Evaluating health visitor parenting support: validating outcome measures for parental self-efficacy
Karen A. Whittaker and Sarah Cowley

*J Child Health Care* 2006 10: 296
DOI: 10.1177/1367493506067882

The online version of this article can be found at:
http:// chc.sagepub.com/content/10/4/296

Published by:
SAGE
http:// www.sagepublications.com

On behalf of:
Association of British Paediatric Nurses

Additional services and information for *Journal of Child Health Care* can be found at:

Email Alerts: http:// chc.sagepub.com/cgi/alerts

Subscriptions: http:// chc.sagepub.com/subscriptions

Reprints: http:// www.sagepub.com/journalsReprints.nav

Permissions: http:// www.sagepub.com/journalsPermissions.nav
Evaluating health visitor parenting support: validating outcome measures for parental self-efficacy

KAREN A. WHITTAKER, BN, MSc, PGCE, RHV
Senior Lecturer, Department of Nursing, University of Central Lancashire, Preston, Lancashire, UK

SARAH COWLEY, BA, PGDE, PhD, DipHV
Head of Public Health and Health Services Research Section, Florence Nightingale School of Nursing and Midwifery, King's College, London, UK

Abstract

Parenting support has become an increasing feature of child health services within the United Kingdom but typically, outcome measures available for testing the effectiveness of parenting interventions have been developed and validated elsewhere. This article reports the results of a feasibility study testing the Parenting Self-Agency Measure (PSAM) and subscales from the Self-Efficacy for Parenting Tasks Index (SEPTI) as outcome measures for UK-based parenting support programmes. Forty-six mothers and 10 fathers accessing routine health visitor and school nurse services participated in the test–re-test of the scales and commented separately on the acceptability of scale questions. Very large intra-class correlation results indicated good repeatability but alpha coefficient scores and factor analysis results suggest that UK respondents may not recognize SEPTI subscales items as measuring single dimensions. The PSAM was a more stable measure of parenting self-beliefs than the SEPTI subscales when tested with a UK sample of parents.

Keywords  child health • evaluation • measurement • parenting support • self-efficacy
Introduction

Parenting education and family support has become one of the fastest growing policy and practice topic areas in recent years in the United Kingdom (Moran et al., 2004). It is a key feature of social programmes such as Sure Start in the UK and Head Start in the United States (Henricson, 2003) that, if successful, can have major implications for improving the health and well-being of whole communities (Stewart-Brown, 2000). However, before research studies can begin to determine the success of such programmes, there is a need to question the usefulness of available outcome measures that may have been developed originally with different cultures to the ones currently being studied. This article acknowledges this concern by presenting the findings from a feasibility study designed to test the suitability of US-developed parental self-efficacy scales for use with parents in the UK. It begins with an introduction to the concept of self-efficacy and its application to parenting. The reliability and validity test results for the scales used with a UK sample of parents are presented. These are discussed and suggestions are made for how the assessed scales might be used with future samples of parents.

Self-efficacy and parenting

Bandurian perspectives of self-efficacy are derived from assumptions about the agent’s ability to anticipate and control personal actions and thereby act intentionally (Bandura, 1997, 2001a). Therefore a sense of personal efficacy – that is, a belief in personal capability to make things happen – is derived from human agency. Parental self-efficacy can be used as an expression of the extent to which parents have developed increasing self-belief in their ability to carry out different parenting tasks and responsibilities.

