

Assessing Parenting Practices Through Parent-Report and Direct Observation During Parent-Training

David J. Hawes, Ph.D.^{1,3} and Mark R. Dadds, Ph.D.²

Published online: 20 May 2006

The specific parenting domains measured by the Alabama Parenting Questionnaire (APQ) make it particularly relevant to interventions concerned with the modification of parenting practices. This study assessed the validity and clinical utility of parent reports on the APQ using observational data of parents and children (N = 56, aged 4–8 years) participating in a parent training intervention for childhood conduct problems. Parent reports on the measure were found to converge well with observations of parents' use of praise, and harsh/laversive parenting. APQ scores also reflected change in parenting practices across treatment, and were associated with clinical child outcomes. Comparisons of the five original APQ subscales with a three-factor empirically-derived form of the measure indicated greater support for the original subscales, which were found to be valid and clinically informative in the treatment of childhood conduct problems.

KEY WORDS: Alabama parenting questionnaire; observational assessment; parenting practices; parent training; conduct problems.

Poor parenting practices represent some of the most robust risk factors for conduct problems in childhood and adolescence (see [Hawes & Dadds, 2005](#)). Lack of parental involvement, poor monitoring and supervision, and harsh and inconsistent discipline, have all been established as strong predictors of antisocial outcomes in children and adolescents. Accordingly, the most effective interventions for conduct problems are those that modify such practices ([Brestan & Eyberg, 1998](#)). Despite the central role of parenting practices in models of antisocial behaviour and associated treatments, the measurement of parenting practices has

¹Research Associate, School of Psychology, University of New South Wales, Sydney, Australia.

²Professor of Psychology, School of Psychology, University of New South Wales, Sydney, Australia.

³Correspondence should be directed to David J. Hawes, School of Psychology, University of New South Wales, Sydney, NSW, 2052, Australia; e-mail: D.Hawes@unsw.edu.au.

received relatively limited attention. This study examined the psychometric properties of self-reported parenting practices and data from the direct observation of parents participating in a parent-training intervention.

Observational measures have played a prominent role in the study of childhood conduct problems, providing the data upon which Patterson's influential coercion model was based. In an effort to collect objective behavioural data concerning family processes, Patterson, Reid and colleagues developed coding systems for recording the moment-to-moment interactions between parents and children during innovative naturalistic-observational studies. This research often utilised experimental treatment designs to study the effects of various parenting behaviours on child behaviour (see Reid, Patterson, & Snyder, 2002 for detailed descriptions). These studies revealed that the parents of conduct problem children often employ methods of discipline that model antisocial behaviour and inadvertently reinforce deviant child behaviour. This research also highlighted the transactional processes by which children's responses to these methods reinforce their continued use by parents (Snyder & Stoolmiller, 2002). This transactional model underlies a range of efficacious behavioural parent training interventions for conduct problems, (e.g., McMahon & Forehand, 2003; Webster-Stratton & Hammond, 1997), and remains central to current theories of antisocial behaviour.

While observational methods are regarded as a gold standard in the assessment of parenting, the complexity and expense associated with these methods often preclude their use in clinical settings. Self-report measures represent a more feasible alternative, however those available have been largely limited to global measures of constructs such as parenting stress or competence. The recently developed Alabama Parenting Questionnaire (APQ; Shelton, Frick, & Wootton, 1996) is an exception, assessing specific parenting domains in both child and parent report formats. The APQ consists of five subscales corresponding to the dimensions of parenting associated with risk for conduct problems: poor parental monitoring and supervision, inconsistent punishment, corporal punishment, positive parenting, parental involvement.

Scores on the APQ have been found to discriminate between the parenting received by children (aged 6–13 years) in clinic and community samples (Shelton et al., 1996). The reliability of the measure has also been established within a representative community sample ($n = 1359$; aged 4–9 years), with Dadds, Maujean, & Fraser (2003) reporting moderate to high internal consistency and good test-retest reliability for the APQ subscales. Evidence of external validity was also reported, with each of the subscales correlating in predicted directions with parents' reports of child conduct problems on the Strengths and Difficulties Questionnaire.

