Commentary on Kallestad and Olweus (2003)

Implementation as a Second Stage in Prevention Research

Marion S. Forgatch
Oregon Social Learning Center

ABSTRACT
Recognizing the protracted effort and rigorous research required to produce evidence-based programs, this commentary challenges prevention scientists to apply similar standards to the study of program implementation. The pioneering work of Dan Olweus and his colleagues is used to illustrate ways in which methodology common in efficacy research could be applied to implementation research.

The Olweus Bullying Prevention Program, first developed in Sweden in 1970 and evaluated in Norway in the early 1980s, represents a model program that has evolved over decades of programmatic research and has withstood the test of time. The bully/victim program stands on a foundation of epidemiological, basic, and applied research in which theory and practical application are thoughtfully integrated. In the years since its inception, several investigators, including Olweus and his colleagues (Kallestad & Olweus, 2003), have conducted replication studies of the bully/victim intervention. The high quality of this body
of work contributed to its selection by a U.S. expert committee as 1 of 10 blueprint programs to be implemented in many U.S. communities. The fact that even so careful a program of research can produce highly variable outcomes when tested in multiple communities and cultures underlines the absolute necessity of embarking on a second stage in the application of science to such social problems. Researchers must apply stringent standards to the study of implementation processes to understand why replication works in some conditions and not in others.

Implementation research can be undertaken only after a long journey has been well underway. Before dissemination can begin, rigorous epidemiological and longitudinal studies, program development, and efficacy and effectiveness research must be conducted. Only programs that survive this process can move on to implementation and the study of factors that facilitate or impede its success. During dissemination, however, investigators who once enjoyed the shelter of the Ivory Towers must adjust to dynamic community forces that can become uncontrollable.

The science of implementation includes systemic study of the organizational structures in which changes occur (Price & Lorion, 1989). For example, the nationwide implementation of the Olweus program in Norway benefited from ideal organizational circumstances. Norway is a social democracy, small and rather homogenous in population, rich in resources, and strong in commitment to social change. Furthermore, the Norwegian Ministries provided powerful policy support as well as financial resources for the program. Most implementations, on the other hand, take place in underresourced, overextended communities without the benefit of strong leadership and political advocacy at the highest levels.

Even in Norway, the greatest program effects were seen in schools with the strongest commitment to implementing the program (Roland, 1993). What factors contributed to reception of the bully/victim program by the teachers in experimental schools in the original Olweus study (Olweus, 1991)? Kallestad and Olweus (2003) found certain teacher characteristics and school climate factors to be relevant to program effectiveness. Teachers reported that their perceptions, personal experiences, and affective involvement, as well as their reading of program materials, all contributed significantly to use of the classroom intervention measures. It was somewhat surprising to note that professional and personal characteristics (e.g., gender, age, education, and years of service) were not contributors to program use. Between-schools predictors of classroom intervention measures were openness in communication and school attention to bullying problems. Given that these data were collected during the first careful test of the program, it is remarkable that Olweus had the foresight to collect implementation data such as these.

In efficacy research, theoretically based mechanisms are specified for targeted outcomes, experimentally manipulated through the intervention, and evaluated in tests of mediational models. The critical test of the intervention and the theory, of course, requires that the mediators be successfully altered in controlled trials, with concomitant benefits to the outcomes. This sort of precision within intervention research can only grow out of an iterative process in which theory, measurement, analysis, and intervention are subjected to progressively rigorous trials. Efficacy and effectiveness studies have come to require that these standards be applied to models, methods, and analytic strategies to evaluate whether programs work and, if they do, why. Now it is time to apply the same meticulous procedures to implementation programs to ensure that model programs can be successfully disseminated. In this regard, a next step in the Olweus body of work would be to experimentally manipulate the predictors of classroom intervention measures in
a randomized trial. For example, given that teachers’ perceptions and affective involvement predicted use of classroom intervention measures, one study design could involve random assignment of teachers to receive an intervention that would increase their sensitivity to bullying and to the experience of victims. Given that the intervention increased perceived levels of bullying, staff importance, and affective involvement, would teachers increase their use of classroom intervention? In all candor, it must be said that if or when Olweus and his colleagues carry out such studies, they will define the cutting edge of implementation research. As of yet, there are no systematic examples of this sort.

The same rigorous standards for measurement required in efficacy and effectiveness trials must be applied in implementation research. Measures should draw from multiple perspectives, agents, and settings and should reflect microsocial processes as well as molar dimensions of the social environment and its contexts. Assessments should tap issues such as severity and frequency of outcomes and be well timed to capture critical implementation processes and outcomes. High-quality assessment will decrease bias, increase predictive validity as well as sensitivity to change, and benefit the generalizability of findings.

In the Kallestad and Olweus (2003) analysis, the sole basis of independent and dependent variables came from teacher reports, a reasonable starting point. Teachers, however, may be subject to bias, limited information, and narrow perspectives. For example, asking teachers about the manner in which youngsters are supervised during break times may provide little validity if aids do the supervision. Reports provided by aids, on the other hand, may be biased by factors such as their frustration with limited resources and may not include accurate accounts of the strategies they use to prevent and intervene in bully/victim episodes. Collecting direct observational data in settings where these episodes occur is likely to yield valuable information about bullies/victims and their interpersonal behaviors, as well as the actions that adults take or fail to take—data that can then be applied to strengthen the program. Key settings for observation might include classrooms, hallways, lunchrooms, and playgrounds.

The Olweus Bullying Prevention Program rests on a theoretical foundation that requires restructuring the social environment of entire schools to encourage youngsters’ prosocial behavior and to discourage bully/victim episodes, an organizational challenge for schools. To the extent that teachers actively applied the intervention tools, bully/victim incidents decreased within the school environment and did not increase outside of school. One area in which the program failed to meet its goals was for teachers to reach outside the school to make direct contact with the parents of bullies and victims, a problem reported by other investigators attempting to replicate the program (Stevens, Bourdeaudhuij, & Oost, 2001). Failure to extend interventions beyond an intervention setting is a problem shared by programs conducted from the school settings as well as from mental health settings.

Kallestad and Olweus (2003) reported that the core components teachers were most likely to use involved reading the program literature and providing the program information folder to parents. Other effective yet perhaps more demanding intervention components (e.g., role-play, class meetings) received much less use. Implementation studies from other prevention arenas indicate that even when program delivery manuals were highly specific and user friendly, service agencies were more likely to adopt model programs when staff were (a) provided intensive training in intervention delivery, (b) backed up with individualized assistance, and offered telephone consultations (Kelly et al., 2000). An experiment to test
increased use of the full range of intervention components could involve random assignment of some
schools and their staff to receive intensive training and backup support, as well as telephone consultation
to the school. Increased use of each component could then be evaluated for increased participation within
schools and improved results.

Although the last two decades have yielded noteworthy development in the art and science of prevention
research with respect to efficacy and effectiveness trials, implementation research is still in its infancy. The data in the Kallestad and Olweus (2003) article suggest why the first intervention trial of the bully/victim program succeeded when carried out by the program developer. To understand the variation in replication attempts by others, researchers must apply experimental design and rigorous methodology to study implementation. Olweus rightfully occupies the role of pioneer in this arena.

References


