The evaluation of school-based social and emotional learning interventions: Current issues and future directions

Ann Lendrum & Michael Wigelsworth

Recent large-scale studies in the US have produced results suggesting that school-based, universal social and emotional learning (SEL) interventions may impact positively on a range of outcomes, including social and emotional skills, mental health difficulties, school attitudes and academic performance. However, many interventions fail to achieve the same impressive results when implemented and evaluated on a smaller-scale. Although research findings have indicated several possible explanations for these inconsistent findings, including implementation failure, the potential impact of other factors such as differences between efficacy and effectiveness trials, the level of involvement of programme developers in evaluations and issues around cultural transferability remain unexplored. This article examines each of these issues in the evaluation of school-based SEL interventions, highlighting important considerations for researchers working in this field. Directions for future research are also discussed.

IN RECENT YEARS, schools have increasingly been expected to play a role in supporting children's emotional education and development (Greenberg, 2010; Weare, 2010). Specifically, there has been a particular focus on the benefits of social and emotional learning (SEL) programmes, which aim to improve learning, promote emotional well-being and prevent problem behaviours through the development of social and emotional competencies (Elias et al., 2001; Greenberg et al., 2003). Although SEL programmes vary widely in their emphasis on the relative importance of these competencies, it is generally acknowledged that there are five crucial, inter-related, cognitive, affective and behavioural competencies: self-awareness, self-management, social awareness, relationship skills and responsible decision-making (Collaborative for Academic, Social and Emotional Learning, 2002). Recent meta-analyses in the US and the Netherlands have suggested that high quality, well-implemented universal SEL interventions, in which all children participate in practices designed to facilitate their intra- and inter-personal competence, may impact on a range of pertinent outcomes, including social and emotional skills, mental health difficulties, school attitudes and academic performance (Wilson & Lipsey, 2007; Durlak et al., 2011; Sklad et al., 2012). This broad range of expected outcomes has led to the use of school-based SEL programmes across the world including in the US (e.g. Greenberg et al., 1995), Australia (e.g. Graetz et al., 2008), across Europe (e.g. Holsen, Smith & Frey, 2008), and in the UK (e.g. DfES, 2007).

Despite the success reported in the meta-analyses (Wilson & Lipsey, 2007; Durlak et al., 2011; Sklad et al., 2012), SEL programmes are not always able to produce the same impressive results when adopted and implemented by practitioners in schools (Social and Character Development Research Consortium, 2010). Research in prevention science suggests a number of possible reasons for this, including implementation failure (Durlak & DuPre, 2008), a reliance on the results of efficacy trials (Flay et al., 2005), developer involvement in

trials (Eisner, 2009) and a lack of cultural transferability of interventions (Castro, Barrera & Martinez, 2004). Although the importance of implementation fidelity is increasingly acknowledged as critical and was included as a factor in Durlak et al.'s meta-analysis (2011), there has as yet been little consideration of the other factors. This paper will now consider these issues and discuss the implications for research.

Stage of evaluation:
Efficacy vs. effectiveness

An intervention should be tested at several stages between its initial development and its broad dissemination into routine practice (Greenberg et al., 2005). Frameworks for complex heath interventions provide a 'road-map' for this journey, and draw important distinctions between various key stages in developing and implementing an initiative. Campbell et al. (2000) provide guidance on specific sequential phases for developing complex interventions: developing theory (pre-phase), modelling empirical relation-ships consistent with intended outcomes (phase I), exploratory trialling (phase II), randomised control trials under optimum conditions (phase III) and long-term imple-mentation in uncontrolled settings (phase IV). Ideally, an intervention should pass through all phases to be considered truly effective and evidence-based. Of particular relevance to SEL interventions are the differences between phases III and IV. Phase III is evidenced by efficacy or 'pilot' studies which are typically conducted by the programme developers under highly controlled and optimal conditions in order to maximise outcomes (Durlak, 1998; Dusenbury et al., 2003) and demonstrate the efficacy and internal validity of a programme. Phase IV may be represented by effectiveness trials, which should be conducted in multiple natu-ralistic settings, using just the staff and resources that would be normally available (Dane & Schneider, 1998) to test whether and how an intervention works in real-world contexts (Durlak, 1998; Greenberg et al.,

2005). Effectiveness trials, which may also be described as 'Type Two translational research', should also identify factors that may influence the successful adoption, imple-mentation and sustainability of interventions when they 'go to scale' (Greenberg, 2010).