Theoretically, perceptions of parental self-efficacy for certain tasks will predict intended parenting behaviour. Problems arise when there is a failure to match intentions with behaviours either due to a lack of understanding or knowledge. For example, a parent may fail to appreciate that a toddler is unable to understand instructions when the parent shouts angrily. Instead, the child cries in fright and fails to pick up the message intended by the parent. The crying child is not an intended outcome but, in Bandurian terms, is a consequence of the parent behaviour. Successful parenting support programmes would help parents to recognize such mismatches by offering education about child age-appropriate abilities and guidance with adult behaviours. When parents then practise and subsequently master what they have learned, parental self-efficacy improves and reinforces the continued use of more positive parenting practices (Sanders and Woolley, 2005; Teti and Gelfand, 1991).
Measuring parental self-efficacy

Bandura (2001b) asserts that there is no general measure of self-efficacy, since the concept should be understood in terms of a particular domain of functioning. Certainly, the general self-efficacy measures that are available (Schwarzer, 2003; Sherer and Adams, 1983), do seem to focus on a broad sense of competence rather than functioning within specified domains, such as parenting. They are problematic because they assume that successful experiences can be ‘banked’ and automatically transferred to new situations with good effect regardless of altered contexts, when in fact parenting is a context-specific activity (Ghate and Hazel, 2002; Quinton, 2004), constantly subject to external influences that can challenge performance.

In critiquing parenting research using the self-efficacy construct, Coleman and Karraker (1997) offer a comprehensive review of published scales. They distinguish between ‘task-specific’, with a tight focus on specified parenting tasks, and ‘domain general’, which considers the broad domain of parenting. Examples of the former (Ballenski and Cook, 1982; Deutsch et al., 1988) are limited by their failure to capture the many and varied components of parenting, while the latter, (Dumka et al., 1996; Johnson and Mash, 1989) ignore the chance that some parents may feel capable in different areas of parenting. The alternative favoured and used by Coleman and Karraker (2000, 2003) is a ‘domain-specific’ approach, which groups the tasks into sub-themes and combines these to give an overall score.

Method

Study aim

The failure to identify any available UK-developed parenting self-efficacy scales despite a search of the electronic databases Medline, Cinahl and Psych-info, emphasized the need to test those that are available before using them in the UK. The study aim was to assess whether identified scales were suitable for use within a parenting support evaluation study questionnaire. More specifically, the study tested the reliability and validity of chosen parental self-efficacy scales and identified respondent views about the chosen scales.

The scales chosen for use in this study were the domain-general Parenting Self-Agency Measure (PSAM) (Dumka et al., 1996) and subscales from the toddler and school versions of the domain-specific Self-Efficacy for Parenting Tasks Indexes (SEPTI) (Coleman and Karraker, 2000, 2003). Final decisions about their use were based on the reliability and validity data available, the scale length and style of English language used.
Ethical considerations

Permission to proceed with the feasibility study was granted by the local research ethics committee; in addition, a study advisory group consisting of lay, professional and academic representatives provided additional governance.

Study design

The feasibility study was designed as a test–re-test of two (toddler and school versions) parenting self-completion questionnaires, incorporating the PSAM and SEPTI subscales. The repeat testing was used to assess the degree of consistency in responses over time. Eighteen health visitors and school nurses practising within geographical zones providing a positive parenting service in a north-west England city were each asked to recruit four parents who were accessing child health services routinely. Each practitioner was provided with a pack containing a practitioner information sheet and details about study exclusion criteria. Parents were excluded who were under 16 years of age, involved in child protection proceedings, had identified learning needs or mental health difficulties. Also excluded were those unable to speak or read English, as the questionnaire terminology could not be translated directly.

Data collection

The practitioners provided parents with a study information sheet and consent form and those recruited were asked to complete either a toddler or school version of the questionnaire twice, one week apart. This time period was felt to be sufficiently long for respondents not to recall their original responses while being short enough to avoid intervening circumstances influencing self-efficacy experiences. The participants also completed a questionnaire comments form after the first application. This was designed to provide feedback from parents on the acceptability, relevance and meaning of questionnaire items.

