



# Innovation and Integrity: Desiderata and Future Directions for Prevention and Intervention Science

Andreas Beelmann<sup>1</sup> · Tina Malti<sup>2</sup> · Gil G. Noam<sup>3</sup> · Simon Sommer<sup>4</sup>

© Society for Prevention Research 2018

This is a peer-reviewed version of the following article: Beelmann, A., Malti, T., Noam, G.G., & Sommer, S. (2018). Innovation and integrity: Desiderata and future directions for prevention and intervention science. *Prevention Science*. doi.org/10.1007/s11121-018-0869-6, which has been published by Springer. The final publication is available at Springer via <https://doi.org/10.1007/s11121-018-0869-6>. Please refer to Springer Terms and Conditions of Archiving for more information: <http://www.springer.com/gp/open-access/authors-rights/self-archiving-policy/2124>

## Abstract

This article summarizes essential implications of the papers within this special issue and discusses directions for future prevention and intervention research on conceptual issues, methodological and transfer-related challenges and opportunities. We identify a need to move from programs to principles in intervention research and encourage the implementation of research on potential mechanisms underlying intervention effectiveness. In addition, current methodological issues in intervention research are highlighted, including advancements in methodology and statistical procedures, extended outcome assessments, replication studies, and a thorough examination of potential biases. We further discuss transfer-related issues, for example the need for more research on the flexibility and adaptability of programs and intervention approaches as well as more general problems in knowledge translation reasoning the need for enhanced communication between practitioners, policy makers, and researchers. Finally, we briefly touch on the need to discuss the relation between single intervention programs, the mental health system, and changes of contextual conditions at the macro level.

**Keywords** Research integrity · Child and adolescent mental health · Evidence-based intervention · Knowledge translation · Intervention theory

The last 50 years have seen enormous progress in prevention and intervention research, including hundreds of high-quality evaluations studies; various evidence-based programs across diverse fields for promoting physical health, mental health, and positive development of children, adolescents, and their families; active prevention and intervention communities around the globe; and countless initiatives to implement prevention and

intervention programs into the routine health and educational systems (see e.g., O'Connell et al. 2009). These are all really strong markers of an innovative and impactful research field. Hence, could a preliminary conclusion simply be: So far so good?

Our answer is: Yes and no. Yes, because the field has established itself as a productive area of research with important implications for practices aimed at creating better lives for children, adolescents, and adults. No, because we still face significant challenges. These include the quality and integrity of prevention and intervention research, the effectiveness in real-world practice settings, and implementation of evidence-based policies. For example, epidemiological studies show constantly high prevalence rates of emotional and behavior problems across childhood and adolescence (see Belfer 2008; Reiss 2013) despite all efforts to reduce mental health problems and promote healthy development in prevention and intervention design and implementation over the last decades. The aim of this article is to summarize the articles of this special issue, discuss some of these conceptual, methodological, and transfer-related desiderata, and identify promising areas and opportunities for future work in these areas.

---

✉ Andreas Beelmann  
andreas.beelmann@uni-jena.de

<sup>1</sup> Institute of Psychology, Department of Research Synthesis, Intervention, Evaluation, Friedrich-Schiller-University Jena, Humboldtstrasse 26, 07743 Jena, Germany

<sup>2</sup> Department of Psychology and Department of Psychiatry, University of Toronto, Mississauga, Canada

<sup>3</sup> Harvard Medical School, Harvard University, Boston, MA, USA

<sup>4</sup> Jacobs Foundation, Zürich, Switzerland

## Conceptual Issues

There has been significant progress in the theoretical foundation of prevention programs and intervention approaches in the areas of child and youth development, education, and health. Many successful programs have relied on developmental theory and research (e.g., Durlak et al. 2011; Malti et al. 2016a; Malti and Noam 2016; Wilson et al. 2014)—marking this movement as one of the most noteworthy improvements over the last decades. Nonetheless, despite the fact that many programs are evidence-based and rooted in developmental work, the term evidence-based is still mostly restricted to outcome evaluation and typically not applied to the theoretical and conceptual intervention foundation (see Beelmann 2011). Beside the problem of which exact results should build the basis for an evidence-based status, this neglects important aspects of the scientific foundation of programs such as an evidence-based legitimation (e.g., is a new intervention really needed?), the use of a validated change model (e.g., Prochaska et al. 2015), and the foundation of intervention content (*what* should be changed or promoted?) and implementation concept (*how* should it be changed or promoted?) according to developmental principles and intervention research (Beelmann 2011).