The importance of both efficacy and effectiveness trials for determining the utility of an intervention is recognised across disci-plines, including health (Campbell et al., 2000) and social and emotional learning (Flay et al., 2005). Despite calls for preven-tion programmes to be tested in multiple contexts before they are described as 'evidence-based' (Kumpfer et al., 2012), the validation of an intervention at just efficacy stage is sometimes enough to achieve 'exem-plary' or 'gold-standard' designation (Greenberg et al., 2005) and inclusion on a register of recommended programmes. However, the adoption and implementation of such a 'model' programme by a school does not guarantee that the outcomes achieved in an efficacy trial will be replicated (Elliott & Mihalic, 2004). Research indicates that practitioners in 'real-world' settings are generally unable to duplicate the favourable conditions and access the technical expertise and resources that were available to researchers and programme developers at the efficacy stage (Hallfors & Godette, 2002; Greenberg et al., 2005) and thus fail to implement programmes to the same stan-dard. This reduced quality and level of implementation is related to a failure to achieve the intended outcomes (Durlak & Du Pre, 2008). An example of this is an effec-tiveness trial of the Promoting Alternative Thinking Strategies (PATHS) curriculum in the Netherlands (Goossens et al., 2012), which failed to replicate the high levels of implementation and the successful outcomes demonstrated in earlier efficacy trials conducted by the programme devel-opers (Greenberg et al., 1995).

The stage at which an intervention is evaluated and classed as 'successful' is thus critical. There is, as yet, however, little clarifi-cation in the SEL literature, including the

major reviews, regarding the stage of evaluation of programmes and whether those classified as 'successful' or 'exemplary', have achieved this status on the basis of efficacy (phase III) trials alone or have also undergone effectiveness (phase IV) trials. This has implications not only for the classification of SEL programmes as effective, but also for practitioners who need to make informed choices about which programmes may be successfully adopted and implemented in their own context.

In summary, a lack of adherence to a clear road map in establishing the stage of evaluation at which a SEL programme has been deemed 'effective', means that the potential success of long term implementation of some programmes is unclear. A comparison of the effects of programmes at both efficacy and effectiveness stages is required in order to fully understand how interventions should be implemented in order to achieve the intended outcomes in real world contexts. Evaluation studies should be more explicit about the type of support that was available in both efficacy and effectiveness trials so that practitioners may make more informed decisions about their capacity to create optimal conditions for successfully implementing interventions in their own settings.

The involvement of programme developers

It is logical, and perhaps inevitable, that efficacy trials (which are intended to explore the feasibility and internal validity of an intervention) are typically conducted by its developers. However, evidence suggests that this may lead to inflated outcomes. In fields such as criminology, psychiatry and substance abuse education which, like SEL, make use of prevention programmes, interventions appear to achieve considerably larger effect sizes when programme developers are involved in evaluation studies (Eisner, 2009). For example, in a review of psychiatric interventions, studies where a conflict of interest was disclosed (e.g. programme developers

were directly involved in the study) were nearly five times more likely to report positive results when compared to truly 'independent' trials (Perlis et al., 2005). An independent effectiveness study of Project ALERT, a substance abuse prevention programme, failed to find positive outcomes, despite successful efficacy and effectiveness studies conducted by the programme developer (St. Pierre et al., 2005). Eisner (2009) offers two explanations for the startling differences in outcomes of these two types of study. The *cynical view* proposes that the more favourable results in developer-led trials stem from systematic biases that influence decision-making during a study. The *high fidelity view* argues that implementation of a given intervention is of a higher quality in studies in which the programme developer is involved, leading to better results. In either case, developer involvement leads to an inflation of outcome effect sizes compared to that which might be be expected from a 'real world' implementation of a given programme. The *high fidelity view* also has implications for the scalability of interventions. If the intended outcomes of an intervention may only be achieved if the programme developer is available to enforce the highest levels of fidelity, then its broad dissemination and sustainability across multiple setting is unlikely to be feasible. This reinforces the need for effectiveness studies, designed to demonstrate the validity of a programme under natural conditions, to be conducted by independent research teams.

The recent reviews and meta-analyses of SEL programmes do not distinguish between evaluations conducted by external researchers and those led by, or with the involvement of, programme developers or their representatives. There is thus a need for future literature to explicitly state who is involved in a study to allow comparison of effect sizes obtained in different types of evaluations. Moreover, reports of evaluations studies should be required to be more explicit about the role of developers and other interested parties.