Using an alternative factor structure of the APQ, Hinshaw et al. (2000) found it to be useful in explaining clinical outcomes in the long-term multimodal treatment of ADHD. Principle components analysis conducted on the pre-treatment parent-report APQ scores in this sample

produced an interpretable three factor structure comprising: Positive Involvement (Cronbach's $\alpha = .85$), Negative/Ineffective Discipline (Cronbach's $\alpha = .70$), and Deficient Monitoring (Cronbach's $\alpha = .72$). This factor structure was shown to have some predictive validity in this clinical context, with the effects of combined pharmacological and behavioural intervention on children's teacher-reported social skills found to be partially mediated by change in parents' scores on the Negative/Ineffective Discipline factor.

A similar three-factor structure was reported by [Elgar, Waschbusch, and Dadds \(submitted\)](#). In the course of developing a short form of the APQ, Elgar et al. (submitted) factor analysed the community sample data reported in [Dadds et al. \(2003\)](#) and found limited support for the Parental Involvement and Corporal Punishment factors. Parental Involvement was found to overlap with Positive Parenting, and the three items comprising the Corporal Punishment scale were deemed insufficient to reflect a unified component.

Despite the role that observational measures have played in establishing the importance of parenting practices in relation to child conduct problems, parent report data on the APQ has not been validated against data from the direct observation of parent behaviour. Intervention designs, such as those utilised by Patterson and colleagues represent an ideal context in which to examine convergence between observational data and parent-reports of parenting practices. Such designs allow also for further examination of the relationship between self-reported change in parenting practices and clinical child outcomes.

The first aim of this study was to assess the external validity of self-reported parenting practices on the APQ against observations of parent-child interaction in families of young conduct problem boys participating in a parent training intervention. An additional aim of the study was to examine the clinical utility of the measure for use in parent-training. In such a context, the clinical utility of a measure such as the APQ could be defined largely by the extent to which it is sensitive to change in parenting. To address this aim, pre- to post-treatment change in APQ scores was examined, as was the relationship between APQ-measured change in parenting and child outcomes. Finally, due to the evidence that a three-factor structure may be more applicable to the APQ than the five theoretically-based subscales ([Hinshaw et al., 2000](#); [Elgor et al., submitted](#)), a further aim was to examine the comparative validity of both alternatives.

In relation to both the original five subscales of the APQ, and the alternative three-factor structure proposed by [Hinshaw et al. \(2000\)](#), the following was hypothesised. Firstly, subscale scores were predicted to converge with concurrent observations of positive and negative parent behaviour, both at pre-treatment and post-treatment. Secondly, it was predicted that the APQ would be sensitive to clinical change in parenting from pre- to post-treatment. Finally, it was predicted that change in APQ scores across treatment would be associated with clinical child outcomes, after controlling for baseline levels of conduct problems.

METHOD

Participants

Participants were boys aged 4 to 8 years who met DMS-IV criteria for either Oppositional Defiant Disorder or Conduct Disorder. Treatment was conducted in the psychology clinics of Griffith University and the University of New South Wales, in Brisbane and Sydney, Australia. Permission to conduct the research was provided by the human research ethics committees of both universities. Participants self-referred or were referred by community health services, between April 2002 and October 2003. Children receiving concurrent psychological treatment were not eligible, nor were those with developmental disabilities. In order to focus on conduct problems most suited to behavioural (rather than pharmacological) intervention, cases with primary diagnoses of ADHD were excluded. Secondary features of ADHD were permitted if currently medicated.

Fifty-six families commenced treatment, with the target children having a mean age of 6.29 years ($SD = 1.55$). Total family income ranged from <\$20,000 (7%), \$20–30,000 (12%), \$30–50,000 (26%), to over \$50,000 (55%). Education in parents ranged from junior certificate (16%) through a mode of 'finished high school' (40%), to university educated (31%). The majority of families (76%) comprised two caregivers. Six families dropped out of treatment within the first three sessions, and the data from one further case were excluded due to marital stress requiring a significant deviation from the treatment protocol. The intention to treatment sample ($n = 56$) was split into completer ($n = 49$) and non-completer ($n = 7$) groups, and compared across demographic variables, child age, and pre-treatment conduct problem measures in a MANOVA. As no between group differences were found, non-completers were excluded from statistical analysis.

Measures

The Alabama Parenting Questionnaire (APQ; Shelton et al., 1996) was used to assess self-reported parenting practices. As the focus of this study was on the use of the parent-report form with younger children, only the parent self-report form was used. The APQ was completed by the child's primary caregiver, who in the vast majority of cases was the mother. The APQ consists of 42 items presented with a 5-point endorsement scale (Never, Almost Never, Sometimes, Often, Always). As stated earlier, this study examined the measure in both the original five-subscale form (Poor Monitoring and Supervision, Inconsistent Discipline, Corporal Punishment, Positive Parenting Techniques, Parental Involvement), as well as the three-factor structure (Positive Involvement, Negative/Ineffective Discipline, Deficient Monitoring) as reported by Hinshaw et al. (2000).