Questionnaire scales

The scales used within the questionnaire were placed on separate pages and small illustrations were interspersed between them to indicate a change of questionnaire theme and improve the overall appeal. Permission to use the scales and confirmation of question wording as detailed in earlier publications (Coleman and Karraker, 2000, 2003; Dumka et al., 1996) was granted from the original authors. Both the toddler and school versions of the questionnaire included the five-item PSAM, making it possible to assess the suitability of this shorter domain general scale for use with parents of different-aged children. Respondents identified how often their reaction corresponded to each item listed using a five-point Likert scale, ranging from never = 1 to always = 5. Total scores could
range from 5 to 25, with lower scores indicating lower self-efficacy. Dumka et al. (1996) had validated the PSAM previously against the adult coping scale containing the constructs of active coping, parenting acceptance and parenting inconsistent discipline and achieved a Cronbach’s alpha coefficient of 0.70 when testing reliability.

By comparison, the two SEPTI scales were age-specific either to parenting toddlers (aged one to four years) or schoolchildren (aged five to 10 years), creating the need for two versions of the questionnaire. In an effort to keep the questionnaire short and user-friendly, the full SEPTI scales were not used and only subscales with content most relevant to the positive parenting course (Taylor, 2000), already used by the host health care trust, were included. However, this did alter the SEPTI from ‘domain-specific’ to ‘task-specific’ measures of self-efficacy.

The SEPTI subscales use a six-point Likert scale, ranging from ‘strongly agree’ to ‘strongly disagree’. Coleman and Karraker’s (2000, 2003) study validated these against the Johnson and Mash (1989) Parenting Sense of Competence Scale, reporting Cronbach’s alpha coefficients for each of the subscales. For the SEPTI-toddler scale (SEPTI-TS) these were: discipline (seven items) = 0.81; play (seven items) = 0.92; routine (eight items) = 0.46. The SEPTI-School Cronbach’s alpha coefficients were: discipline (eight items) = 0.86; recreation (seven items) = 0.82; nurturance (seven items) = 0.77. Again, to minimize questionnaire length, items with reported factor loadings below 0.40 (Coleman and Karraker, 2000) were omitted from the school subscales for discipline (one item), recreation (one item) and nurturance (three items). Total possible scores for toddler subscales ranged from 7 to 42 for discipline and play; and 8 to 48 for routine. For school subscales, these were: discipline, 7 to 42; recreation, 6 to 36; nurturance, 4 to 24.

Data analysis

The inclusion of both the PSAM and SEPTI scales allowed comparisons to be made between the two different measures of self-efficacy, but strong relationships were not expected. This was because one was ‘domain general’ and the others were ‘task-specific’ measures. Equally, it was anticipated that there would be variance between the task-specific scales (SEPTI subscales), since parents might not believe themselves to be equally self-efficacious in all parenting tasks. In order to assess the suitability of the PSAM for parents of toddlers and school-age, children a comparison of total mean scores was made using the unrelated t-test.

Demographic and outcome measure data were analysed using the Statistical Package for the Social Sciences (SPSS) version 10. To examine the repeatability of scales used with these non-random and small samples, the intra-class correlation coefficients (ICC) with 95 percent confidence intervals (CI) were computed as indicators of reliability (Bland and Altman, 1996; Hopkins, 2000). To interpret the ICC results the descriptors for different-sized correlation coefficients detailed by Hopkins (2000) were used: 0.3–0.5 = moderate, 0.5–0.7 = large, 0.7–0.9 = very
large and above 0.9 = nearly perfect. Differences between repeat measures were analysed using the paired $t$-test. Intercorrelations between the measurement scales were computed to assess convergent validity and an indication of scale similarities (Streiner and Norman, 2003).

Alpha coefficients were computed to identify scale internal reliability and consistency, where higher values of at least 0.7 were considered to be acceptable. The principle components analysis with varimax rotation (Bryman and Cramer, 2001) was used to highlight any unnecessary questions and thus the construct validity of each scale. Finally, the parents’ opinions about the suitability of the questions and general questionnaire layout were assessed using a short respondents’ comments form.