Cultural transferability

As discussed above, the reduced effectiveness of programmes outside efficacy trials may relate to differences between the context of validation and that of later implementation. Further to this, studies of programmes transferred across contexts within the US have indicated that they are unlikely to succeed if they do not fit with the cultural needs, values and expectations of the adopters (Castro, Barrera & Martinez, 2004). Given that research suggests that variations in cultural contexts *within* countries may impact on effectiveness *(ibid)*, this raises questions around the extent to which programmes can achieve the intended outcomes when transported to countries with different education systems, cultural beliefs and expectations of the roles of teachers and pupils. This has been identified as a potential problem for other prevention and promotion interventions such as anti-bullying programmes (Ttofi, Farrington & Baldry, 2008) and sex and drugs education (Wiggins et al., 2009) both of which report null results for US programmes transported into the UK. Similarly, lack of impact has been reported for some SEL initiatives in some contexts. Little et al. (2012) failed to find an effect for either PATHS (Greenberg & Kusché, 2002) or Triple-P (Bodenman et al., 2008), despite an established international literature base supporting these programmes. Findings are mixed, however; some positive effects for PATHS have been found in Switzerland (Malti, Ribeaud & Eisner, 2011) and the SEL programme 'Second Step' (Committee for Children, 2011) has been shown to have positive effects across several sites in the US (e.g. Frey et al., 2005; Cooke, et al., 2007) and in Europe (Holsen et al., 2008).

A major factor in the successful transportability of interventions is their adaptability (Castro et al., 2004). Adaptations vary however, and although surface level changes (e.g. modified vocabulary, photographs or names) may be beneficial and enhance cultural acceptability, deeper structural modifications, (e.g. different pedagogical approaches or modified programme delivery), may compromise the successful implementation of the critical components of an intervention. This may have serious negative consequences, to the extent that change is not triggered and the outcomes of the programme are not achieved. Indeed, there is arguably the potential for programmes to be adapted to cultural contexts to such an extent that they become, in effect, new programmes requiring re-validation, ideally through the use of an evidence framework, such as Campbell et al's. (2000) in order to test the underlying programme theory and internal validity.

Issues of cultural transferability have particular implications for SEL programmes. Many have been developed and tested in the US and their perceived success has resulted in their rapid global dissemination and adoption. For example, in England, the failure of national strategies such as the social and emotional aspects of learning (SEAL) programme has led to an increasing interest in the adoption of evidence-based approaches developed outside the UK, often at considerable expense (for example, the National Lottery Fund's recent £7m investment through the Realising Ambition project).

Despite the adoption of SEL programmes around the world, however, the vast majority of the SEL literature originates from the US and it is noteworthy that around 90 per cent of the studies included in Wilson and Lipsey's (2007) and Durlak et al.'s (2011) reviews originated there. Although some of the more established programmes with strong evidence bases in the US, such as Promoting Alternative Thinking Strategies (PATHS) (Greenberg & Kusché, 2002), Second Step (Committee for Children, 2011) and Incredible Years (Webster-Stratton, 2011) have been adopted and implemented across the world (e.g. Holsen et al., 2008; Malti, Ribeaud & Eisner, 2011; Henningham, 2013), the varying levels of success and the extent to which this may be

due to adaptations or poor fit with cultural contexts is not yet clear. This suggests that there is a need for future research to carefully examine the extent to which programmes implemented outside their country of origin are able to achieve their expected outcomes with or without adaptations and explicitly report any modifications that are made

Conclusion

In summary, although recent reviews of SEL programmes have indicated that they may be effective in the achievement of a broad range of outcomes (Wilson & Lipsey, 2007; Durlak et al., 2011; Sklad et al., 2012), questions remain about the basis of these claims. The influence of factors such as stage and type of evaluation, level of developer involvement, and cultural transferability is, as yet, unclear. An examination of the possible impact of these factors is vital and may offer significant insight into current tensions around inconsistent findings. This is especially true given the potential for interaction between these factors. For example, the involvement of a programme developer in an effectiveness evaluation may mediate issues associated with transferability and scalability of an intervention, or over-

adaptation to fit cultural contexts may result in a 'new' programme that requires efficacy trials to assess feasibility and internal validity. This highlights the importance of examining these issues, not only in future research, but also when reviewing existing studies. To this end, the authors are currently conducting a systematic review and meta-analysis of the SEL literature base with a specific focus on the issues highlighted in this paper. The findings of this project, alongside similar considerations embedded in future studies, should present an advancement in the knowledge base and serve to deepen understanding of the potential of SEL to effect positive change in the lives of children and young people.