Observational data of parent-child interactions were collected using the Behavioral Observation Coding System: Family Observation Schedule (FOS 5th ed.; Dadds & McHugh, 1992). This time-sampling protocol has been used often in parent training research (e.g., Sanders, Markie-Dadds, Tully, & Bor, 2000), with good reliability and validity consistently demonstrated. The system provides a framework for the scheduling of family interaction tasks, and the recording and categorising of both parent and child behaviours. Observation periods were divided into 'observe' and 'record' intervals, lasting 20 s and 10 s respectively. This cycle repeated for the total duration of the observation. Using headphones, observers listened to a CD recording that signalled the start and end of each time interval with a series of tones. During the 'record' interval, observers ticked the codes for the behaviors occurring in the previous 'observe' interval. Using this method, only the presence or absence of these behaviors was noted, not the frequency of each behavior during the interval. Child behaviors coded included non-compliance, complaints, demands, physical aggression, and general oppositional behavior. Parent behaviors included praise, physical contact, questions, instructions, and social attention. All of these (except for praise) were recorded with affect indicators, with ' - ' signifying an aversive tone (e.g., frustrated, angry, rebuking), and ' + ' indicating a positive or neutral affect. Observations of parent implementation of the techniques taught in treatment were also recorded.

A number of variables were calculated from the raw observational data. 'Conduct problems' was the percentage of total observation intervals during which time any child behavior codes were recorded. 'Aversive parent behavior' was the percentage of total parent-child interaction intervals during which any parent behaviors with negative affect indicators were recorded. 'Praise' was the percentage of total parent-child interaction intervals during which praise was used by parents, and 'correct implementation' was the percentage of conduct problem intervals in which parents initiated the behaviour correction routine without engaging in aversive behaviour. Each observation involved two components, a play and a dinner setting. In the play observation, the primary caregiver parent was observed interacting with the referred child in periods of free play (10 min), structured play (10 min), and tidying up (5 min). For the dinner observation, all family members were observed during their typical dinner routine, with observational data recorded only for the parents and referred child. One third of all observations were conducted by two observers for the purpose of calculating inter-rater reliability. All observers were blind to parents' APQ scores.

Diagnostic interviews were conducted using the Diagnostic Interview Schedule for Children, Adolescents, and Parents (DISCAP) (Holland & Dadds, 1997). This semi-structured interview is based on DSM-IV (American Psychiatric Association, 1994) criteria for childhood disorders, and demonstrates good reliability and validity (Johnson, Barrett, Dadds, Fox, & Shortt, 1999). The DISCAP is designed to assign DSM-IV diagnoses, and identify sub-clinical features of DSM-IV disorders, providing both categorical (i.e., diagnosis/no diagnosis) and continuous

(i.e., clinical symptom severity from 0–6) data. Pre-treatment DISCAP interviews were conducted by the treating therapist, while those at post-treatment and follow-up conducted by clinical psychologists unfamiliar with the case. 30% of interviews were conducted by two interviewers, positioned on separate telephone lines and kept blind to each other's written notes and diagnoses, in order to check inter-rater reliability of diagnoses.

Procedure

Following a screening interview in which inclusion criteria were addressed and the ODD component of the DISCAP administered, eligible families attended an initial assessment session. During this session a comprehensive diagnostic interview was conducted, including full administration of the DISCAP in order to confirm the screening diagnosis and identify additional psychopathology relevant to the inclusion criteria (e.g., developmental delay, ADHD). Parent-reports on the APQ were also completed at this time, as was the first home observation assessment. Informed consent was also obtained. Post-treatment assessment occurred in the week following the final treatment session, and consisted of DISCAP interview, home observation, and parent-report APQ.

Intervention

Treatment consisted of a fully manualised parent training intervention based on the empirically-validated intervention by Sanders and Dadds (1993). The intervention commenced with a 1.5 hr assessment session with parents, followed by nine weekly 1 hr sessions. In addition to the primary focus on child behavior management, the protocol provides a systemic intervention addressing parent and family issues impacting on child adjustment (e.g., parent stress, relationship discord). In order to maintain the flexibility with which parent training is delivered in the real world, treatment sessions were repeated with participants when appropriate, up to a limit of 3 repeated sessions. Treatment was conducted by clinical psychologists with at least one year of clinical experience in child and family therapy.

