Abstract

Effective intervention for children who present challenging behaviours remains a practical and theoretical priority. Previous research has investigated interventions which draw on a range of different psychological models, including behaviourist, psychodynamic, cognitive-behavioural and systemic perspectives. Exploring the impact of school-delivered targeted interventions extends previous research into effectiveness of intervention in real-world settings and offers insight into feasibility issues and the impact of early intervention. The current review aimed to update and extend the work of Evans, Harden and Thomas (2004) in exploring how to support children with behavioural difficulties in mainstream primary schools. A search of the literature identified seventeen journal articles which met inclusion criteria. These studies explored a variety of targeted individual and small group interventions to address externalising behaviour problems. The target behaviours the research addressed included both promoting positive behaviours, such as time on-task, and reducing negative behaviours, such as disruptive or distracting episodes. The studies were evaluated and critiqued and the evidence subsequently weighted according to Gough's (2007) 'Weight of Evidence' Framework. Convincing support for the effect of intervention on externalising behaviour was found, particularly in terms of outcomes for increased positive behaviours. Therefore, it seems that intervening to improve positive outcomes for children with externalising behaviour difficulties may be both feasible and an efficient use of resources. Recommendations for implementation of interventions for externalising behaviour difficulties and areas for future research are outlined.

Introduction

Social, emotional and mental health

The way in which behavioural difficulties are conceptualised and understood continues to evolve, reflecting shifts in psychological paradigms and special educational needs legislation (Frederickson & Cline, 2009). The Special Educational Needs and Disability Code of Practice (CoP; DfE, 2015) suggests that children may display 'challenging, disruptive or disturbing' behaviour due to a wide range of needs, including mental health difficulties. Effective intervention involves exploration of the causes for problematic behaviour and treatment which enables the child to achieve in school both academically and socially.

Establishing the effectiveness of school-based interventions for children with behaviour difficulties is a significant ethical and practical goal. The Equality Act (2010) and the CoP (2015) emphasise the imperative for inclusion, which includes ensuring that children with special educational needs and disabilities (SEND) are effectively supported in mainstream schools.

Basis in psychological theory

The distinction between students with challenging behaviour and those for whom persistent disruptive behaviour is indicative of an underlying learning or mental health difficulty is problematic for both psychologists and educators. Jones (2003) argues for further integration of psychological theory and educational practice so that the needs of pupils with emotional and behavioural difficulties can be met. Jones (2003) suggests that embedding psychological theory, explanatory frameworks and research-based evidence in a school context is critical in order to

reduce ‘pragmatic eclecticism’ and ensure that students’ psychological needs are not overlooked when trying to adapt their behaviour to fit school expectations.

Early intervention and preventative work continues to be a priority when considering how best to improve outcomes for children with additional needs (CoP, 2015). Intervening early in the onset of difficulties is considered an effective strategy given that the cost to individuals and society of managing and treating established disorders is high and such disorders are often hard to treat (Baxter & Frederickson, 2005). Targeted interventions can, therefore, play an important role in ensuring that children’s additional needs are met at an early stage.

A range of strategies reflecting different strands of psychological thinking have been employed as interventions for children with behaviour difficulties. Behavioural, cognitive-behavioural, psychotherapeutic and systemic approaches have all been adopted when considering how to intervene to address problematic behaviour (Evans, Harden & Thomas, 2004). Functional behaviour assessment (FBA) has emerged as one of the key ways in which information is gathered with regard to the particular nature and severity of a child’s behaviour and can make a significant contribution to decision making with regard to the effectiveness and efficiency of intervention (Horner, 1994).

Intervention implementation is an important consideration for researchers in a school-based setting. School staff are rarely utilised in intervention delivery, despite the fact that they have been found to be at least as effective as other professionals in this respect (O’Mara, Marsh, Craven & Debus, 2006). Carter, Stephenson and Clayton (2008) explored the support services

accessed by schools when dealing with challenging behaviour and found that within-school support, such as that provided by other teachers or management, was accessed comparatively more frequently and perceived to be more effective than other sources, such as external professional support services. Monitoring implementation quality must, however, be given due consideration given research suggesting that programme implementation across change agents can be of varying effectiveness and is sometimes seriously compromised, which in turn influences outcomes (Durlak, 1998).

Some important evidence relating to the impact of school-based interventions on externalising behaviours has come from meta-analyses exploring the broader remit of social and emotional learning. Universal programmes were found to be effective in helping students develop improved social and emotional skills, attitudes, behaviour and academic performance (Durlak, Dymnicki, Taylor, Weissberg & Schellinger, 2011). In addition, there was evidence that school teaching staff successfully conducted social and emotional learning programmes. Two systematic reviews of research into the effectiveness of strategies to support students with emotional and behavioural difficulties in mainstream primary schools found that a number of strategies, based on a range of theoretical frameworks, showed some positive impacts on student behaviour (Evans et al., 2004; Harden, Thomas, Evans, Scanlon & Sinclair, 2003). There has been limited consideration of whether strategies to reduce problem behaviour or support appropriate behaviour are more effective (Partin, Robertson, Maggin, Oliver & Wehby, 2009).

For the purposes of this review, Shucksmith et al.'s (2007) use of the term 'externalising behaviours' will be employed. However, with an early identification and intervention purpose in mind, this review will focus on studies which intervene prior to the diagnosis of disability or

disorder. This review seeks to extend the findings of Evans et al. (2004) and therefore will focus on interventions carried out in mainstream primary schools. Research into the types of outcomes for which intervention is most effective is needed and, with this in mind, the present review explored whether interventions are more effective in increasing students' positive behaviours or reducing problematic behaviours.