In addition, meta-analytic results reveal that, over the last 30 years, prevention and intervention science has focused mainly on brand name programs and less on prevention or intervention principles. As a result, we now have dozens of programs within each prevention or intervention field that are essentially similar to each other. Keeping this situation in mind, would it not be better to rely or focus on what their essential elements and principles are? For example, prevention meta-analyses confirm that interactive learning is better than providing enlightenment simply via information giving (e.g., Tobler et al. 2000). Other authors have established lists about what they have called intervention kernels (Embry and Biglan 2008), i.e., methods that could be used in interventions because they have been proven to be effective over various applications. Hence, there is a need for further research on sound prevention and intervention principles that are independent from concrete application, rather than a need for an extended list of programs, especially when several programs are already available. However, principles are more abstract by nature and need more competencies and professional education to transfer these principles to concrete application in certain cases.

## Methodological Challenges and Opportunities

From a methodological standpoint, the current state of research in prevention and intervention science is characterized by a high number of randomized trials, extended knowledge

in data analysis and design issues for evaluation research, and an impressive number of high-quality reviews and meta-analyses summarizing the results of international prevention and intervention efforts (see, Farrington et al. 2016, for the prevention of antisocial behavior). However, while these achievements are impressive, methodological challenges remain and opportunities for future research are emerging.

The first challenge concerns the *use* of the full range of methodological and statistical innovations and opportunities in intervention research. For example, Lang and Little (2016) show that prevention science needs to improve the elaborated use of missing data techniques—especially as these techniques have now been available for a relatively long time. Other methods not used sufficiently are multilevel analysis, latent class models, and propensity score matching, to mention a few. Especially when it comes to the roll-out of interventions within large-scale research, traditional ways of performing data analysis should at least be complemented by new, innovative procedures, such as using archival measures and large-scale social data sets on a societal level (e.g., crime or health statistics) to evaluate the outcomes of programs and interventions.

This is also true for an extended use of a greater variety of research designs beyond randomized controlled trials (RCTs) and the use of additional control strategies for selection bias especially for large-scale dissemination studies (see Hallberg et al. 2016; Henry et al. 2017; Ghate 2016; Pawson and Tilley 1997). Journal editors have a responsibility here, namely that they have to find reviewers who are willing and qualified to embrace methodological innovations applied in manuscripts they review.

A second issue concerns the assessment of outcomes in prevention and intervention studies. Several meta-analyses have shown that distal and “hard” outcomes that correspond to the real prevention and intervention targets are mostly a rarity especially when looking at long-term effects. More generally, there appears to be an overrepresentation of proximal outcome assessments that may result in misleading effect size approximations. For example, crime prevention studies have often assessed proximal risk or protective factors such as social-cognitive information processing (e.g., Beelmann and Lösel 2006) or parenting skills (Beelmann et al. 2017) but not official crime. Of course, short-term effects on proximal outcomes are (in most cases) a necessary condition for long-term effects on distal outcomes. In addition, long-term evaluations are time consuming, expensive, and harbor a risk of failing to verify long-term effects. Thus, we recommend a more face valid measurement of prevention and intervention effects in future work, such as incidence rates or percentages of healthy children and adolescents. This would also strengthen the reputation of the field outside of science and enhance a more integrated, successful and sustainable knowledge transfer to practice and policy (see Wathen and MacMillan 2015).

A related challenge is the relatively rare use of multi-method and multi-informant assessments. For example, a meta-analysis of parent training programs (Beelmann et al. 2017) showed that more than 80% of all dependent variables are derived from information provided by the participating mothers and fathers via questionnaires. In future work, it will therefore be important to utilize a comprehensive outcome measurement framework (e.g., multi-method, multi-informant data sets). Without doubt, a more rigorous assessment approach will yield valuable information when testing and specifying the effects of any intervention program or approach.