Treatment integrity was monitored using therapist self-report scales, previously developed and validated for use in controlled trials using multiple therapists (e.g., Barrett, Dadds, & Rapee, 1996). These scales assessed adherence to each session plan, knowledge of session material, interpersonal effectiveness, and participant engagement and comprehension. Ratings were monitored by the project coordinator in supervision sessions, with any reports of deviation from the treatment protocol or related problems addressed directly with the clinician. Using this method, one case was excluded from the sample following an excessive departure from the treatment protocol due to the parents' concurrent marital stress.

RESULTS

Inter-rater reliability for the observational assessments was high, with inter-rater data correlating $r = .71$ for observations of conduct problems, $r = .80$ for aversive parent behaviour, $r = .78$ for correct implementation of discipline strategies, and $r = .79$ for praise. Inter-rater reliability for the diagnostic interviews was also high, with a Cohen's Kappa value of 1 at post-treatment indicating perfect agreement between inter-rater diagnoses. Likewise, a strong correlation was seen between inter-rater diagnostic severity ratings at post-treatment ($r = .90$).

The internal reliability of the original APQ subscales was modest to high. The highest internal reliability was seen for the Inconsistent Discipline subscale ($\alpha = .80$), and the weakest for Corporal Punishment ($\alpha = .53$). The internal reliability of the three scales from Hinshaw et al's (2000) factor structure were generally superior, with alpha's ranging from $\alpha = .69$ (Deficient monitoring, Negative/Ineffective Discipline) to $\alpha = .74$ (Positive Involvement). All alpha coefficients are presented in Table I, as are the means for the five and three subscale structures at both pre- and post-treatment.

The convergent validity of the APQ was examined using correlations between subscale scores and observational data of parent-child interaction collected within one week of parents' completion of the APQ. As these two measures were completed concurrently at both pre-treatment and post-treatment, two sets of correlations were available for this purpose. Table II shows correlations between the original APQ subscales and rates of observed aversive parent behaviour and praise, with the same correlations presented for the three-factor APQ scales in Table III. Correlations are presented separately for each of the two observational settings (play and dinner).

Table I. Means, Coefficient Alphas, and T Tests for the Original and Three-Factor APQ Subscales at Pre- and Post-treatment

| Subscale | Alpha | Pre-treatment | | Post-treatment | | T | df | d |
|---------------------------------|-------|---------------|------|----------------|------|----------|----|------|
| | | Mean | SD | Mean | SD | | | |
| Original APQ | | | | | | | | |
| Parental involvement | .64 | 37.87 | 4.50 | 39.73 | 5.13 | 29.46** | 42 | 4.09 |
| Poor monitoring & supervision | .66 | 12.49 | 2.83 | 12.53 | 3.65 | 16.73** | 44 | 2.93 |
| Positive parenting techniques | .74 | 25.19 | 2.78 | 25.71 | 2.69 | -26.96** | 43 | 6.45 |
| Inconsistent discipline | .80 | 16.35 | 3.77 | 13.37 | 3.72 | 5.83** | 44 | 0.07 |
| Corporal punishment | .53 | 5.93 | 1.60 | 5.00 | 1.39 | 5.76** | 43 | 0.64 |
| 3-Factor structure | | | | | | | | |
| Positive involvement | .74 | 63.11 | 6.14 | 65.44 | 7.11 | -2.62* | 42 | 0.36 |
| Negative/ineffective discipline | .69 | 24.69 | 4.37 | 21.00 | 4.81 | 5.61** | 43 | 0.69 |
| Deficient monitoring | .69 | 10.09 | 2.5 | 9.8 | 3.1 | 1.89 | 43 | 0.2 |

* $p < .05$. ** $p < .01$.