Critical Review of Evidence Base

A comprehensive literature search was carried out in July 2014 and updated in July 2015. A search was undertaken on four databases (PsychINFO, ERIC, British Education Index and Medline) using search terms for subject headings related to intervention, population and outcomes. The intervention terms related to programme effectiveness or evaluation, population focused on elementary or primary education and outcomes explored a range of terms related to behaviour or conduct, as well as social, communication or interpersonal skills. An example of the search terms used in the PsychINFO database is outlined in Table 1.1. Search terms were combined so that each article contained at least one term from each column as a subject heading.

The searches were limited so that only articles in peer-reviewed journals (as a control for quality of research) and written in English (as resources for translation were not available) were included in the search results. The search excluded articles published prior to 1999 because a systematic review which focused on a similar research question included articles published between 1975 and 1999 in the search strategy (Evans et al., 2004). A range of inclusion and exclusion criteria applied to articles in terms of intervention, participants, setting and outcome measures are outlined in Appendix A. See Figure 1.1 for a flow diagram of the study selection process.

The seventeen studies identified were critically analysed using Gough's (2007) 'Weight of Evidence' Framework, summarised in Table 1.2. The weighting criteria are detailed in Appendix B. The weighting of each study is presented in Table 1.3.

Participants

The number of participants in each study ranged from one to two hundred and fifty three and ages ranged from four to eleven years old. Males formed the majority of participants across the selected studies as a whole. In terms of setting, all schools were general education primary or elementary schools and were located in predominantly urban districts. The majority of the studies included in this review were undertaken in the USA.

Participants in most studies received all academic instruction in a general education classroom. Many of the studies explicitly stated that participants had not received a diagnosis of any physical, learning, neurological or psychiatric disability that might affect their presenting behaviour. Where this information was not given, it was inferred from the study's focus on typically developing children. The rationale for including participants in the study was usually based on teacher report of persistent or multiple behaviours that were found to be problematic in a classroom context.

Design & Measures

Four studies employed a group experimental or quasi-experimental design and were evaluated in terms of methodological quality using an adapted version of Gersten et al.'s (2005) quality indicators for such research. Thirteen papers used a single-case design (SCD), which is a rigorous, scientific methodology that can explore treatment effects with a small sample size (Horner et al., 2005). SCD involves manipulation of an independent variable together with repeated measurement of a dependent variable before, during and after introduction of the independent variable. The systematic structure of such a design provides a strong basis for establishing causal inference (Kratochwill et al., 2010). Each study's weighting for methodological quality was assessed using an adapted version of the SCD coding protocol devised by Kratochwill et al. (2010).

Two of the pre-post quasi-experimental design studies received a low weighting for methodological quality due to a lack of a control group, which is an important indicator of experimental quality (McIntosh et al., 2009; Hawken, O'Neill & MacLeod, 2011). Humphrey et al. (2010) and Wyman et al. (2010) received a higher rating for methodological quality because they employed appropriate procedures to increase the likelihood that relevant characteristics of participants in the sample were comparable across experimental and control conditions.

Experimental control in SCD is established through three demonstrations of experimental effect, which can be achieved in a range of design structures (Horner et al., 2005). Many studies employed a multiple baseline design, either across settings (Lane, Smither, Huseman, Guffey & Fox, 2007) or across participants (Todd et al., 2008), which enables intervention effectiveness to

be demonstrated by introducing it after numerous baseline periods. However, a number of studies did not document three demonstrations of experimental effect and therefore received a low rating for methodological quality (Lane et al., 2007; Wilkinson, 2003; Wood, Umbreit, Liaupsin & Gresham, 2007). Most studies received a high rating for the quality of baseline data, which in SCD functions as a form of within-participant control.

Many of the single-case studies did not triangulate observation data with another method of measuring the same dependent variables, such as a checklist or rating scale, which would have further increased the reliability of results obtained. Therefore, only five single-case studies received a high weight of evidence for reliability and validity of measurement (Wilkinson, 2003; Mong, Johnson & Mong, 2011; De Martini-Scully, Bray & Kehle, 2000; Restori et al., 2007; Todd et al., 2008). The reliability and validity of measures used in the group experimental and quasi-experimental studies were similarly variable. For example, Hawken et al. (2011) received a low weighting due to the limited validity and reliability information available.

Intervention

There were a range of interventions explored by the studies, though all were targeted and delivered by school staff. A number of studies implemented antecedent- or consequent-based approaches, including those focused on adjustments, self-monitoring, task modification, reinforcement or extinction (Janney et al., 2013; De Martini-Scully et al., 2000; Umbreit et al., 2004; Kamps et al., 2006; Restori et al., 2007; Lane et al., 2007; Wood et al., 2007; Shumate & Wills, 2010). Wilkinson (2003) explored the use of behavioural contracts.

Many of the studies undertaken in the USA followed a model of three-tiered support and focused on tier two interventions, intended to benefit the 5-15% of students at risk for problem behaviour whose needs are not met by universal programmes (Sugai et al., 2000). The Check-In/Check-Out (CICO) programme was one tier two intervention that several studies investigated (Mong et al., 2011; Todd et al., 2008; McIntosh et al., 2009; Campbell & Anderson, 2008; Miller et al., 2015). Hawken et al. (2011) investigated the effectiveness of the Behaviour Education Programme (BEP), which is another name used for CICO. In these studies, the intervention consisted of the student checking in and out of school with a member of school staff at the start and end of each day and collecting feedback throughout the day on a daily report card. Meeting behaviour goals, collecting points towards rewards and receiving positive or constructive feedback were all components of the intervention.