**Results**

**Respondents**

In total, 72 questionnaire packs (38 toddler and 34 school-age) were given to practitioners for distribution to parents. Fifty-six parents (78%) returned the first application questionnaires (30 toddler and 26 school-age) but fewer ($N=45$, 62%) returned the second application. Forty-eight parents (67%) completed the comments form. The feasibility study demographic data illustrate that respondents were largely white women ($N=46$, 82%) who lived with a partner ($N=49$, 88%) and who had not experienced household changes in the previous six months ($N=53$, 95%). Half ($N=29$, 52%) were in paid employment and relatively few ($N=9$, 16%) cared for additional children on a regular basis. The age of parents ranged from 16–50 years, although half were aged between 26 and 35 years. Comparison of PSAM scores for parents of toddlers and school-age children (see Table 1) indicate that there were no statistical differences ($t = -0.035$, $df = 54$, $p = 0.97$) for these two types of respondent.

**Scale repeatability**

Descriptive statistics were calculated to summarize PSAM and SEPTI subscales scores (see Table 1). The test–re-test ICCs were within the very large range for all the scales, although the wide CIs for the SEPTI-TS subscales and the SEPTI-School nurturance scale demonstrate the potential for the ICCs to fall below the 0.7 level. The observed mean differences were moderate for these subscales and the SEPTI-School recreation scale, but small for the PSAM and the SEPTI-School discipline scales.

The small observed differences were confirmed by the paired samples $t$-test results. None of the paired $t$-test results were statistically significant; however, where the observed mean differences were greater, corresponding $t$-values tended to be larger.
Scale internal consistency

Alpha coefficient scores help to identify whether a multiple-itemed scale is measuring a single concept, where higher scores indicate greater consistency among the items. The results were: PSAM = 0.76, SEPTI-TS subscales for discipline = 0.66 (improved to 0.79 with questions 4 and 7 removed), play = 0.84 and routine = 0.7 (improved to 0.77 with questions 17 and 20 removed). SEPTI-School subscales for discipline = 0.86, recreation = 0.88 and nurturance = 0.58.

Scale construct validity

The factor analysis exercise was used to provide evidence of whether subscale items were measuring elements of the same components, or whether there was a
tendency for them to measure different things. The results indicated that the PSAM factored to a single component, thus indicating that scale items were measuring the same parenting dimension; whereas the SEPTI-TS discipline subscale factored to two components, suggesting that two of the items were not related to each other five items and may not have been measuring discipline in parenting. These were the reverse scored items, 4: ‘Setting limits for my toddler is relatively easy for me’, and 7: ‘I allow my child enough freedom to actively explore the environment’. The play subscale factored to a single component and the routine subscale to three components. Those on the routine subscales factoring to the third component were both items concerning the child’s diet i.e. items 17: ‘I have been successful in getting my child to eat on a fairly regular schedule’ and 20: ‘Although I have tried to train my child to eat well, my efforts have been met with very little success’. Otherwise, four of the routine subscale items factored to the first component and two items to the second component.

Of the SEPTI-School subscales, only that measuring recreation factored to a single component, while the discipline and nurturance subscales factored to two components. Here it was the reversed scored items in both scales (one item in the discipline scale and two items in the nurturance scale) that factored to the second component.

**Relations between self-efficacy measures (convergent validity)**

The intercorrelations between the scales (Tables 2 and 3) ranged between small (≤0.3) and large (0.5–0.7) relationships. For the toddler group, the strongest relationship existed between the PSAM and play subscale. Discipline and routine were moderately related and each demonstrated only weak relationships with the PSAM and play subscale. For the school-age group, discipline and nurturance demonstrated stronger relationships with the PSAM.