Table II. Correlations Between Original APQ Subscales and Concurrent Observations of Parent-Child Interaction at Pre- and Post-Treatment

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|----------------------------------|-------|-------|-------|-------|---------|-------|-------|--------|--------|-------|-------|-------|-------|-------|-------|------|------|----|
| Pre-treatment APQ | | | | | | | | | | | | | | | | | | |
| 1. Parental involvement | 1 | | | | | | | | | | | | | | | | | |
| 2. Monitoring and supervision | -0.27 | 1 | | | | | | | | | | | | | | | | |
| 3. Positive parenting techniques | -0.15 | .36** | 1 | | | | | | | | | | | | | | | |
| 4. Inconsistency in discipline | -0.22 | .91** | .30* | 1 | | | | | | | | | | | | | | |
| 5. Corporal punishment | -0.15 | .38** | 0.18 | 0.03 | 1 | | | | | | | | | | | | | |
| Post-treatment APQ | | | | | | | | | | | | | | | | | | |
| 6. Parental involvement | .60** | -0.25 | -0.17 | -0.09 | -0.48** | 1 | | | | | | | | | | | | |
| 7. Monitoring and supervision | 0.00 | .33* | .74** | 0.20 | 0.28 | -0.18 | 1 | | | | | | | | | | | |
| 8. Positive parenting techniques | .44** | -0.28 | -0.23 | -0.15 | -0.41** | .61** | -0.27 | 1 | | | | | | | | | | |
| 9. Inconsistency in discipline | -.35* | .72** | .33* | .68** | .36* | -.30* | 0.29 | -.51** | | | | | | | | | | |
| 10. Corporal punishment | -0.09 | 0.20 | .48** | -0.03 | .74** | -.31* | .49** | -.44** | .34* | 1 | | | | | | | | |
| Pre-treatment observation | | | | | | | | | | | | | | | | | | |
| 11. Parent aversive (Dinner) | -0.01 | 0.09 | 0.02 | 0.09 | 0.13 | .30* | 0.08 | 0.13 | 0.06 | 0.14 | 1 | | | | | | | |
| 12. Use of praise (Dinner) | .32* | -0.25 | -0.09 | -0.23 | -0.08 | .33* | -0.14 | .29* | -.32* | 0.01 | 0.20 | 1 | | | | | | |
| 13. Parent aversive (Play) | -0.02 | 0.03 | 0.19 | -0.01 | .29* | -0.02 | 0.23 | 0.04 | 0.00 | .40** | .54** | 0.26 | 1 | | | | | |
| 14. Use of praise (Play) | .31* | -0.19 | -0.18 | -0.09 | -0.22 | .30* | -0.14 | 0.10 | -0.19 | -0.20 | 0.20 | .37** | 0.24 | 1 | | | | |
| Post-treatment observation | | | | | | | | | | | | | | | | | | |
| 15. Parent aversive (Dinner) | 0.22 | -0.12 | 0.07 | -0.06 | -0.11 | 0.15 | 0.17 | -0.04 | -0.01 | 0.06 | 0.27 | -0.05 | 0.04 | 0.05 | 1 | | | |
| 16. Use of praise (Dinner) | -.37* | 0.04 | 0.12 | 0.03 | 0.12 | -0.17 | 0.14 | 0.01 | 0.01 | -0.08 | 0.07 | -0.09 | 0.12 | 0.20 | -0.09 | 1 | | |
| 17. Parent aversive (Play) | 0.18 | 0.00 | 0.12 | 0.00 | 0.22 | 0.12 | 0.22 | -0.03 | 0.08 | 0.28 | .50** | 0.08 | .49** | 0.18 | .56** | 0.05 | 1 | |
| 18. Use of praise (Play) | .33* | -.34* | -0.16 | -0.19 | -.39** | .45** | -0.14 | .48** | -.41** | -0.28 | 0.03 | .33* | 0.26 | .45** | 0.03 | .30* | 0.12 | 1 |

** $p < .01$ level (2-tailed).

* $p < .05$ level (2-tailed).

Table III. Correlations Between the Three-Factor APQ Subscales and Concurrent Observations of Parent-Child Interaction at Pre- and Post-Treatment