Many of the studies employed a functional behavioural assessment (FBA) for each participant in order to investigate the context and function of the problematic behaviour prior to intervention implementation. This process enabled the researchers to design an intervention package that was tailored to meet the individual needs of each participant and therefore these studies received a higher rating for ecological validity.

Two studies explored interventions which were designed to promote social and emotional competence in children. Wyman et al. (2010) investigated the impact of the Rochester Resilience Project on the classroom behaviours and social-emotional functioning of children with elevated behavioural and social classroom problems. The intervention consisted of lessons and practise opportunities with school-based mentors in which participants were taught skills including monitoring of emotions, self-control and regaining equilibrium. Humphrey et al.

(2010) conducted an effectiveness trial of New Beginnings, a short social-emotional intervention for children delivered in a school setting and facilitated by a member of school staff. The programme involved the exploration of a range of feelings, as well as learning about and practising skills for 'calming down' and 'problem solving'.

In order to receive a high weight of evidence rating for methodological relevance, studies needed to document the way in which the intervention was delivered. This could be achieved by including information on how school staff were trained in the intervention or supported in its delivery. Alternatively, the study could have included information on treatment integrity or implementation fidelity as this is considered a key quality indicator in intervention research (Gersten et al., 2005; Horner et al., 2005). As well as detailing both treatment integrity and staff training, McIntosh et al. (2009) established a context for intervention by reporting participants' scores on a standardised assessment of behaviour, the Behaviour Assessment Scale for Children 2 (BASC-2; Reynolds & Kamphaus, 2004), which allows for comparisons to be made between participants and other children.

A further factor that was taken into consideration when evaluating methodological relevance was whether follow-up data had been collected rather than only immediate intervention effects. This was deemed to be important because it is an additional indicator of cost and intervention effectiveness (Gersten et al., 2005).

Outcomes

All of the single-case studies collected observational data on incidence rates of target behaviours. The group experimental and quasi-experimental studies employed measures of social and emotional competence or collected information on office discipline referrals (ODRs) to explore the effect of intervention on target behaviours.

The types of behaviour targeted by each study are shown in Table 1.4. Outcomes can be considered in terms of whether the introduction of the intervention led to a reduction in a problematic behaviour (e.g. time off-task) or an increase in a positive behaviour (e.g. time on-task), or a combination of the two. Effect sizes for each outcome, calculated for each case where this information was not present in the original study, are also presented in the table. The group experimental and quasi-experimental studies tended to report Cohen's *d* effect sizes and these were interpreted according to Cohen's (1992) recommendations. The exception was Hawken et al. (2011), but as the raw data was reported in the study a Cohen's *d* effect size could be calculated for the purposes of the current review. Non-overlap of all pairs (NAP; Parker & Vannest, 2009) was the effect size calculated in the Miller et al. (2015) study. Parker and Vannest (2009) suggest how to weigh the relative strength of NAP scores and these interpretations are included in the table. Percentage of non-overlapping data (PND) was used as the evaluative effect size measure for the single-case studies included in this review because of its validity and demonstrated applicability in a wide range of contexts (Scruggs, Mastropieri & Casto, 1987; Scruggs & Mastropieri, 1994). Scruggs and Mastropieri (1998) suggest descriptive labels that interpret the relative strength of a PND effect size, and these interpretations are also included in the table.

Increase in positive behaviours

Five studies explored whether a targeted individual or small group intervention could help participants with externalising behaviour difficulties develop more positive behaviours. The positive behaviours targeted by these studies included on-task, behaviour control, peer social skills and social and emotional competence. Overall, the impact of intervention on a positive behaviour outcome appeared to be very effective, with PND across studies ranging from 75-100% (*M*=94%). Cohen's *d* effect sizes ranged from 0.31-0.99.

The only study in the current review that received an overall high weight of evidence rating researched the extent to which a function-based intervention increased levels of on-task behaviour and reported results representing 90-100% PND across participants (Janney et al., 2013). On the whole, studies exploring the effect of intervention on positive outcomes were deemed to be fairly sound in terms of weight of evidence ratings. Although the research undertaken by Wyman et al. (2010) received an overall low weight of evidence rating, this was largely due to judgements regarding the relevance of both the methodology and the study focus, rather than the quality of the methodology.

Three of the studies which explored the effect of intervention on positive behaviours received a high weight of evidence rating for study focus relevance (Janney et al., 2013; Wood et al., 2007; Umbreit et al., 2004). Issues of social and ecological validity were considered when evaluating studies because the meaning and impact of the intervention for participants, particularly representing the views of children and including them in decision-making, is both a legislative and ethical imperative (CoP, 2015; BPS Code of Ethics and Conduct, 2009). In

order to receive a high weighting, studies had to collect information on intervention acceptability and the social importance of the effects of intervention from the point of view of both a staff member and the participant, which might be conceptualised as a consumer satisfaction survey (Kratochwill et al., 2010).

Decrease in problematic behaviours

Seven studies explored whether a targeted intervention led to a reduction in the problematic behaviours of participants with externalising behaviour difficulties. The problematic behaviours targeted by these studies included off-task, disruptive and problem behaviour, which were defined in greater detail in each study (for example, non-compliance or talking out of turn). Overall, the impact of intervention on a problematic behaviour outcome appeared to be questionable, with PND across studies ranging from 0-100% (*M*=68%). Cohen's *d* effect sizes ranged from 0.3-1.04. On the whole, studies which explored a problematic behaviour outcome received an overall weight of evidence rating of medium. Where there were critical considerations with regard to the study quality or methodology, this was in relation to the stability or variability of participants' data (Shumate & Wills, 2010; Todd et al., 2008; Hawken et al., 2011). In terms of study focus relevance, the majority of studies demonstrated acceptable relevance, including with regard to treatment acceptability (Hawken et al., 2011) and social and ecological validity (Wilkinson, 2003; De Martini-Scully et al., 2000).