Third, like in other fields, there are still diverse sources of bias within prevention and intervention science (Ioannidis et al. 2014). Prominent and often related issues are publication bias and conflict of interest (see Gorman 2016; Eisner 2009; Rothstein et al. 2005). These biases pose serious concerns because they inherently relate to customs in the publication system, to norms within the scientific community, and to the corresponding needs and requirements of individual scientists' careers (Bakker et al. 2012). Of course, the standardization of reporting (e.g., APA, CONSORT; see APA Publications and Communication Board Working Group on Journal Article Reporting Standards 2008; Gottfredsson et al. 2015), as well as trial registers (see Nosek and Lakens 2014; and <https://clinicaltrials.gov>; [www.clinicaltrialsregister.eu](http://www.clinicaltrialsregister.eu), for examples), have contributed greatly to increasing the transparency and credibility of prevention and intervention findings, although these measures do not have the same relevance in psychosocial intervention science yet compared to that found in medical science. However, even these measures cannot fully eliminate conflict of interest and biased data analyses. In fact, alongside intentional data manipulation, there are hundreds of ways to arrange data in a specific manner and select those presentations that best fit one's own interests (see Simmons et al. 2011). Unfortunately, one can only remind scientists of their commitment to integrity and objectivity and point out how unexpected data are usually of greater heuristic value than expected data. For example, for future prevention and intervention research, we should put high emphasis on the publication of null results, negative findings, or—in general terms—non-confirmative findings for at least four reasons: First, the non-publication of negative results leads to biased outcome summaries of a special program or intervention approach (Kepes et al. 2014). Second, because it is unlikely that interventions will be always or totally successful, the presentation of null results correspond particularly to practical experience and, as a result, can receive more credibility. This is of major importance when it comes to transferring and implementing scientific knowledge (see below). Third, bearing the principle of falsification in mind (see Earp and Trafimow 2015), null results and negative findings are probably of higher value, especially when it comes to testing a program's underlying theoretical model. Fourth, in more practical terms, null

results and negative findings are likely to encourage more in-depth reflection about why a special program (although already positively tested) has no or even negative effects. Ultimately, this may turn out to be more productive than a very simple statement that a program or approach works as initially hypothesized. Therefore, especially reviewers and editors should be encouraged to appreciate the value of articles that try to publish “bad news” and be open to what can be even more inspiring and important for the advancement of science.

Fourth, and relatedly, the current state of replication studies warrants discussion. These studies are generally an important means to control and confirm existing results and are especially needed in prevention and intervention science because they especially strengthen the generalizability of results (Earp and Trafimow 2015; Nosek and Lakens 2014; Valentine et al. 2011). However, conducting replication studies are frequently of limited attractiveness (Everett and Earp 2015) because they are harder to publish and do not often enhance reputations as much as the publication of original research. These obstacles together with the “replication crisis” have cast much doubt on the credibility of psychological science in recent years (Lilienfeld 2017). This, in turn, can negatively impact the status of science when it comes to decision-making in practice and policies because practitioners and politicians do not expect heterogeneity in scientific results (see Bromme and Beelmann 2016).

We urge researchers to consider conducting replication studies. Like the publication of null findings or negative effects, replications are truly needed for both theoretical and practical reasons. Nevertheless, alongside asking what is meant by replication and which function it fulfills (Aos et al. 2011), there are still further questions such as whether ten at least comparable studies, each with a sample size of 100 participants, provide a better evidence base for a program than one large-scale study with 1000 participants. Again, from the standpoint of critical positivism and falsification, the first alternative would be better, because we can run several tests of falsification and generalization across contexts, samples, and conditions of research. High-quality meta-analyses can then be used to summarize these results in an appropriate manner. In addition, large single studies often suffer from idiosyncratic characteristics (e.g., a certain set of variables and specific contextual conditions) and methodological problems, such as higher requirements for a successful implementation. Therefore, doing prevention and intervention replications is one of the most needed pathways to future intervention research, even if it is time consuming and dependent on sophisticated scientific networks and resources.

Finally, when an intervention should be viewed as evidence-based? What are the thresholds for recommending one program, one special type of program, or interventions in a field in general? High variability and inconsistent findings even in established programs (Gruner Gandhi et al. 2007;

Wilson et al. 2012) can make it challenging for policy makers and practitioners to understand the nature of scientific results. In addition, independent testing of program effectiveness mostly yield considerable lower effect size estimations compared to evaluations where program originators are involved (Eisner and Humphreys 2011). Although, these results can be interpreted differently (e.g., high implementation quality guaranteed by originators vs. conflict of interest), the issue remains that the label “evidence-based” does not refer to a unified characteristic of a program or intervention.

Of course, evidence-based program registers are widely disseminated (Burkhardt et al. 2015) and of great value for practitioners and policymakers. In addition, sophisticated guidelines for conducting sound prevention research facilitate its evaluation according to quality standards (Gottfredsson et al. 2015). Yet, even these lists and guidelines often still suffer from subjectivity when defining the criteria for evidence (Gruner Gandhi et al. 2007; Means et al. 2015) and they do not give specific thresholds for evidence. From our perspective, it is somewhat hard to imagine that it will be possible to find an objective solution for all stakeholders, i.e., generalizable, yet concrete outcome standards to assess efficacy, effectiveness, and dissemination quality. Because intervention outcomes are a result of a complex interaction between characteristics of the intervention, the sample, methodological features of the study, and the type and method of dependent variables, it will be rather difficult to condense this complexity to single numbers. And even if we can define a unified threshold for outcomes, “small” effects may be impressive when the intensity of a treatment or the necessary efforts and investments required to deliver it are low (Prentice and Miller 1992). In general, effect sizes should be interpreted with respect to empirical benchmarks that are meaningful in the context of the intervention domain and the nature of the outcomes being examined (e.g., Hill et al. 2008).