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|------------------------------------|-------|-------|-------|--------|--------|-------|-------|-------|-------|-------|-------|------|------|----|
| Pre-treatment APQ | | | | | | | | | | | | | | |
| 1. Positive involvement | 1 | | | | | | | | | | | | | |
| 2. Negative ineffective discipline | -0.27 | 1 | | | | | | | | | | | | |
| 3. Deficient monitoring | -0.15 | .36** | 1 | | | | | | | | | | | |
| Post-treatment APQ | | | | | | | | | | | | | | |
| 4. Positive involvement | .60** | -0.28 | -0.21 | 1 | | | | | | | | | | |
| 5. Negative ineffective discipline | -0.27 | .67** | .49** | -.45** | 1 | | | | | | | | | |
| 6. Deficient monitoring | -0.07 | .29* | .77** | -0.24 | .42** | 1 | | | | | | | | |
| Pre-treatment observation | | | | | | | | | | | | | | |
| 7. Parent aversive (Dinner) | -0.01 | 0.09 | 0.02 | 0.27 | 0.16 | 0.01 | 1 | | | | | | | |
| 8. Use of praise (Dinner) | .32* | -0.25 | -0.09 | .35* | -0.26 | -0.14 | 0.20 | 1 | | | | | | |
| 9. Parent aversive (Play) | -0.02 | 0.03 | 0.19 | 0.00 | 0.15 | 0.23 | .54** | 0.26 | 1 | | | | | |
| 10. Use of praise (Play) | .31* | -0.19 | -0.18 | 0.26 | -0.21 | -0.13 | 0.20 | .37** | 0.24 | 1 | | | | |
| Post-treatment observation | | | | | | | | | | | | | | |
| 11. Parent aversive (Dinner) | 0.22 | -0.12 | 0.07 | 0.10 | 0.10 | 0.05 | 0.27 | -0.05 | 0.04 | 0.05 | 1 | | | |
| 12. Use of praise (Dinner) | -.37* | 0.04 | 0.12 | -0.11 | 0.00 | 0.16 | 0.07 | -0.09 | 0.12 | 0.20 | -0.09 | 1 | | |
| 13. Parent aversive (Play) | 0.18 | 0.00 | 0.12 | 0.08 | 0.20 | 0.18 | .50** | 0.08 | .49** | 0.18 | .56** | 0.05 | 1 | |
| 14. Use of praise (Play) | .33* | -.34* | -0.16 | .51** | -.40** | -0.16 | 0.03 | .33* | 0.26 | .45** | 0.03 | .30* | 0.12 | 1 |

** $p < .01$ level (2-tailed).

* $p < 0.05$ level (2-tailed).

Among the original APQ subscales, Parental Involvement correlated positively with observations of parents' use of praise in both the play ($r = .31, p < .05$) and dinner ($r = .32, p < .05$) settings at pre-treatment, and in the play setting at post-treatment ($r = .45, p < .01$). Positive Parenting Techniques also correlated with parents' use of praise in the post-treatment play observation ($r = .48, p < .01$). Scores on the Corporal Punishment subscale correlated positively with observations of aversive parent behaviour in the pre-treatment play observation ($r = .29, p < .05$), and scores on Inconsistent Discipline correlated negatively with observations of praise in the post-treatment play setting ($r = -.41, p < .01$).

Among the three-factor subscales, Positive Involvement correlated positively with rates of observed praise in both the play ($r = .31, p < .05$) and dinner ($r = .32, p < .05$) pre-treatment observation settings; and in the post-treatment play observation ($r = .51, p < .01$). The Negative/Ineffective Discipline subscale correlated negatively with rates of praise observed in the post-treatment play observation ($r = -.40, p < .01$). None of the three-factor subscales were found to correlate with observations of aversive parent behaviour.

To assess sensitivity of the APQ to clinical change in parenting, differences between pre- and post-treatment APQ scores were assessed using paired samples T tests. As seen in Table I, scores on each of the five original APQ subscales changed significantly from pre- to post-treatment. The effect sizes for these tests ranged from medium for Corporal Punishment (Cohen's $d = .64$), to very large for Parental Involvement (Cohen's $d = 4.09$). Among the three-factor subscales, significant differences between pre- and post-treatment scores were found for the Parental Involvement $t(42) = -2.62, p < .05$, and Negative/Ineffective Discipline $t(43) = 5.61, p < .01$, subscales, the effect sizes of which were small (Cohen's $d = 0.36$) and medium (Cohen's $d = 0.69$) respectively. Only scores on the Deficient Monitoring subscale exhibited no significant change across treatment.

To test the hypothesis that change in APQ scores across treatment would be associated with clinical child outcomes, change scores were calculated for each of the APQ subscales. For ease of interpretation, these scores were calculated to reflect the amount of positive change to the respective parenting practices (i.e., for negative subscales pre-treatment scores were subtracted from post-treatment scores, and vice versa for positive scales). Partial correlations were calculated between change scores and rates of child conduct problems observed at post-treatment (controlling for rates of conduct problems observed pre-treatment). These correlations are shown in Table IV.