Positive and problematic behaviours

Five studies operationalised both positive and problematic behaviours as outcome variables. Kamps et al. (2006), Lane et al. (2007) and Restori et al. (2007) reported larger effect sizes for the effect of intervention on promoting a positive behaviour (83-100% PND, *M*=95%) rather than decreasing a problematic one (58-100% PND, *M*=83%). McIntosh et al. (2009) found that CICO was more effective for problem (*d*=1.04) rather than prosocial (*d*=0.99) behaviour, according to teacher ratings on the BASC-2 for participants where the function of the problematic behaviour is to obtain attention. However, both outcomes represent large effect sizes. Miller et al. (2015) reported similar results, with CICO slightly more effectively in reducing problem behaviours than increasing academic engagement.

Restori et al. (2007), who received a high rating for methodological quality, provide an interesting example of the overall pattern emerging in the data, in that intervention was overall found to be more effective in increasing academic engaged time than decreasing disruptive behaviour. In this study, antecedent- and consequent- based strategies were compared, with an antecedent-based strategy found to be highly effective for both academic engagement (*M*=100% PND) and disruptive behaviour (*M*=100% PND). A consequent-based strategy, while slightly less effective, still led to significant improvements in participants' engagement (*M*=93% PND) and behaviour (*M*=88% PND). When considered in light of the current review question, these results contribute to the finding that targeted individual and small group interventions are more effective when a positive rather than a problematic behaviour is defined as the outcome variable.

On the whole, studies which explored both behaviour outcomes received an overall weight of evidence rating of medium, although some were more variable in terms of methodological

quality due to factors such as sample size, rigour of a multiple baseline design or quality of the data (McIntosh et al., 2009; Lane et al., 2007; Kamps et al., 2006). Two of the studies that explored both positive and problematic behaviour outcomes received a medium or high rating for study focus relevance due to their exploration of treatment acceptability and effectiveness (Lane et al., 2007; Miller et al., 2015).

Conclusion and Recommendations

This review examined seventeen studies which explored the effects of targeted individual or small group school-based interventions on outcomes for primary-aged children with externalising behaviour problems. Overall, intervention was found to be more effective for positive behaviour outcomes rather than problematic behaviours. Positive outcomes included on-task and academically engaged behaviour, as well as peer social skills and behaviour control. A range of interventions were utilised by researchers, including those designed following FBA, CICO and individual and small group social and emotional programmes.

Studies in the current review were grouped according to whether they explored a positive or problematic behaviour outcome, or both. In each of the three groups, there was one study that was given a low overall weight of evidence rating. While this implies that all conclusions should be tentatively drawn, it is also means that there was a fairly even spread of study quality across the three groups. Two studies rated highly for methodological or overall study quality demonstrated very effective treatment impact for on-task behaviours and academic engagement (Janney et al., 2013; Restori et al., 2007).

The finding that intervention is more successful in promoting positive behaviours chimes with the current shift towards preventative and positive psychology, which prioritises supporting children to develop skills rather than focusing on extinguishing perceived deficiencies (Baxter & Frederickson, 2005). The emphasis on individual and collective strengths rather than deficits and on competency building rather than pathology is supported by research demonstrating the protective impact of resilience factors and positive experiences (Noble & McGrath, 2008; Seligman, Ernst, Gillham, Reivich & Linkins, 2009). There is also a fit with research indicating the effectiveness of positive reinforcement (Maag, 2001).

The range of interventions included in the current review varied in effectiveness. Studies which demonstrated the effectiveness of an intervention based on the principles of functional assessment were, on the whole, of sound design. This, therefore, represents a promising behaviour change intervention to research further. The evidence base for the CICO intervention was of more variable quality and therefore interpretations of its positive impact should be treated with more caution.

One imperative for the current review was that pragmatic issues should not be overlooked in school-based research. Despite the fact that all studies featured interventions delivered by typical school personnel, and many demonstrated social and ecological validity, only four included follow-up data. Demonstrating maintenance of treatment effects is an important element of intervention research (Yeaton & Sechrest, 1981) and is particularly relevant in the case of short-term interventions where, without some idea of longer term effects, it is difficult to form judgements of treatment effectiveness. These are obviously important considerations for schools, where there is a need to justify investment of time or other resources. Of the four

studies that did collect follow-up data, maintenance of treatment effects appeared to be satisfactory.

A salient finding from the current review is the overwhelming number of studies which operated from a behaviourist paradigm for intervention. Many of the studies adopted the principles of functional assessment, and its connections to applied behaviour analysis, in their conceptualisation of the nature of problematic behaviour (Horner, 1994). This may in part be due to the criteria against which studies were judged for inclusion in the review, in that more intensive interventions for children with identified special educational needs and those delivered by external professionals were not included. Therefore, the range of interventions featured is perhaps limited and not representative of the multi-component interventions that have been found to be effective for children with diagnosed conduct problems (Webster-Stratton, Reid & Hammond, 2004), or those stemming from a cognitive-behavioural or psychotherapeutic approach (Evans, Harden & Thomas, 2004).