<table>
<thead>
<tr>
<th>Spearman’s rho</th>
<th>PSAM</th>
<th>SEPTI-TS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Discipline</td>
<td>Play</td>
</tr>
<tr>
<td>SEPTI-TS</td>
<td>.255</td>
<td>.530**</td>
</tr>
<tr>
<td>N</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>Play</td>
<td>.318</td>
<td>.457*</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

* Correlation is significant at the .05 level (2-tailed)
** Correlation is significant at the .01 level (2-tailed)
Respondent comments

The average time reported to complete the questionnaire was 15 minutes (range 2–35) and was felt to be generally acceptable. Four respondents indicated that questions were not always straightforward and therefore needed careful thought, as one stated:

Well thought out questions, but not straightforward to answer, I needed to really think about my own behaviour/relationship with [their] daughter (it’s given me food for thought!)

Another inferred that the choice of language could pose challenges for some potential respondents:

Some of the language may be complex when English is not [the] first language.

Most parents returning comments forms indicated that the questionnaire content was satisfactory, with only one parent identifying a question for removal. Eleven (23%) parents identified that one or more of the scale items were confusing. This was commonly item 7: ‘I allow my child enough freedom to actively explore the environment’ from the SEPTI-TS, or item 14: ‘I meet my own expectations in terms of providing emotional support for my child’ from the SEPTI-School scale.

When asked about the illustrations scattered throughout the questionnaire, 30 (73%) parents commented how they had liked them or found them to be fine. One-quarter (N = 12) of those returning comments forms stated that they had not noticed the illustrations and three parents made no comment at all.

Discussion

This study assessed the feasibility of using US-developed outcome measures with a UK population accessing National Health Service (NHS) parenting and child

Table 3: PSAM and SEPTI-School subscales correlation

<table>
<thead>
<tr>
<th>Spearman’s rho</th>
<th>PSAM Discipline</th>
<th>SEPTI-School Recreation</th>
<th>SEPTI-School Nurturance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discipline</td>
<td>.657**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recreation</td>
<td>.382</td>
<td>.250</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>26</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Nurturance</td>
<td>.590**</td>
<td>.531**</td>
<td>.641**</td>
</tr>
<tr>
<td>N</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
</tbody>
</table>

** Correlation is significant at the .01 level (2-tailed)
health services. The absence of available UK-developed scales had made this necessary. However, a team in the south of England (Bloomfield et al., 2005) are rectifying this with the Tool to Measure Parenting Self-Efficacy (TOPSE) pilot.

The PSAM attempts to tap the existence of human agency and therefore the extent to which parents recognize themselves as instruments of their own actions. While it is not identical to self-efficacy, it offers a conceptual parallel, as personal self-efficacy beliefs are dependent on the existence of human agency. Simultaneous use of PSAM and SEPTI subscales thus provided an opportunity to appraise convergent validity. As anticipated, and similar to Coleman and Karraker’s (2000) results, most of the relationships with the PSAM were at a moderate level. Further evidence for the validity of the PSAM was provided by the results from the factor analysis. These, along with the alpha coefficients and the test–re-test results from this UK sample, add support to Dumka et al.’s (1996) findings about the robustness of the five-item version of this scale as a measure of parents’ confidence in their own abilities to act successfully.

In contrast with the PSAM, the SEPTI scales are age-related and task-specific to acknowledge how parenting tasks vary with a child’s age. While the SEPTI-TS ICCs seem reassuring, the discipline was the least stable subscale, with weaker alpha coefficient and test–re-test reliability statistics than those for play and routine. The internal consistency of the discipline subscale was improved with the removal of items 4: ‘Setting limits for my toddler is relatively easy for me’ and 7: ‘I allow my child enough freedom to actively explore the environment’, which not only factored differently to other items but were also the only reversed scored items for this subscale. As previously noted, item 7 was identified in the respondents’ comments as a confusing statement, so it is plausible that reversing the statements threatened the reliability of responses.

The factor analysis results raise questions about the appropriateness of including items that concerned the child’s diet within the routine subscale. It is possible that child eating habits are construed in terms of diet quality as opposed to mealtime routines, hence these items represent a separate construct. Given this, the content validity of the routine subscale might be improved if items 17: ‘I have been successful in getting my child to eat on a fairly regular schedule’ and 20: ‘Although I have tried to train my child to eat well, my efforts have been met with very little success’ were expressed as mealtime routines as opposed to eating habits.