Among the original five APQ subscales, rates of conduct problems observed in parent-child interaction at post-treatment correlated significantly with change in parents' scores on Parental Involvement (play observation: $r = -.46, p < .01$; dinner observation: $r = -.27, p < .05$), Positive Parenting (play observation: $r = -.41, p < .01$), and Inconsistent Discipline (play observation: $r = -.28, p < .05$). For the three-factor subscales, change scores on Positive Involvement correlated with rates of conduct problems observed in the post-treatment play

Table IV. Correlations Between Change Scores for the Original and Three-Factor APQ Subscales and Post-Treatment Observations of Conduct Problems

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|------------------------------------|--------|--------|--------|-------|--------|---------|--------|-------|-------|----|
| Original APQ | | | | | | | | | | |
| 1. Parental involvement | 1 | | | | | | | | | |
| 2. Positive parenting | 0.48** | 1 | | | | | | | | |
| 3. Inconsistent discipline | 0.30* | 0.27* | 1 | | | | | | | |
| 4. Corporal punishment | -0.12 | 0.06 | 0.15 | 1 | | | | | | |
| 5. Monitoring and supervision | 0.42** | 0.38* | 0.15 | -0.05 | 1 | | | | | |
| 3-Factor APQ | | | | | | | | | | |
| 6. Positive involvement | 0.89 | 0.83 | 0.33* | -0.05 | 0.49** | 1 | | | | |
| 7. Negative ineffective discipline | 0.30* | 0.35* | 0.91** | 0.36* | 0.31* | 0.38* | 1 | | | |
| 8. Deficient monitoring | 0.27* | 0.20 | 0.00 | 0.05 | 0.71** | 0.28* | -0.03 | 1 | | |
| Post-treatment observation | | | | | | | | | | |
| 9. Conduct problems (dinner) | -0.27* | -0.14 | -0.25 | -0.01 | -0.11 | -0.25 | -0.24 | -0.05 | 1 | |
| 10. Conduct problems (play) | -0.46* | -0.41* | -0.28* | 0.00 | -0.25 | -0.50** | -0.36* | 0.04 | 0.36* | 1 |

Note. Partial correlations were conducted controlling for pre-treatment observations of conduct problems.

** $p < .01$ level (2-tailed).

* $p < 0.05$ level (2-tailed).

observation ($r = -.50, p < .01$), as did change scores on Negative/Ineffective Discipline ($r = -.36, p < .05$).

DISCUSSION

The aims of this study were to validate self-reported parenting practices on the APQ against observations of parent-child interaction, and to compare the clinical utility of five-factor and three-factor forms of the APQ, in families of young boys being treated for conduct problems. It was predicted firstly that parents' self-reports on the APQ subscales would converge with concurrent observations of harsh/aversive parent behaviour and use of praise. Evidence of convergence was found for both the original and three-factor forms of the APQ, with all significant correlations between APQ scores and observational data occurring in the predicted directions. Four of the five original APQ subscales correlated with the observational data. Parents scoring higher on Positive Parenting Techniques and Parental Involvement were observed to use more praise across observational settings, as were those who reported higher consistency in discipline (i.e., low scores on Inconsistent Discipline). Parents reporting higher rates of corporal punishment were observed to engage in higher rates of harsh/aversive parenting. While scores on two of the three-factor subscales correlated with praise (Positive Involvement correlated positively, and Negative/Ineffective Discipline negatively), none of these subscales were related to observations of aversive parent behaviour.

It was also predicted that scores on the APQ subscales would be sensitive to clinical change in parenting across the parent training intervention. T tests conducted on pre- and post-treatment APQ scores generally supported this prediction. Compared to the three-factor APQ scales however, the original subscales were more consistently sensitive to change in parenting, and demonstrated larger effect sizes.

The final hypothesis was that change in APQ scores from pre- to post-treatment would be associated with clinical child outcomes, after controlling for baseline levels of conduct problems. Change scores on four out of the five original APQ subscales were associated with clinical child outcomes, with the children of parents recording the greatest positive change on Parental Involvement, Positive Parenting, and Inconsistent Discipline, exhibiting the lowest rates of conduct problems in post-treatment observations. Support for this prediction was also seen for the three-factor APQ subscales, with higher change scores on both the Positive Involvement and Negative/Ineffective Discipline associated with lower rates of conduct problems in post-treatment observations.

The measurement evidence presented here allows for comparison of the clinical utility of both the original five subscales of the APQ, and the alternative three-factor structure proposed by Hinshaw et al. (2000). While scores on both variations of the APQ converged with observations of parents' use of praise, only scores on one of the original APQ subscales (Corporal Punishment) were

associated with observations of harsh/aversive parent behaviour. The original APQ subscales also demonstrated greater sensitivity to change in parenting than those of the three-factor structure. This evidence suggests that scores on the five original APQ subscales may provide somewhat greater clinical utility than those from a three-factor structure.