Regardless, the findings of the current review suggest that there is a limited range of interventions carried out by school staff and that these interventions are usually behaviourist in design. Only two studies explored an intervention focused on developing social and emotional competence (Humphrey et al., 2010; Wyman et al., 2010). While there has been a legislative move towards considering the mental health difficulties that may underlie challenging behaviour (CoP, 2015), it seems there is still scope for research and practice to consider intervening for positive behaviour outcomes from a much broader perspective.

Methodological limitations found in some studies included use of a wait-list control group, small sample sizes, lack of information about how the impact of different intervention agents was accounted for and, for the single-case designs, baseline data of variable quality. This review considered a minimum of three baseline data points satisfactory, following Kratochwill et al. (2010). The limitations of the quality of evidence of some studies included in this review mean that caution should be exercised when interpreting conclusions.

The present review aimed to extend the work of Evans et al. (2004) in exploring how to support children with behavioural difficulties in mainstream primary schools. Several of the conclusions drawn by Evans et al. are similar to findings of this review, in particular the limited evidence base, restricted range of interventions investigated and lack of high quality research. However, the current review contributes a number of interesting insights to the evidence base on interventions for children with externalising behaviour difficulties. The first is that interventions tend to be more effective when students are supported to develop positive behaviours, rather than reduce problematic behaviours. Additionally, in the majority of studies, interventions implemented by school staff appeared to result in significant and meaningful change for participants. The role of support staff in schools is a topical issue given recent recommendations with regard to their effective deployment (Russell et al., 2012) and therefore the findings of this review are timely.

Recommendations for further research

Further research could explore a wider range of interventions in terms of whether it is feasible for school staff to be trained in their implementation, including those based on

psychotherapeutic, cognitive-behavioural and systemic models. In addition, although some studies included follow-up data, there is a clear need for more extensive and longer term investigation of intervention effects over time. Generalisation data could also be scrutinised, given that the generalisation of skills from a specific intervention context to other areas of functioning is considered another key indicator of high quality school intervention implementation (Domitrovich et al., 2008). Finally, the range of possible contributing factors to externalising behaviour difficulties should not be ignored. Mental health difficulties or undetected speech, language and communication needs may underlie challenging behaviours and should always be explored.

References

Baxter, J., & Frederickson, N. (2005). Every child matters: Can educational psychology contribute to radical reform?. *Educational Psychology in Practice, 21(*2), 87-102.

British Psychological Society. (2009). *Code of Ethics and Conduct*. Leicester: BPS.

Campbell, A., & Anderson, C. M. (2008). Enhancing Effects of Check-In/Check-Out with Function-Based Support. *Behavioral Disorders, 33*(4), 233-245.

Carter, M., Stephenson, J., & Clayton, M. (2008). Students With Severe Challenging Behaviour in Regular Classrooms: Support and Impacts. *Australian Journal of Guidance and Counselling, 18*(02), 141-159.

Cohen, J. (1992). A power primer. *Psychological bulletin, 112*(1), 155.

De Martini-Scully, D., Bray, M. A., & Kehle, T. J. (2000). A packaged intervention to reduce disruptive behaviors in general education students. *Psychology in the Schools, 37*(2), 149-156.

DfE (2015). *Special Educational Needs and Disability Code of Practice: 0 to 25 years*. London: DfE.

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., ... & Ialongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion, 1*(3), 6-28.

Durlak, J. A. (1998). Why program implementation is important. *Journal of Prevention & Intervention in the Community, 17*(2), 5-18.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child development, 82*(1), 405-432.

Evans, J., Harden, A., & Thomas, J. (2004). What are effective strategies to support pupils with emotional and behavioural difficulties (EBD) in mainstream primary schools? Findings from a systematic review of research. *Journal of Research in Special Educational Needs, 4*(1), 2-16.

Frederickson, N. & Cline, T. (2009). *Special Educational Needs, Inclusion and Diversity*. Berkshire: Open University Press.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional children, 71*(2), 149-164.

Gough, D. (2007). Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education, 22*(2), 213-228.

Great Britain. Parliament. (2010). *Equality Act 2010: Elizabeth II*. Chapter 15. London: Stationery Office.

Harden, A., Thomas, J., Evans, J., Scanlon, M., & Sinclair, J. (2003). Supporting pupils with emotional and behavioural difficulties (EBD) in mainstream primary schools: a systematic review of recent research on strategy effectiveness (1999 to 2002). In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

Hawken, L. S., O'Neill, R. E., & MacLeod, K. S. (2011). An investigation of the impact of function of problem behavior on effectiveness of the Behavior Education Program (BEP). *Education and Treatment of Children, 34*(4), 551-574.

Horner, R. H. (1994). Functional assessment: Contributions and future directions. *Journal of Applied Behavior Analysis, 27*(2), 401-404.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single subject research to identify evidence-based practice in special education. *Exceptional Children, 71,* 165-179.

Humphrey, N., Kalambouka, A., Wigelsworth, M., Lendrum, A., Lennie, C., & Farrell, P. (2010). New Beginnings: evaluation of a short social–emotional intervention for primary-aged children. *Educational Psychology, 30*(5), 513-532.

Janney, D. M., Umbreit, J., Ferro, J. B., Liaupsin, C. J., & Lane, K. L. (2013). The Effect of the Extinction Procedure in Function-Based Intervention. *Journal of Positive Behavior Interventions, 15*(2), 113-123.

Jones, R. A. (2003). The construction of emotional and behavioural difficulties. *Educational Psychology in Practice, 19*(2), 147-157.

Kamps, D., Wendland, M., & Culpepper, M. (2006). Active Teacher Participation in

Functional Behavior Assessment for Students with Emotional and Behavioral Disorders Risks in General Education Classrooms. *Behavioral Disorders, 31*(2), 128-146.

