development, so they should not be interfered with haphazardly. Other research methods used to study peer relations may also interfere with children's social lives (e.g., reflecting on own social functioning), but they are not explicitly designed to do so. Experimental researchers clearly have an important responsibility to adhere to ethical standards and guidelines. For useful advice on ethical issues in experimental research with children, see also a special issue of *Ethics and Behavior* (2005).

Moreover, experiments often involve temporary deception of participants. Deception in empirical research is only considered ethical if there is no alternative and the research is urgently needed to address important issues. This tradeoff needs to be considered carefully for each individual study. In addition to ethical concerns, deception of participants may also complicate the repeated use of a single experimental procedure in a longitudinal study. Once participants have been informed of the deception in the first take of the experiment, they cannot be considered naïve participants anymore in a consequent assessment. Using alternative experimental manipulations on different measurement occasions may provide a solution to this problem.

Notwithstanding these concerns, experimental manipulations are not necessarily taxing to children. In many cases, experimental manipulations may simply consist of systematic manipulation of events children face routinely in their everyday lives, such as winning or losing a game, getting specific information about peers, or being evaluated by peers in the real or virtual world. Moreover, with some creativity, manipulations can often be designed to affect children in a positive, rather than in a negative manner, for example, by assigning them to interventions. Whether manipulations of an independent variable to the positive provide valid information about the effects of (unethical) manipulations to the negative depends on the specific independent variable concerned. In some cases, it seems reasonable to assume that change to the positive and change to the negative are on a continuum, but in other cases effects of manipulations to the positive may not necessarily be generalized to reverse effects of negative events. For example, effects of positive peer feedback on children's self-views do not necessarily imply that negative peer feedback would have the opposite effect. When interpreting effects of manipulations to the positive, generalization to effects of changes to the negative need to be made with caution.

CONCLUSION

Experimental designs can help understand causal processes in the development of peer relations. Development of peer relations involves transactional processes. In the course of development, experiments can be considered as one iteration in a series of these transactions. Ideally, integrating experiments in longitudinal studies will enable the prediction of developmental trajectories from the responses to the experiment carried through for a large number of iterations.

Many different designs have already been used in experimental studies of peer relations. Manipulations of group composition, social tasks, social cognitions, and peer feedback are among the many possibilities. The popularity of social media with children provides many new opportunities for experimental procedures. Important challenges for future research include designing experiments with social constructs that may not directly appear to lend themselves for manipulation, such as social ambitions or resource control strategies. When direct manipulation of putative causes is not possible, triggered mediation and moderation design may serve as a proxy.

Social relations are at the core of children's lives. Experimentation with these relations is delicate. Ethical considerations should be a primary concern for experimental researchers. Experimental manipulations should preferably be made in ways that yield positive influences on social relations, as is the case in intervention research. When this is not possible, manipulations should not exceed the impact of experiences that children routinely experience in their day-to-day lives. In some cases, manipulations of hypothetical situations (e.g., social information processing vignettes) or virtual environments (e.g., the Survivor Internet game) may be possible.

Intervention studies may help demonstrate the actual impact of established causal factors on development over a longer period of time. Assessing the 5 Ws—what works when for whom and why—may help to get the most information about causality from intervention studies. However, intervention trials provide less specific information about causal processes, because interventions usually include a combination of many manipulations. Experimental research on specific intervention components may clarify the relative contribution of each intervention component to the effects of an intervention.

With this article, we hope to have inspired readers to combine longitudinal research with experimental research. We want to understand something, so let us change it.