The Authors
Ann Lendrum & Michael Wigelsworth
University of Manchester, UK.

Correspondence
Dr Ann Lendrum
School of Education,
University of Manchester,
Oxford Road,
Manchester, M13 9PL, UK.
Email: ann.lendrum@manchester.ac.uk

References

Big Lottery Fund (2013). *Funding: Big lottery funding*. Retrieved 17 July 2012, from: http://www.biglotteryfund.org.uk/ prog_realising_ambition

Bodenman, G., Cina, A., Ledenmann, T. & Sanders, M.R. (2008). The efficacy of positive parenting programme (Triple P) in improving parenting and child behaviour: A comparison with two other treatment conditions. *Behaviour Research and Therapy, 46*, 411–42.

Campbell, M., Fitzpatrick, R., Haines, A., Kinmouth, A., Sandersock, P., Spiegelhalter, D. & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to improve health. *British Medical Journal, 321*, 694–696.

Castro F.G., Barrera, M. & Martinez, C.R. (2004). The cultural adaptation of prevention interventions: Resolving tensions between fidelity and fit. *Prevention Science, 5*, 41–45.

Collaborative for Academic, Social and Emotional Learning [CASEL] (2002). *Guidelines for social and emotional learning: High quality programmes for school and life success.* Chicago, IL: CASEL.

Committee for Children (2011). *Second Step: A violence prevention curriculum.* Seattle, WA: Committee for Children.

Cooke, M., Ford, J., Levine, J., Bourke, C., Newell, L. & Lapidus, G. (2007). The effects of city-wide implementation of 'Second Step' on elementary school students' prosocial and aggressive behaviours. *The Journal of Primary Prevention, 28*, 93–115.

Dane, A.V. & Schneider, B.H. (1998). Programme integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18*, 23–45.

Department for Education and Skills (2007). *Social and Emotional Aspects of Learning for Secondary Schools (SEAL) guidance booklet.* London: DfES.

Durlak, J.A. (1998). Why programme implementation is important. *Journal of Prevention and Intervention in the Community, 17*, 5–18.

Durlak, J. & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on programme outcomes and the factors affecting implementation. *Journal of Community Psychology, 41*, 327–350.

Durlak, J., Weissberg, R., Dymnicki, A., Taylor, R. & Schellinger, K. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*, 405–432.

Dusenbury, L., Brannigan, R., Falco, M. & Hansen, W.B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health and Education Research, 18*, 237–256.

Dusenbury, L., Brannigan, R., Hansen, W.B, Walsh, J. & Falco, M. (2005). Quality of implementation: Developing measures crucial to understanding the diffusion of preventive interventions. *Health Education Research, 20*, 308–313.

Eisner, M. (2009). No effects in independent prevention trials: Can we reject the cynical view? *Journal of Experimental Criminology, 5*, 163–183.

Elias, M.J., Lantieri, L., Patti, J., Shriver, T.P., Walberg, H. Weissberg, R.P. & Zins, J.E. (2001). No new wars needed. *Reclaiming Children and Youth*, Spring, 29–32.

Elliott, D.S. & Mihalic, S. (2004). Issues in disseminating and replicating effective prevention programmes. *Prevention Science, 5*, 47–53.

Flay, B.R., Biglan, A., Boruch, R.F., Castro, F.G., Gottfredson, D. & Kellam, S. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6*, 151–175.

Frey, K., Nolen, S., Edstrom, L. & Hirschstein, M. (2005). Effects of a school-based social-emotional competence programme: Linking children's goals, attributions, and behaviour. *Journal of Applied Developmental Psychology, 26*, 171–200.

Goossens, F., Gooren, E., de Castro, B., van Overveld, K., Buijs, G., Monshouwer, K., Onrust, S. & Paulussen, T. (2012). Implementation of PATHS through Dutch municipal health services: A quasi-experiment. *International Journal of Conflict and Violence, 6*, 234–248.

Graetz, B., Littlefield, L., Trinder, M., Dobia, B., Souter, M., Champion, C. & Boucher, S. (2008). KidsMatter: A population health model to support student mental health and well-being in primary schools. *The International Journal of Mental Health Promotion, 10*, 13–20.