Kratochwill, T. R. (2003). Task force on evidence-based interventions in school psychology. Retrieved April 16, 2014, from http://madison.k12.in.us/MCSWeb/CSSU/EBI%20Manual.pdf

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L. & Rindskopf, D. M. et al. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Lane, K. L., Smither, R., Huseman, R., Guffey, J., & Fox, J. (2007). A Function-Based Intervention to Decrease Disruptive Behavior and Increase Academic Engagement. *Journal of Early and Intensive Behavior Intervention, 3*(1), 348-364.

Maag, J. W. (2001). Rewarded by punishment: Reflections on the disuse of positive reinforcement in schools. *Exceptional Children, 67*(2), 173-186.

McIntosh, K., Campbell, A. L., Carter, D. R., & Dickey, C. R. (2009). Differential effects of a tier two behavior intervention based on function of problem behavior. *Journal of Positive Behavior Interventions, 11*(2), 82-93.

Miller, L. M., Dufrene, B. A., Olmi, D. J., Tingstrom, D., & Filce, H. (2015). Self-monitoring as a viable fading option in check-in/check-out. *Journal of school psychology, 53*(2), 121-135.

Mong, M. D., Johnson, K. N., & Mong, K. W. (2011). Effects of Check-In/Checkout on Behavioral Indices and Mathematics Generalization. *Behavioral Disorders, 36*(4), 225-240.

Noble, T., & McGrath, H. (2008). The positive educational practices framework: A tool for facilitating the work of educational psychologists in promoting pupil wellbeing. *Educational and Child Psychology, 25*(2), 119-134.

O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist, 41*(3), 181-206.

Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*(4), 357-367.

Partin, T. C. M., Robertson, R. E., Maggin, D. M., Oliver, R. M., & Wehby, J. H. (2009). Using teacher praise and opportunities to respond to promote appropriate student behavior. *Preventing School Failure: Alternative Education for Children and Youth, 54*(3), 172-178.

Restori, A. F., Gresham, F. M., Chang, T., Lee, H. B., & Laija-Rodriquez, W. (2007). Functional assessment-based interventions for children at-risk for emotional and behavioral disorders. *The California School Psychologist, 12*(1), 9-30.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment Scale for Children* (2nd ed.). Circle Pines, MN: AGS Publishing.

Russell, A., Webster, R., & Blatchford, P. (2012). *Maximising the Impact of Teaching Assistants: Guidance for School Leaders and Teachers*. London: Routledge.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research methodology and validation. *Remedial and Special Education, 8*(2), 24-33.

Scruggs, T. E., & Mastropieri, M. A. (1994). The utility of the PND statistic: A reply to Allison and Gorman. *Behaviour Research and Therapy, 32*(8), 879-883.

Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing single-subject research

issues and applications. *Behavior Modification, 22*(3), 221-242.

Seligman, M. E., Ernst, R. M., Gillham, J., Reivich, K., & Linkins, M. (2009). Positive education: Positive psychology and classroom interventions. *Oxford review of education, 35(*3), 293-311.

Shucksmith, J., Summerbell, C., Jones, S., & Whittaker, V. (2007). *Mental wellbeing of children in primary education (targeted/indicated activities).* Teeside: University of Teeside.

Shumate, E. D., & Wills, H. P. (2010). Classroom-based functional analysis and intervention for disruptive and off-task behaviors. *Education and Treatment of Children, 33*(1), 23-48.

Sugai, G., Horner, R. H., Dunlap, G., Hieneman, M., Lewis, T. J., Nelson, C. M., ... Ruef, M. (2000). Applying positive behavior support and functional behavioral assessment in schools. *Journal of Positive Behavior Interventions, 2*(3), 131-143.

Todd, A. W., Campbell, A. L., Meyer, G. G., & Horner, R. H. (2008). The effects of a targeted intervention to reduce problem behaviors: Elementary school implementation of Check In—Check Out. *Journal of Positive Behavior Interventions, 10*(1), 46-55.

Umbreit, J., Lane, K. L., & Dejud, C. (2004). Improving Classroom Behavior by Modifying Task Difficulty Effects of Increasing the Difficulty of Too-Easy Tasks. *Journal of Positive Behavior Interventions, 6*(1), 13-20.

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology, 33*(1), 105-124.

Wilkinson, L. A. (2003). Using behavioral consultation to reduce challenging behavior in the classroom. *Preventing School Failure: Alternative Education for Children and Youth, 47*(3), 100-105.

Wood, B. K., Umbreit, J., Liaupsin, C. J., & Gresham, F. M. (2007). A treatment integrity analysis of function-based intervention. *Education and Treatment of Children, 30*(4), 105-120.

Wyman, P. A., Cross, W., Brown, C. H., Yu, Q., Tu, X., & Eberly, S. (2010). Intervention to strengthen emotional self-regulation in children with emerging mental health problems: Proximal impact on school behavior. *Journal of Abnormal Child Psychology, 38*(5), 707-720.

Yeaton, W. H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology, 49*(2), 156.