Therefore, future research requires a serious discussion about what should be realistically and reasonably expected from prevention and intervention programs. Of course, the negative endpoint of an evaluative scale is reached when programs have no positive evaluation outcome. But even this can be rather fuzzy. What is meant by no positive evaluation outcome? Is a simple pre–post study worthless when there is no other source of information? Most researchers would likely agree, but how would practitioners or target groups respond to this question? In the future, it will become increasingly important for scientists to make the public aware of this complex decision-making process and offer better explanations (e.g., via more practical parameters such as the probability of change) for *why* a concrete program or type of program is to be recommended or not, without exaggerating effects and concealing problems in outcome evaluations.

## Transfer-Related Challenges and Opportunities

Within the last decade, knowledge transfer and implementation have become important topics within prevention and intervention science. We now possess real extended knowledge and models to improve implementation (see Meyers et al. 2012; Wathen and MacMillan 2015), but comprehensive research on the routine implementation of programs within the social and health system is still at the beginning, especially with respect to outcome results on the societal level using population trials. In addition, despite all efforts within the last decade, prevention services are still established mainly as pilot or temporary applications without being integrated into country-wide routine services. Furthermore, the appreciation and perceived public interest in prevention and intervention within the whole society is not overwhelming particularly when compared to international concerns such as right-wing populism, the financial crisis, job loss, or threats of terrorism (Bromme and Beelmann 2016). Therefore, transfer and implementation will remain a constant mission for prevention and intervention science with at least three important aspects:

First, as in the case of the methodological challenges, we have to press for the existing knowledge on transfer and implementation to be applied in modern concepts of prevention and intervention. For example, only a few prevention programs contain an explicit large-scale implementation concept as an integrative part of their approach such as Spiel et al. (2016) have presented. A modern implementation concept contains at least three different facets—a broader concept of dissemination and delivery on a societal level, an implementation concept within each institution and facility, and finally, a way to deal with deviations from the optimal conduct of programs that stem from less than optimal conditions and limited resources.

Second, it still remains a central ambition to communicate scientific knowledge to the public, society, and policymakers (Wathen and MacMillan 2015). However, as Bromme and Beelmann (2016) have shown, this is much more complicated than we usually think. Good arguments (e.g., evidence, positive cost–benefit ratios) are probably not enough. Instead, we have to assume that scientific knowledge and theories are only one—and often, unfortunately, less important—information source for the public, the media, society, and public policymakers. And making policy decisions is different from scientific decision-making because the criteria are not the same (e.g., successful negotiation between stakeholders vs. objectivity and validity). As was discovered long ago, research on the use of evaluation shows that multiple factors affect a successful transfer from science to practice (e.g., Leviton and Hughes 1981), and we have to acknowledge that this requires permanent efforts to convince people, society, and public policy. Science, as a prominent information source

for decision-making, needs to improve its communication with policymakers, practitioners, and the media; for example, by building effective partnerships, as Wathen and MacMillan (2015) have suggested. However, this requires a basic interest in science-based knowledge on the side of the recipients which is, at least in some cases, questionable (Lilienfeld 2012).

Third, another sometimes neglected issue is the transfer from scientific knowledge to practical applications in concrete cases of prevention and intervention programs. There are several reasons why this transfer is difficult and needs constant consideration. For example, we should be aware that scientific results (at least when they come from empirical data) are the outcome of statistically tested mean values whose validity can be questionable in individual cases. Science should help practitioners accomplish this transfer process. One element could be to guide intervention planning with a systematic individualized assessment (see Malti et al. 2017). Another, even bigger topic is the need for cultural adaptations (Okamoto et al. 2014). Current research (e.g., Baumann et al. 2015; Castro et al. 2010; Castro and Yasui 2017, and special issue in *Prevention Science* 6/2017) is concluding that our programs lack cultural sensitivity and need to be modified for applications to different target groups. Meta-analytic data has shown that cultural adaptations normally lead to higher effects of prevention programs (Sundell et al. 2016). Therefore, and in more general terms, prevention and intervention concepts have to be directed actively toward delivering ideas on how to adapt a concept to specific applications and how to individualize the application to improve the implementation process. However, up to now adaptations of evidence-based programs are mainly a result of logistical issues and were made reactively, at least in natural contexts (Moore et al. 2013). In addition, adaptability and flexibility are not the natural antagonism to fidelity if the applicant knows what could and what should be adapted and used with flexibility, and what should necessarily remain unchanged (Castro and Yasui 2017). For example, as Resnicow et al. (2000) have pointed out that changes within the so-called surface structure of programs (e.g., intensity, intervention methods, and materials used) are possible and sometimes probably necessary, but changes in the so-called deep structure of programs (e.g., theoretical underpinnings) are most likely leading to less implementation quality and lower effectiveness. However, if practitioners are unable to decide what to use with flexibility and how to adapt intervention concepts according to specific target groups, they may be at-risk for unreflecting or at-hoc changes that probably lead to serious deviation from the essence of the intervention (see Molloy et al. 2013). Therefore, future prevention and intervention research should invest more on the adaptability of interventions and on opportunities and necessities of using manuals and concepts with flexibility.