REFERENCES

Anderson, C. A., Lindsay, A. J., & Bushman, B. J. (1999). Research in the psychological laboratory: Truth or triviality? *Current Directions in Psychological Science*, *8*, 3–9. doi:10.1111/1467-8721.00002

Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., Pijlman, F. T. A., Mesman, J., & Juffer, F. (2008). Experimental evidence for differential susceptibility: Dopamine D4 receptor polymorphism (DRD4 VNTR) moderates intervention effects on toddlers' externalizing behavior in a randomized controlled trial. *Developmental Psychology*, *44*, 293–300. doi:10.2007-19851-030

Barrett, P. M., Rapee, R. M., Dadds, M. M., & Ryan, S. M. (1996). Family enhancement of cognitive style in anxious and aggressive children. *Journal of Abnormal Child Psychology*, *24*, 187–203. doi:10.1002/ab.20337

Baumeister, R. F., Bushman, B. J., & Campbell, W. K. (2000). Self-esteem, narcissism, and aggression: Does violence result from low self-esteem. or from threatened egotism? *Current Directions in Psychological Science*, *9*, 26–29. doi:10.1111/1467-8721.00053

Baumeister, R., Twenge, J., & Nuss, C. (2002). Effects of social exclusion on cognitive processes: Anticipated aloneness reduces intelligent thought. *Journal of Personality and Social Psychology*, *83*, 817–827. doi:10.2002-18351-003

Boxtel, H. W., De Castro, B. O., & Goossens, F. A. (2004). High self-perceived social competence in rejected children is related to frequent fighting. *European Journal of Developmental Psychology*, *1*, 205–214. doi:10.1080/17405620444000102

Brendgen, M., Vitaro, F., Turgeon, L., Poulin, F., & Wanner, B. (2004). Is there a dark side of positive illusions? Overestimation of social competence and subsequent adjustment in aggressive and nonaggressive children. *Journal of Abnormal Child Psychology*, *32*, 305–320. doi:10.1023/b:jacp.0000026144.08470.cd

Bronfenbrenner, U. (1979). Een experimentele ecologie van de menselijke ontwikkeling. In W. Koops & J. J. van der Werff (Eds.), *Overzicht van de ontwikkelingspsychologie* (pp. 407–423). Groningen, The Netherlands: Wolters-Noordhoff.

Brummelman, E., Thomaes, S., Overbeek, G. J., De Castro, B. O., Van den Hout, M. A., & Bushman, B. (2013). On feeding those hungry for praise: Person praise backfires in children with low self-esteem. *Journal of Experimental Psychology: General*. Advance online publication. doi:10.1037/a0031917.

Bryant, P. (1990). Empirical evidence for causes in development. In G. Butterworth & P. Bryant (Eds.), *Causes of development: Interdisciplinary perspectives* (pp. 33–45). London: Harvester Wheatsheaf.

Chorpita, B. F., Rotheram-Borus, M. J., Daleiden, E. L., Bernstein, A., Cromley, T., Swendeman, D., & Regan, J. (2011). The old solutions are the new problem: How do we better use what we already know about reducing the burden of mental illness? *Perspectives on Psychological Science*, *6*, 493–497. doi:10.1177/1745691611418240

Christakis, D. A., Garrison, M. M., Herrenkohl, T., Haggerty, K., Rivara, F. P., Zhou, C., & Liekweg, K. (2013). Modifying media content for preschool children: A randomized controlled trial. *Pediatrics*, *131*, 431–438. doi:10.1542/peds.2012-1493

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*, 1307–1310. doi:10.1126/science.1170769

Coie, J. D., Dodge, K. A., Terry, R., & Wright, V. (1991). The role of aggression in peer relations: An analysis of aggression episodes in boys' play groups. *Child Development*, *62*, 812–826. doi:10.2307/1131179

Colvert, E., Rutter, M., Beckett, C., Castle, J., Groothues, C., Hawkins, A., ... Sonuga-Barke, E. J. (2008). Emotional difficulties in early adolescence following severe early deprivation: Findings from the English and Romanian adoptees study. *Development and Psychopathology*, *20*, 547–567. doi:10.S0954579408000278

Crick, N. C., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin*, *115*, 74–101. doi:10.1037//0033-2909.115.1.74

De Castro, B. O., Bosch, J. D., Veerman, J. W., & Koops, W. (2003). The effects of emotion regulation, attribution, and delay prompts on aggressive boys' social problem solving. *Cognitive Therapy and Research*, *27*, 153–166. doi:10.1023/a:1023557125265