Table 1.1

Search Terms

Intervention	Population	Outcomes
School based intervention / Classroom behaviour modification	Elementary school students / Childhood development	Social skills / Interpersonal communication / Communication skills / Social interaction / Social skills training / Behaviour problems / Classroom behaviour / Aggressive behaviour / Behaviour disorders / Conduct disorder

Table 1.2

Weight of Evidence Framework (Gough, 2007)

Weight of Evidence A	**Weight of Evidence B**	**Weight of Evidence C**	**Weight of Evidence D**
Generic judgement about the coherence, quality and integrity of the evidence	Review-specific judgement about appropriateness of the evidence for	Review-specific judgement about the relevance of the focus of the	Overall assessment of the extent to which a study contributes

(Methodological Quality)	answering the review question (Methodological relevance)	evidence for the review question (Study focus relevance)	evidence to answer the review question (Overall weighting)

Table 1.3

Weight of Evidence

Authors	**Methodolog-ical Quality**	**Methodologic-al Relevance**	**Study Focus Relevance**	**Overall Weight of Evidence**
Campbell and Anderson (2008)	Low	Low	Low	**Low**
De Martini-Scully et al. (2000)	Medium	Medium	Medium	**Medium**
Hawken et al. (2011)	Low	Low	High	**Medium**
Humphrey et al. (2010)	High	Medium	Low	**Medium**
Janney et al. (2013)	Medium	High	High	**High**
Kamps et al. (2006)	Low	Low	Low	**Low**
Lane et al. (2007)	Low	Low	High	**Medium**
McIntosh et al. (2009)	Low	High	Low	**Medium**
Miller et al. (2015)	Medium	Medium	Medium	**Medium**
Mong et al. (2011)	Medium	Medium	High	**Medium**
Restori et al. (2007)	High	High	Low	**Medium**
Shumate and Wills (2010)	Medium	Medium	Medium	**Medium**
Todd et al. (2008)	Medium	Medium	Medium	**Medium**
Umbreit et al. (2004)	Medium	Low	High	**Medium**
Wilkinson (2003)	High	Low	Medium	**Medium**
Wood et al. (2007)	Medium	Low	High	**Medium**
Wyman et al. (2010)	Medium	Low	Low	**Low**

Table 1.4

Summary Table of Effect Sizes by Study

Type of Outcome	Primary Outcome*	Study	Sample Size	Effect Size	Effect Size Type**	Effect Size Interpretation	Study Quality
Reducing problematic behaviours	Decrease in disruptive behaviour P1	**De Martini-Scully et al. (2000)**	3	75%	PND	Effective	Medium
	Decrease in disruptive behaviour P2			82%	PND	Effective	
	Decrease in office discipline referrals (ODRs)	**Hawken et al. (2011)**	17	.30	*d*	Small to medium	Medium
	Decrease in problem behaviour	**McIntosh et al. (2009)**	34	1.04	*d*	Large	Medium
	Decrease in disruptive behaviour with an antecedent-based strategy	**Restori et al. (2007)**	8	100%	PND	Very effective	Medium
	Decrease in disruptive behaviour with a consequent-based strategy			88%	PND	Effective	
	Decrease in disruptive behaviour	**Wilkinson (2003)**	1	88%	PND	Effective	Medium
	Decrease in disruptive behaviour	**Lane et al. (2007)**	1	58%	PND	Questionable	Medium
	Decrease in problem behaviour P1	**Todd et al. (2008)**	4	26%	PND	Ineffective	Medium
	Decrease in problem behaviour P2			47%	PND	Ineffective	

Reducing problematic behaviours	Decrease in problem behaviour P3			17%	PND	Ineffective	
	Decrease in problem behaviour P4			0%	PND	Ineffective	
	Decrease in problem behaviour P1	**Campbell and Anderson (2008)**	2	44%	PND	Ineffective	Low
	Decrease in problem behaviour P2			12%	PND	Ineffective	
	Decrease in disruptive behaviours P1 group activities & independent work	**Kamps et al. (2006)**	2	69%	PND	Questionable	Low
	Decrease in disruptive behaviours P2 group activities & independent work			100%	PND	Very effective	
	Decrease in problem behaviour P1	**Miller et al. (2015)**	4	67% .92	PND NAP	Questionable Strong effect	Medium
	Decrease in problem behaviour P2			100% 1.00	PND NAP	Very effective Strong effect	
	Decrease in problem behaviour P3			100% 1.00	PND NAP	Very effective Strong effect	
	Decrease in problem behaviour P4			100% 1.00	PND NAP	Very effective Strong effect	
	Decrease in problem behaviour P1	**Mong et al. (2011)**	4	100%	PND	Very effective	Medium
	Decrease in problem behaviour P2			100%	PND	Very effective	
	Decrease in problem behaviour P3			63%	PND	Questionable	
	Decrease in problem behaviour P4			75%	PND	Effective	
	Decrease in disruptive behaviour P1	**Shumate and Wills (2010)**	3	44%	PND	Ineffective	Medium

	Decrease in disruptive behaviour P2			100%	PND	Very effective	
	Decrease in disruptive behaviour P3			57%	PND	Questionable	
	Decrease in off-task behaviour P1			89%	PND	Effective	
	Decrease in off-task behaviour P2			100%	PND	Very effective	
	Decrease in off-task behaviour P3			100%	PND	Very effective	
Increasing positive behaviours	Increase on-task behaviour in maths and reading	**Umbreit et al. (2004)**	1	100%	PND	Very effective	Medium
	Increase on-task behaviour	**Wood et al. (2007)**	1	75%	PND	Effective	Medium
	Increase on-task behaviour P1	**Janney et al. (2013)**	3	90%	PND	Effective	High
	Increase on-task behaviour P2			100%	PND	Very effective	
	Increase on-task behaviour P3			100%	PND	Very effective	
	Increase in child-reported social and emotional competence scores	**Humphrey et al. (2010)**	253	0.44	*d*	Small to medium	Medium
	Increase in task orientation	**Wyman et al. (2010)**	226	0.33	*d*	Small to medium	Low
	Increase in behaviour control			0.31	*d*	Small to medium	
	Increase in assertive vs withdrawn behaviours			0.37	*d*	Small to medium	
	Increase in peer social skills			0.47	*d*	Small to medium	