## Outlook

In this paper, we have argued that future generation of prevention and intervention research will need to systematically integrate theoretical models of change and use innovative methodological approaches to move the field forward. Integrated, comprehensive knowledge transfer is needed to create sustainable effects. At the same time, researchers and the scientific community have to remind themselves that these ambitions can only be fulfilled if we are oriented toward the ultimate goal of prevention and intervention research, namely to deliver valid unbiased basic and applied knowledge that result in models of interventions that are applicable and effective.

Of course, prevention programs and intervention approaches are only one aspect of mental health systems (see Malti et al. 2016b). Therefore, it remains an open question if individual programs are enough, or if we need to restructure components of our mental health systems. In a current paper, Malti et al. (2016b) have argued that we need to move from programs to systems in order to deliver better services to children, adolescents, and families (see also Ghate 2016). This calls for some final reflection on the larger theme: In modern society, many examples point to the fact that a sole focus on individual competence promotion is limited because macro-level influences interfere with preventive efforts and can limit their effectiveness, for example, in the prevention of negative consequences of poverty (Beelmann et al. 2015). The same may well also apply to other areas of prevention and intervention. For example, child and adolescent obesity are highly increasing in Western societies (e.g., Ogden et al. 2010). A simple question could then be: Should we initiate changes to the delivery of food (by e.g., stricter regulations, better labeling of food, or prohibition of advertising for unhealthy food), or should we promote individual competencies to cope with the (unhealthy) food products on the market. At first glance, this seemingly simple question reveals fundamental conflicts about the values, norms, and aims within our society. More generally, we may ask: Are individual programs enough or do we need a change of conditions even at the societal level? Biglan (2016) and others (Wilson et al. 2014) argue for so-called nurturing environments, an approach that is oriented systematically toward positive youth development and what children and adolescents need to grow up successfully, both in terms of productivity and as ethically responsible citizens (see also Komro et al. 2011). This may also require societal changes if conditions prove to be systematically responsible for individual social and health problems. For example, systematic comparisons between developed societies have shown that social inequality is a major risk factor for a wide range of social and health problems (see Wilkenson and Pickert 2009). Thus, it remains an open question if, and to what extent, we need to change the roots of social inequality before we start to engage in prevention of its negative consequences

for youth health. This also holds for future prevention and intervention research that will have to face new challenges in a fast changing world, including global violence, unprecedented numbers of refugees, worldwide immigration, and political and religious radicalization and terrorism.

**Funding** The second author was funded, in part, by a New Investigator Award from the Canadian Institutes of Health Research.

## Compliance with Ethical Standards

**Conflict of Interest** The authors declare that they have no conflict of interest.

**Ethical Approval** This article does not contain any studies with human participants or animals performed by any of the authors.