For the school questionnaire the recreation subscale appeared the weakest, with poor alpha coefficient scores, factor analysis results and respondent comments which cast some doubt over items that are perhaps linguistically more suited to US than UK audiences. However, a premature reduction of this scale to four items (based on Coleman and Karraker’s earlier data), did limit the opportunity to explore whether other items would have factored in a similar way. Despite this limitation, all the school subscale ICC results do suggest sufficient test–re-test reliability. It also seemed that the school discipline subscale was
more stable than its toddler counterpart. This might be explained by the greater experience of, and thereby possible clarity over, discipline practice that parents of older children are likely to have.

In general, it would seem that the questionnaires that contained the self-efficacy scales and were set out with illustrations to improve visual appeal were acceptable to the UK sample. The feedback from the respondents also highlights how care must be taken when introducing the questionnaire to parents, given that many of the items require careful thought. Failure to acknowledge this could threaten reliability, and is therefore an important consideration when planning data collection methods.

Study limitations

The omission of specific guidance on when to complete each application of the questionnaire may have allowed the introduction of bias. The concern is that if the first and second applications were completed in differing social situations, exposure to events such as the presence of a practitioner or a recent toddler tantrum could have an impact on self-efficacy experiences. Indeed, Sanders and Woolley (2005) recently highlighted how context can influence maternal efficaciousness, particularly when entering situations that present competing demands. Moreover, the small and likely heterogeneous nature (in terms of educational and socio-economic situations) of the sample used in this UK study should be borne in mind when considering the scale validity and reliability for different subgroups of the UK population.

Conclusions

The data provide some reassuring results for using the US-developed PSAM and SEPTI subscales as outcome measures with a UK population accessing routine pre-school and school health services. They also confirm the importance of pilot testing when the study participants are likely to differ culturally from those involved in the original development of the scale.

The test–re-test results gave an indication of the reliability of the scales and the extent to which a parenting programme evaluator could be confident in the ability of the scales to measure in the same way on repeated applications. In this study greater confidence was found in the repeatability of the PSAM than the SEPTI scales. The potential bias introduced by measuring on separate occasions suggests that careful instruction regarding when each self-report measure should be completed might be helpful in encouraging questionnaire completion during similar parenting circumstances, for example, ‘while your child takes a nap’. However, during a parenting programme this might be inherently difficult if the parenting experience has altered between measurements.
Equally, the comparison of the PSAM and SEPTI scales produced results that contribute to the debate about domain general or task-specific measures for parenting self-efficacy. No evidence was found in this study that the domain general measure would discriminate between parents of differently-aged children. However, correlation coefficient data suggest that domain general scores do not relate equally to the different task-specific scores. Moreover, the results from the UK sample support Coleman and Karraker’s (2000) claim that parents can have varying degrees of self-efficacy dependent on parenting tasks, hence the need for a domain-specific parenting self-efficacy measure that combines task-specific scales into a single measure of parental self-efficacy.

Acknowledgements
The authors are grateful to Dr Chris Sutton for provision of statistical guidance and to Preston Primary Care Trust and the National Coordinating Centre for Research Capacity at the Department of Health for research funding.

References


KAREN A. WHITTAKER lectures at the University of Central Lancashire and holds an NHS research training fellowship to support part-time doctoral study at King’s College, London. This research has been completed as part of her doctoral study examining parental self-efficacy experiences for users of positive parenting services.

SARAH COWLEY is Professor of Community Practice Development at King’s College, London. Her research, teaching and publications reflect her professional background in health visiting and interest in public health and positive health, especially in relation to needs assessment, families and the social environment.

**Correspondence to:**

Karen A. Whittaker, Department of Nursing, University of Central Lancashire, Preston PR1 2HE, UK. [email: kwhittaker1@uclan.ac.uk]