De Castro, B. O., Brendgen, M., Van Boxtel, H., Vitaro, F., & Schaepers, L. (2007). "Accept me, or else.": Disputed overestimation of social competence predicts increases in proactive aggression. *Journal of Abnormal Child Psychology*, *35*, 165–178. doi:10.1007/s10802-006-9063-6

De Castro, B. O., Slot, N. W., Bosch, J. D., Koops, W., & Veerman, J. W. (2003). Negative feelings exacerbate hostile attributions of intent in highly aggressive boys. *Journal of Clinical Child and Adolescent Psychology*, *32*, 56–65. doi:10.1207/s15374424jccp3201_06

Dodge, K. A. (1980). Social cognition and children's aggressive-behavior. *Child Development*, *51*, 162–170. doi:10.2307/1129603

Dodge, K. A. (1986). A social information processing model of social competence in children. In M. Perlmutter (Ed.), *Cognitive perspectives on children's social and behavioral development*. *Minnesota symposia on child psychology*, 27, 18, (pp. 77–125). Hillsdale, NJ: Erlbaum. doi:10.1017/cbo9780511752834.013

Dodge, K. A. (2006). Translational science in action: Hostile attributional style and the development of aggressive behavior problems. *Development and Psychopathology*, *18*, 791–814. doi:10.1017/S0954579406060391

Dodge, K. A., & Frame, C. L. (1982). Social cognitive biases and deficits in aggressive boys. *Child Development*, *53*, 620–635. doi:10.2307/1129373

Dodge, K. A., & Pettit, G. S. (2003). A biopsychosocial model of the development of chronic conduct problems in adolescence. *Developmental Psychology*, *39*, 349. doi:10.1037//0012-1649.39.2.349

Dodge, K. A., Pettit, G. S., McClaskey, C. L., & Brown, M. M. (1986). *Social competence in children*. Monographs of the Society for Research in Child Development, *51*. Hoboken, NJ: John Wiley. doi:10.2307/1165906

Dodge, K. A., & Somberg, D. R. (1987). Hostile attributional biases among aggressive boys are exacerbated under conditions of threats to the self. *Child Development*, *58*, 213–224. doi:10.2307/1130303

Elder, G., Van Nguyen, T., & Caspi, A. (1985). Linking family hardship to childrens lives. *Child Development*, *56*, 361–375. doi:10.2307/1129726

Ethics and Behavior (2005). Special section: Deception and observation. *Ethics and Behavior*, *15*(3).

Facchin, F., Margola, D., Molgora, S., & Revenson, T. A. (2013). Effects of benefit-focused versus standard expressive writing on adolescents' self-concept during the high school transition. *Journal of Research on Adolescence*. Advance online publication. doi:10.1111/jora.12040.

Hankinson, J. R. (1998). *Cause and explanation in ancient Greek thought*. P. Guyer & A. W. Wood, Ed. and Trans. Oxford, UK: Oxford University Press.

Hinshaw, S. P. (2005). Objective assessment of covert antisocial behavior: Predictive validity and ethical considerations. *Ethics and Behavior*, *15*, 259–269. doi:10.1207/s15327019eb1503_6

Horsley, T. A., De Castro, B. O., & Van der Schoot, M. (2010). In the eye of the beholder: Eye-tracking assessment of social information processing in aggressive behavior. *Journal of Abnormal Child Psychology*, *38*, 587–599. doi:10.1007/s10802-009-9361-x

Kant, I. (1999/1781). *Critique of pure reason*. P. Guyer & A. W. Wood, Ed. and Trans. Cambridge, UK: Cambridge, UK: Cambridge University Press.

Karna, A., Voeten, M., Little, T. D., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2011). A large-scale evaluation of the KiVa antibullying program: Grades 4–6. *Child Development*, *82*, 311–330. doi:10.1111/j.1467-8624.2010.01557.x

Kempes, M. M., De Castro, B. O., & Sterck, E. H. M. (2008). Conflict management in 6–8-year-old aggressive dutch boys: Do they reconcile? *Behaviour*, *145*, 1701–1722. doi:10.1163/156853908786131306

Kuterovac-Jagodic, G. (2003). Posttraumatic stress symptoms in Croatian children exposed to war: A prospective study. *Journal of Clinical Psychology*, *59*, 9–25. doi:10.1002/jclp.10114