**Informed Consent** Informed consent is not applicable.

## References

- Aos, S., Cook, T. D., Elliott, D. S., Gottfredson, D. C., Hawkins, J. D., Lipsey, M. W., & Tolan, P. (2011). Commentary on Valentine, Jeffrey, et al. Replication in prevention science. *Prevention Science, 12*, 121–122. <https://doi.org/10.1007/s11121-011-0219-4>.
- APA Publications and Communication Board Working Group on Journal Article Reporting Standards. (2008). Reporting standards for research in psychology. Why do we need them? What might they be? *American Psychologist, 63*, 839–851. <https://doi.org/10.1037/0003-066X.63.9.839>.
- Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science, 7*, 543–554. <https://doi.org/10.1177/1745691612459060>.
- Baumann, A. A., Powell, B. J., Kohl, P. L., Tabak, R. G., Penalba, V., Protor, E. K., Domenech-Rodriguez, M. M., & Cabassa, L. J. (2015). Cultural adaptation and implementation of evidence-based parent-training: A systematic review and critique of guiding evidence. *Children and Youth Services Review, 53*, 113–120. <https://doi.org/10.1016/j.childyouth.2015.03.025>.
- Beelmann, A. (2011). The scientific foundation of prevention. The status quo and future challenges of developmental crime prevention. In T. Bliesener, A. Beelmann, & M. Stemmler (Eds.), *Antisocial behavior and crime. Contributions of developmental and evaluation research to prevention and intervention* (pp. 137–164). Cambridge: Hogrefe Publishing.
- Beelmann, A. (2017). The effects of parent training programmes in the prevention and treatment of antisocial behavior in childhood and adolescence. An international comprehensive meta-analysis. Paper presented at the What Works Global Summit - Conference, London, UK.
- Beelmann, A., Maichrowitz, S., Schulz, S. & Arnold, L. (2015). Wirksamkeit von Programmen zur Prävention und Kompensation von Armut- und Migrationsfolgen bei Kindern und Jugendlichen. Ein Forschungsüberblick und eine Meta-Analyse internationaler Evaluationsforschung [The effectiveness of programs for buffering negative consequences of poverty and migration in children and adolescents. A review and a meta-analysis of the international evaluation research]. Final report to the Federal Ministry of Education and Research, Berlin, Germany.
- Belfer, M. L. (2008). Child and adolescent mental disorders: the magnitude of the problem across the globe. *Journal of Child Psychology and Psychiatry, 49* (3), 226–236. <https://doi.org/10.1111/j.1469-7610.2007.01855.x>
- Biglan, A. (2016). The ultimate goal of prevention and the larger context for translation. *Prevention Science. https://doi.org/10.1007/s11121-016-0635-6*.
- Bromme, R., & Beelmann, A. (2016). Transfer entails communication: The public understanding of (social) science as a stage and a play for implementing evidence-based prevention knowledge and programs. *Prevention Science. https://doi.org/10.1007/s11121-016-0686-8*.
- Burkhardt, J. T., Schröter, D. C., Magura, S., Means, S. N., & Coryn, C. L. S. (2015). An overview of evidence-based program registers (EBPRs) for behavioral health. *Evaluation and Program Planning, 48*, 92–99. <https://doi.org/10.1016/j.evalprogplan.2014.09.006>.
- Castro, F. G., & Yasui, M. (2017). Advances in EBI development for diverse populations: Towards a science of intervention adaption. *Prevention Science, 18*, 623–629. <https://doi.org/10.1007/s11121-017-0809-x>.
- Castro, F. G., Barrera Jr., M., & Holleran-Steiker, L. K. (2010). Issues and challenges in the design of culturally adapted evidence-based interventions. *Annual Review of Clinical Psychology, 6*, 213–239. <https://doi.org/10.1146/annurev-clinpsy-033109-132032>.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schnellinger, K. B. (2011). The impact of enhancing students' social and emotional learning. A meta-analysis of school-based universal interventions. *Child Development, 82*, 405–432. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>.
- Earp, B. D., & Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology, 6*, article 621. <https://doi.org/10.3389/psyg.2015.00621>.
- Eisner, M. (2009). No effects in independent prevention trials: Can we reject the cynical view? *Journal of Experimental Criminology, 5*, 163–183. <https://doi.org/10.1007/s11292-009-9071-1>.
- Eisner, M., & Humphreys, D. (2011). Measuring conflict of interest in prevention and intervention research. A feasibility study. In T. Bliesener, A. Beelmann, & M. Stemmler (Eds.), *Antisocial behavior and crime. Contributions of developmental and evaluation research to prevention and intervention* (pp. 165–180). Cambridge: Hogrefe.
- Embry, D. D., & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clinical Child and Family Review, 11*, 75–113. <https://doi.org/10.1007/s10567-008-0036-x>.
- Everett, J. A. C. & Earp, B. D. (2015). A tragedy of the (academic) commons: Interpreting the replication crisis in psychology as a social dilemma for early-career researchers. *Frontiers in Psychology, 6*, article 1152. <https://doi.org/10.3389/fpsyg.2015.01152>.
- Farrington, D. P., Gaffney, H., Lösel, F., & Ttofi, M. M. (2016). Systematic reviews of the effectiveness of developmental prevention programs in reducing delinquency, aggression, and bullying. *Aggression and Violent Behavior, 33*, 91–106. <https://doi.org/10.1016/j.avb.2016.11.003>.
- Ghate, D. (2016). From programs to systems: Deploying implementation science and practice for sustained real world effectiveness in services for children and families. *Journal of Clinical Child and Adolescent Psychology, 45*, 812–826. <https://doi.org/10.1080/15374416.2015.1077449>.
- Gorman, D. M. (2016). Can we trust positive findings of intervention research? The role of conflict of interest. *Prevention Science. https://doi.org/10.1007/s11121-016-0648-1*.
- Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., Sandler, I. N., & Zafft, K. M. (2015). Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: Next generation. *Prevention Science, 16*, 893–926. <https://doi.org/10.1007/s11121-015-0555-x>.
- Gruner Gandhi, A., Murphy-Graham, E., Petrosino, A., & Weiss, C. H. (2007). The devil is in the detail. Examining the evidence for “proven” school-based drug abuse prevention programs.