Greenberg, M.T. & Kusché, C. (2002). *Promoting Alternative Thinking Strategies (PATHS), blueprints for violence prevention*. Boulder, CO: Center for the Study and Prevention of Violence, University of Colorado.

Greenberg, M.T., Kusche, C.A., Cook, E.T. & Quamma, J.P. (1995). Promoting emotional competence in school-aged children: The effects of the PATHS curriculum. *Development and Psychopathology, 7*, 117–136.

Greenberg, M.T., Weissberg, R., O'Brien, M., Zins, J., Fredericks, L., Resnik, H. & Elias, M. (2003). Enhancing school-based prevention and youth development through co-ordinated social, emotional, and academic learning. *American Psychologist, 58*, 466–474.

Greenberg, M.T., Domitrovich, C.E., Graczyk, P.A. & Zins, J.E. (2005). *The study of implementation in school-based preventive interventions: Theory, practice and research*. Washington, DC: Department of Health and Human Services.

Greenberg, M.T. (2010). School-based prevention: Current status and future challenges. *Effective Education, 2*, 27–52.

Hallfors, D. & Godette, D. (2002). Will the 'Principles of Effectiveness' improve prevention practice? Early findings from a diffusion study. *Health Education Research, 17*, 461–470.

Henningham, H. (2013). *Transporting an evidence-based programme to Jamaican preschools: Cluster randomised controlled trial*. Paper presented at the 21st Society for Prevention Research Conference, San Francisco.

Holsen, I., Smith, B.H. & Frey, K. (2008). Outcomes of the social competence programme Second Step in Norwegian elementary schools. *School Psychology Internationa, 29*, 71–88.

Kumpfer, K.L., Magalhães, C. & Xie, J. (2012). Cultural adaptations of evidence-based family interventions to strengthen families and improve children's developmental outcomes. *European Journal of Developmental Psychology, 9*, 104–116.

Little, M., Berry, V., Morpeth, L., Blower, S., Axford, N., Taylor, R., Bywater, T., Lehtonen, M. & Tobin, K. (2012). The impact of three evidence-based programmes delivered in public systems in Birmingham, UK. *International Journal of Conflict and Violence, 6*, 260–272.

Malti, T., Ribeaud, D. & Eisner, M.P. (2011). The effectiveness of two universal preventive interventions in reducing children's externalising behaviour: A cluster randomised controlled trial. *Journal of Clinical Child and Adolescent Psychiatry, 40*, 677–692.

Perlis, R.H., Perlis, C.S., Wu, Y., Hwang, C., Joseph, M. & Nierenberg, A.A. (2005). Industry sponsorship and financial conflict of interest in the reporting of clinical trials in psychiatry. *American Journal of Psychiatry, 162*, 1957–1960.

Social and Character Development Research Consortium (2010). *Efficacy of schoolwide programmes to promote social and character development and reduce problem behavior in elementary school children (NCER 2011–2001).* Washington, DC: US Government Printing Office.

Sklad, M., Diekstra, R., de Ritter, M. & Ben, J. (2012). Effectiveness of school-based universal social, emotional, and behavioural programmes: Do they enhance students' development in the area of skill, behaviour, and adjustment? *Psychology in the Schools, 49*, 892–909.

St.Pierre, T.L., Osgood, D.W., Mincemoyer, C.C., Kaltreider, D.L. & Kauh, T.J. (2005). Results of an independent evaluation of Project ALERT delivered in schools by co-operative extension. *Prevention Science, 6*, 305–317.

Ttofi, M.M., Farrington, D. & Baldry, A.C. (2008). *Effectiveness of programmes to reduce school bullying: A systematic review of school bullying programmes.* Stockholm: Swedish National Council for Crime Prevention.

Weare, K. (2010). Mental health and social and emotional learning: Evidence, principles, tensions, balances. *Advances in School Mental Health Promotion, 3*, 5–17.

Webster-Stratton, C. (2011). *The Incredible Years: Parents, teachers, and children's training series.* Seattle, WA: Incredible Years Inc.

Wiggins M., Bonnell C., Sawtell M., Austerberry H., Burchett H., Allen E. & Strange V. (2009). Health outcomes of youth development programmes in England: Prospective matched comparison study. *British Medical Journal, 339*, 2543.

Wilson, S.J. & Lipsey, M.W. (2007). School-based interventions for aggressive and disruptive behaviour. Update of a meta-analysis. *American Journal of Preventative Medicine, 3*, 130–143.