Lacourse, E., Cote, L., 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. doi:10.1017/s0954579402004121

Lansford, J. E., Malone, P. S., Dodge, K. A., Crozier, J. C., Pettit, G. S., & Bates, J. E. (2006). A 12-year prospective study of patterns of social information processing problems and externalizing behaviors. *Journal of Abnormal Child Psychology*, *34*, 715–724. doi:10.1007/s10802-006-9057-4

Lansford, J. E., Malone, P. S., Dodge, K. A., Pettit, G. S., & Bates, J. E. (2010). Developmental cascades of peer rejection, social information processing biases, and aggression during middle childhood. *Development and Psychopathology*, *22*, 593–602. doi:10.1017/s0954579410000301

Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development*, *71*, 107–118. doi:10.1111/1467-8624.00124

Leventhal, T., Fauth, R., & Brooks-Gunn, J. (2005). Neighborhood poverty and public policy: A 5-year follow-up of children's educational outcomes in the New York City Moving to Opportunity demonstration. *Developmental Psychology*, *41*, 933–952. doi:10.1037/0012-1649.41.6.933

Lochman, J. E., & Dodge, K. A. (1998). Distorted perceptions in dyadic interactions of aggressive and nonaggressive boys: Effects of prior expectations, context, and boys' age. *Development and Psychopathology*, *10*, 495–512. doi:10.1017/s0954579498001710

Lochman, J., & Wells, K. (2002). Contextual social-cognitive mediators and child outcome: A test of the theoretical model in the coping power program. *Development and Psychopathology*, *14*, 945–967. doi:10.1017/s0954579402004157

Montagne, B., Kessels, R. P. C., De Haan, E. H. F., & Perrett, D. I. (2007). The emotion recognition task: A new paradigm to study the perception of facial emotional expressions at different intensities. *Perceptual and Motor Skills*, *104*, 589–598. doi:10.2466/pms.104.2.589-598

Patterson, G., Dishion, T., & Bank, L. (1984). Family-interaction: A process model of deviancy training. *Aggressive Behavior*, *10*, 253–267. doi:10.1002/1098-2337(1984)10:3<253::aid-ab2480100309>3.0.co;2-2

Peterson, B. S. (2013). Editorial: From correlations to causation: The value of preventive interventions in studying pathogenic mechanisms in childhood psychiatric disorders. *Journal of Child Psychology and Psychiatry*, *54*, 813–815. doi:10.1111/jcpp.12122

Project Multisite Violence Prevention. (2013). Targeting high-risk, socially influential middle school students to reduce aggression: Universal versus selective preventive intervention effects. *Journal of Research on Adolescence*. Advance online publication. doi:10.1111/jora.12067

Reijntjes, A., Thomaes, S., Kamphuis, J. H., Bushman, B. J., De Castro, B. O., & Telch, M. J. (2011). Explaining the paradoxical rejection-aggression link: The mediating effects of hostile intent attributions, anger, and decreases in state self-esteem on peer rejection-induced aggression in youth. *Personality and Social Psychology Bulletin*, *37*, 955–963. doi:10.1177/0146167211410247

Reijntjes, A., Thomaes, S., Kamphuis, J. H., De Castro, B. O., & Telch, M. J. (2010). Self-verification strivings in children holding negative self-views: The mitigating effects of a preceding success experience. *Cognitive Therapy and Research*, *34*, 563–570. doi:10.1007/s10608-009-9289-z

Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior*, *15*, 112–120. doi:10.1016/j.avb.2009.08.007

Sandstrom, M. J., & Herlan, R. D. (2007). Threatened egotism or confirmed inadequacy? How children's perceptions of social status influence aggressive behavior toward peers. *Journal of Social and Clinical Psychology*, *26*, 240–267. doi:10.1521/jscp.2007.26.2.240

Smith, M., Hubbard, J. A., & Laurenceau, J. (2011). Profiles of anger control in second-grade children: Examination of self-report, observational, and physiological components. *Journal of Experimental Child Psychology*, *110*, 213–226. doi:10.1016/j.jecp.2011.02.006