- Evaluation Review*, 31, 43–74. <https://doi.org/10.1177/019841X0687188>.
- Hallberg, K., Cook, T. D., Steiner, P. M., & Clark, M. H. (2016). Pretest measures of the study outcome and the elimination of selection bias: Evidence from three within study comparisons. *Prevention Science*. <https://doi.org/10.1007/s11121-016-0732-6>.
- Henry, D., Tolan, P., Gorman-Smith, D., & Schoeny, M. (2017). Alternatives to randomized control trial designs for community-based prevention evaluation. *Prevention Science*, 18, 671–680. <https://doi.org/10.1007/s11121-016-0706-8>.
- Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2, 172–177.
- Ioannidis, J. P. A., Munafo, M. R., Fusar-Poli, P., Nosek, B. A., & David, S. P. (2014). Publication and other reporting biases in cognitive sciences: Detection, prevalence, and prevention. *Trends in Cognitive Science*, 18, 235–241. <https://doi.org/10.1016/j.tics.2014.02.010>.
- Kepes, S., Banks, G. C., & Oh, I.-S. (2014). Avoiding bias in publication bias research: The value of “null” findings. *Journal of Business and Psychology*, 29, 183–203. <https://doi.org/10.1007/s10869-012-9279-0>.
- Komro, K. A., Flay, B. R., Biglan, A., & Promise Neighborhoods Research Consortium. (2011). Creating nurturing environments: A science-based framework for promoting child health and development within high-poverty neighborhoods. *Clinical Child and Family Psychology Review*, 14, 111–134.
- Lang, K. M., & Little, T. D. (2016). Principled missing data treatments. *Prevention Science*. <https://doi.org/10.1007/s11121-016-0644-5>.
- Leviton, L. C., & Hughes, E. F. X. (1981). Research on utilization of evaluations. A review and synthesis. *Evaluation Review*, 5, 525–548.
- Lilienfeld, S. O. (2012). Public skepticism of psychology. Why many people perceive the study of human behavior as unscientific. *American Psychologist*, 67, 111–129. <https://doi.org/10.1037/a0023963>.
- Lilienfeld, S. O. (2017). Psychology’s replication crisis and the grant culture: Righting the ship. *Perspectives on Psychological Science*, 12, 660–664. <https://doi.org/10.1177/1745691616687745>.
- Malti, T., & Noam, G. G. (2016). Social-emotional development: From theory to practice. *The European Journal of Developmental Psychology*, 13, 652–665. <https://doi.org/10.1080/17405629.2016.1196178>.
- Malti, T., Chaparro, M. P., Zuffianò, A., & Colasante, T. (2016a). School-based interventions to promote empathy-related responding in children and adolescents: A developmental analysis. *Journal of Clinical Child and Adolescent Psychology*, 45, 718–731. <https://doi.org/10.1080/15374416.2015.1121822>.
- Malti, T., Noam, G. G., Beelmann, A., & Sommer, S. (2016b). Toward dynamic adaptation of psychological interventions of child and adolescent development and mental health. *Journal of Clinical Child and Adolescent Psychology*, 45, 827–836. <https://doi.org/10.1080/15374416.2016.1239539>.
- Malti, T., Zuffianò, A., & Noam, G. G. (2017). Knowing every child: Validation of the Holistic Student Assessment (HSA) as a measure of social-emotional development. *Prevention Science*. <https://doi.org/10.1007/s11121-017-0794-0>.
- Means, S. N., Magura, S., Burkhardt, J. T., Schröter, D. C., & Coryn, C. L. S. (2015). Comparing rating paradigms for evidence-based program registers in behavioral health: Evidentiary criteria and implications for assessing programs. *Evaluation and Program Planning*, 48, 100–116. <https://doi.org/10.1016/j.evalprogplan.2014.09.007>.
- Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology*, 50, 462–480. <https://doi.org/10.1007/s10464-012-9522-x>.
- Molloy, L. E., Moore, J. E., Trail, J., van Epps, J. J., & Hopfer, S. (2013). Understanding real-world implementation quality and “active ingredients” of PBIS. *Prevention Science*, 14, 593–605. <https://doi.org/10.1007/s11121-012-0343-9>.
- Moore, J. E., Bumbarger, B. K., & Cooper, B. R. (2013). Examining adaptations of evidence-based programs in natural contexts. *Journal of Primary Prevention*, 34, 147–161. <https://doi.org/10.1007/s10935-013-0303-6>.