Increasing positive behaviours	Increase in academic engaged time	**Lane et al. (2007)**	1	83%	PND	Effective	Medium
	Increase in prosocial behaviour	**McIntosh et al. (2009)**	34	0.99	*d*	Large	Medium
	Increase in time on-task P1 group activities & independent work	**Kamps et al. (2006)**	2	100%	PND	Very effective	Low
	Increase in time on-task P2 group activities & independent work			100%	PND	Very effective	
	Increase in academic engagement P1	**Miller et al. (2015)**	4	100% 1.00	PND NAP	Very effective Strong effect	Medium
	Increase in academic engagement P2			100% 1.00	PND NAP	Very effective Strong effect	
	Increase in academic engagement P3			100% 1.00	PND NAP	Very effective Strong effect	
	Increase in academic engagement P4			33% .87	PND NAP	Ineffective Moderate effect	
	Increase in academic engaged time with an antecedent-based strategy	**Restori et al. (2007)**	8	100%	PND	Very effective	Medium
	Increase in academic engaged time with a consequent-based strategy			93%	PND	Very effective	

**P1 = Participant 1, P2 = Participant 2, etc.*

***PND = percentage of non-overlapping data (Scruggs, Mastropieri & Casto, 1987). See text for explanation of technique. d = Cohen's d (Cohen, 1992). NAP = Non-overlap of all pairs (Parker & Vannest, 2009).*

Appendix A: Inclusion and exclusion criteria

	Inclusion Criteria	Exclusion criteria
1. Intervention	a) The intervention must be targeted individual or small group for children with identified externalising behavioural difficulties	The intervention is not targeted (eg is wholly or in part universal, school-wide or class-wide) and/or is not implemented for children with identified externalising behavioural difficulties (eg a group that may potentially be at-risk for developing difficulties later)
	b) The intervention must be school-based and has no home component other than information or liaison	The intervention includes a parent training component or is a combined home-school intervention
	c) The intervention must be implemented wholly by teachers, learning support assistants (paraprofessionals) or the school's coordinator of special educational needs although intervention may be supported by training or coaching provided by external professionals	The intervention is implemented wholly or partially by external professionals
2. Setting & Participants	a) The intervention must take place in a mainstream school setting	The intervention does not take place in a mainstream school setting (eg a special or residential school)
	b) Participants must be educated within a mainstream classroom for at least some of their timetable	Participants are taught wholly in a special classroom (eg a resource or special education classroom)
	c) Participants must be primary school age children (mean age of participants in the study should be between 4-11 years old)	The mean age of participants in the study is not between 4-11 years old
	d) Participants must be specifically referred for displaying externalising	Participants have a diagnosed neurological, psychiatric, learning or physical disability

		behavioural difficulties but without a diagnosed disorder that might affect their behaviour	that may affect their behaviour (eg autism spectrum disorders, attention deficit hyperactivity disorder, oppositional defiant disorder)
3. Outcome Measures	a)	Primary outcome measures are related to externalising behaviours	Primary outcome measures are not related to externalising behaviours (eg academic achievement or substance abuse)
	b)	Primary outcomes related to measurement of the child's behaviour must be presented	Primary outcomes presented are not measures of the child's behaviour (eg relate to teacher behaviour or perceptions).

Rationale for Inclusion & Exclusion Criteria:

Intervention: (a) Previous meta-analyses have explored universal interventions (eg. Durlak, Dymnicki, Taylor, Weissberg & Schellinger, 2011). (b) School-based interventions are the focus due to concerns of feasibility and effectiveness. (c) School staff that would typically be employed in a UK primary school setting should have delivered the intervention. Studies which were undertaken outside the UK were retained if an equivalent role would be undertaken by a member of school staff typically employed in a UK primary school.

Setting & Participants: The review focuses on early intervention in mainstream primary school settings and therefore studies which sampled children with identified SEND are excluded. Additionally, studies were excluded if participants were educated for their whole academic timetable in special provisions attached to or within mainstream schools.

Outcome measures: (a) The review aims to explore the effects of intervention on externalising behaviour difficulties in particular. (b) Studies which operationalised a change in a teacher's behaviour or attitude as the primary outcome variable were excluded.

Appendix B: Weighting of Studies

A: Methodological Quality

Based on the rating given to each of the studies according to adapted coding protocols for single-participant designs (Kratochwill et al., 2010) and group experimental and quasi-experimental designs (Gersten et al., 2005).

B: Methodological Relevance

Considers whether the methodological design was suitable for evaluating the effect of school-based interventions on the externalising behavioural difficulties of primary-aged children. Ratings were based on evidence of the following: a) maintenance or follow-up sessions, b) a measure for documenting fidelity of implementation for the intervention, c) information about how school staff were trained or supported to deliver the intervention, and d) a context for the need for intervention should be established, either by documenting participants' lack of response to prior interventions or by including data for comparative or control students which indicates target students' elevated behavioural difficulties.

C: Study Focus Relevance

Considers whether the focus and character of the study contribute towards answering the review question. Ratings were based on evidence of the following: a) data on the social importance of the effects of intervention from the perspective of both school staff and the child, perhaps in the form of a consumer satisfaction measure, b) the social appropriateness of procedures established by documenting implementer feedback on the acceptability of the intervention, c) demonstration of ecological validity by establishing the contextual basis for the intervention through consultation or formative research with study participants, and d) consideration of rival hypotheses for the child's behaviour difficulties either in the formative research phase and/or in the interpretation of results phase.