Snoek, H., Van Goozen, S., Matthys, W., Buitelaar, J., & Van Engeland, H. (2004). Stress responsivity in children with externalizing behavior disorders. *Development and Psychopathology*, *16*, 389–406. doi:10.1017/s0954579404044578

Steenbeek, H., & van Geert, P. (2008). An empirical validation of a dynamic systems model of interaction: Do children of different sociometric statuses differ in their dyadic play? *Developmental Science*, *11*, 253–281. doi:10.1111/j.1467-7687.2007.00655.x

Thomaes, S., Bushman, B. J., De Castro, B. O., Cohen, G. L., & Denissen, J. J. A. (2009). Reducing narcissistic aggression by buttressing self-esteem: An experimental field study. *Psychological Science*, *20*, 1536–1542. doi:10.1111/j.1467-9280.2009.02478.x

Thomaes, S., Bushman, B. J., De Castro, B. O., & Reijntjes, A. (2012). Arousing gentle passions in young adolescents: Sustained experimental effects of value-affirmations on prosocial feelings and behaviors. *Developmental Psychology*, *48*, 103–110. doi:10.1037/a0025677

Thomaes, S., Bushman, B. J., Stegge, H., & Olthof, T. (2008). Trumping shame by blasts of noise: Narcissism, self-esteem, shame, and aggression in young adolescents. *Child Development*, *79*, 1792–1801. doi:10.1111/j.1467-8624.2008.01226.x

Thomaes, S., Reijntjes, A., De Castro, B. O., & Bushman, B. J. (2009). Reality bites—or does it? Realistic self-views buffer negative mood following social threat. *Psychological Science*, *20*, 1079–1080. doi:10.1111/j.1467-9280.2009.02395.x

Thomaes, S., Reijntjes, A., De Castro, B. O., Bushman, B. J., Poorthuis, A., & Telch, M. J. (2010). I like me if you like me: On the interpersonal modulation and regulation of preadolescents' state self-esteem. *Child Development*, *81*, 811–825. doi:10.1111/j.1467-8624.2010.01435.x

Tolan, P., Gorman-Smith, D., & Henry, D. (2003). The developmental ecology of urban males' youth violence. *Developmental Psychology*, *39*, 274–291. doi:10.1037/0012-1649.39.2.274

Underwood, M., Scott, B., Galperin, M., Bjornstad, G., & Sexton, A. (2004). An observational study of social exclusion under varied conditions: Gender and developmental differences. *Child Development*, *75*, 1538–1555. doi:10.1111/j.1467-8624.2004.00756.x

Van Geert, P., & Steenbeek, H. (2005). Explaining after by before: Basic aspects of a dynamic systems approach to the study of development. *Developmental Review*, *25*, 408–442. doi:10.1016/j.dr.2005.10.003

Van Lier, P. A. C., Vuijk, P., & Crijnen, A. A. M. (2005). Understanding mechanisms of change in the development of antisocial behavior: The impact of a universal intervention. *Journal of Abnormal Child Psychology*, *33*, 521–535. doi:10.1007/s10802-005-6735-7

Van Nieuwenhuijzen, M., Bijman, E. R., Lamberix, I. C. W., Wijnroks, L., De Castro, B. O., Vermeer, A., & Matthys, W. (2005). Do children do what they say? Responses to hypothetical and real-life social problems in children with mild intellectual disabilities and behaviour problems. *Journal of Intellectual Disability Research*, *49*, 419–433. doi:10.1111/j.1365-2788.2005.00674.x

Williams, K., & Jarvis, B. (2006). Cyberball: A program for use in research on interpersonal ostracism and acceptance. *Behavior Research Methods*, *38*, 174–180. doi:10.3758/bf03192765

Yeager, D. S., Trzesniewski, K. H., Tim, K., Nokelainen, P., & Dweck, C. S. (2011). Adolescents' implicit theories predict desire for vengeance after peer conflicts: Correlational and experimental evidence. *Developmental Psychology*, *47*, 1090–1107. doi:10.1037/a0023769