- Nosek, B. A., & Lakens, D. (2014). Registered reports. A method to increase the credibility of published results. *Social Psychology*, 45, 137–141. <https://doi.org/10.1027/1864-9335/a000192>.
- O’Connell, M. E., Boat, T., & Warner, K. E. (Eds.). (2009). *Preventing mental, emotional, and behavioral disorders among young people. Progress and possibilities*. Washington, DC: The National Academies Press.
- Ogden, C. L., Carroll, M. D., Curtin, L. R., Lamb, M. M., & Flegal, K. M. (2010). Prevalence of high body mass index in US children and adolescents. *Journal of the American Medical Association*, 303, 242–249. <https://doi.org/10.1001/jama.2009.2012>.
- Okamoto, S. K., Kulis, S., Marsiglia, F. F., Steiker, L. K. H., & Dustman, P. (2014). A continuum of approaches toward developing culturally focused prevention interventions: From adaption to grounding. *Journal of Primary Prevention*, 35, 103–112. <https://doi.org/10.1007/s10935-013-0333-z>.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: Sage.
- Prentice, D. A., & Miller, D. T. (1992). When small effects are impressive. *Psychological Bulletin*, 112, 160–164. <https://doi.org/10.1037/0033-2909.112.1.160>.
- Prochaska, J. O., Redding, C. A., & Evers, K. E. (2015). The transtheoretical model and stages of change. In K. Glanz, B. K. Rimer, & K. Viswanath (Eds.), *Health behavior: Theory, research, and practice* (5th ed., pp. 125–148). San Francisco: Jossey-Bass.
- Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: A systematic review. *Social Science & Medicine*, 90, 24–31. <https://doi.org/10.1016/j.socscimed.2013.04.026>.
- Resnicow, K., Soler, R., Braithwaite, R. L., Ahluwalia, J. S., & Butler, J. (2000). Cultural sensitivity in substance use prevention. *Journal of Community Psychology*, 28, 271–290.
- Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.). (2005). *Publication bias in meta-analysis: Prevention, assessment, and adjustment*. Chichester: Wiley.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 1359–1366. <https://doi.org/10.1177/0956797611417632>.
- Spiel, C., Schober, B., & Strohmeier, D. (2016). Implementing intervention research into public policy—The “I<sup>3</sup>-approach.”. *Prevention Science*. <https://doi.org/10.1007/s11121-016-0638-3>.
- Sundell, K., Beelmann, A., Hasson, H., & von Thiele Schwarz, U. (2016). Novel programs, international adoptions, or contextual adaptations? Meta-analytical results from German and Swedish intervention research. *Journal of Clinical Child & Adolescent Psychology*, 45, 784–796. <https://doi.org/10.1080/15374416.2015.1020540>.
- Tobler, N. S., Roona, M. R., Ochshorn, P., Marshall, D. G., Streke, A. V., & Stackpole, K. M. (2000). School-based adolescent drug prevention programs: 1998 meta-analysis. *The Journal of Primary Prevention*, 20, 275–336. <https://doi.org/10.1023/A:1021314704811>.
- Valentine, J. C., Biglan, A., Boruch, R. F., Castro, F. G., Collins, L. M., Flay, B. R., Kellam, S., Mościcki, E. K., & Schinke, S. P. (2011). Replication in prevention science. *Prevention Science*, 12, 103–117. <https://doi.org/10.1007/s11121-011-0217-6>.
- Wathen, C. N., & MacMillan, H. L. (2015). The role of integrated knowledge translation in intervention research. *Prevention Science*. <https://doi.org/10.1007/s11121-015-0564-9>.

- Wilkenson, R. G., & Pickett, K. E. (2009). Income inequality and social dysfunction. *Annual Review of Sociology*, *35*, 493–511. <https://doi.org/10.1146/annurev-soc-070308-115926>.
- Wilson, P., Rush, R., Hussey, S., Puckering, C., Sim, F., Allely, C. S., Allely, C. S., Doku, P., & Gillberg, C. (2012). How evidence-based is an “evidence-based parenting program”? A PRISMA systematic review and meta-analysis of Triple P. *BMC Medicine*, *10*, 130. <https://doi.org/10.1186/1741-7015-10-130>.
- Wilson, D. S., Hayes, S. C., Biglan, A., & Embry, D. D. (2014). Evolving the future: Toward a science of intentional change. *Behavioral and Brain Sciences*, *37*, 395–460. <https://doi.org/10.1017/S0140525X13001593>.