It is noteworthy that despite the modest internal reliability of the Corporal Punishment subscale and evidence that the Corporal Punishment items may not represent a distinct factor (Elgar et al., submitted), scores on this subscale appeared to be particularly meaningful in the current treatment sample, being the only subscale scores to correlate with observations of harsh/aversive parenting. This convergence indicates that parents' self-reports of corporal discipline were not confounded by social desirability. These findings therefore support the clinical utility of this subscale.

A number of limitations should be recognised when interpreting these findings. Firstly, methodological differences associated with the inherent properties of self-report and observational methods precluded the measurement of exactly the same constructs with each. While we examined convergence between conceptually similar constructs, the similarity of these constructs was at times therefore limited. Examples include the examination of convergence between 'parental involvement' and 'use of praise,' and between 'inconsistent discipline' and 'harsh/aversive parent behaviour.' Secondly, participant factors such as the exclusion of cases with untreated ADHD comorbidity, and the use of an exclusively male sample, may limit the generalisation of the current findings to broader clinical samples. It would appear unlikely however that these factors would influence the reliability/validity of either the self-report or observational methods used in the study. It would be beneficial for future research to replicate the current study with a mixed-gender sample, and one large enough to allow for the testing of predictive effects using regression models.

The findings of this study add to existing support for the APQ as a measure of domain-specific parenting practices. Parent reports on the measure were found to converge well with concurrent observational data in the clinical sample, reflected change in parenting, and were associated with clinical child outcomes. Interestingly, the original five theoretically chosen subscales of the APQ demonstrated somewhat greater clinical utility than the alternative empirically-derived factor structure proposed by Hinshaw et al. (2000). This evidence suggests the APQ to be a valid and clinically informative tool in the treatment of childhood conduct problems.

ACKNOWLEDGMENTS

This research was in part supported by funding from the National Health and Medical Research Council. Thanks to all participating families.

REFERENCES

- American Psychiatric Association (APA). (1994). *Diagnostic and statistical manual of mental disorders: fourth edition*. Washington, DC: APA.
- Barrett, P. M., Dadds, M. R., & Rapee R. M. (1996). Family treatment of child anxiety: A controlled trial. *Journal of Consulting and Clinical Psychology, 64*, 333–342.
- Brestan, E. V., & Eyberg, S. M. (1998). Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. *Journal of Clinical Child Psychology, 27*, 180–189.
- Dadds, M. R., & McHugh, T. A. (1992). Social support and treatment outcome in behavioral Family therapy for child conduct problems. *Journal of Consulting and Clinical Psychology, 60*, 252–259.
- Dadds, M. R., Maujean, A., & Fraser, J. A. (2003). Parenting and conduct problems in children: Australian data and psychometric properties of the Alabama Parenting Questionnaire. *Australian Psychologist, 38*, 238–241.
- Elgar, F. J., Waschbusch, D. A., & Dadds, M. R. (submitted manuscript). Short form of the Alabama parenting questionnaire.
- Hawes, D. J., & Dadds, M. R. (2005). Oppositional and Conduct Problems. In J. Hudson & R. Rapee (Eds.), *Current thinking on psychopathology and the family* (pp. 73–91). New York: Elsevier.
- Hinshaw, S. P., Owens, E. B., Wells, K. C., Kraemer, H. C., Abikoff, H. B., Arnold, L. E., Conners, C. K., Elliott, G., Greenhill, L. L., Hechtman, L., Hoza, B., Jensen, P. S., March, J. S., Newcorn, J. H., Pelham, W. E., Swanson, J. M., Vitiello, B., & Wigal, T. (2000). Family process and treatment outcome in the MTA: Negative/ineffective parenting practices in relation to multimodal treatment. *Journal of Abnormal Child Psychology, 28*, 555–568.
- Holland, D., & Dadds, M. R. (1997). *The diagnostic interview schedule for children, adolescents, and parents*. Brisbane: Griffith University.
- Johnson, S., Barrett, P., Dadds, M. R., Fox, T., & Shortt, A. (1999). The diagnostic interview schedule for children, adolescents and parents: Initial reliability and validity data. *Behaviour Change, 16*, 155–164.
- McMahon, R. J., & Forehand, R. L. (2003). *Helping the noncompliant child: Family-based treatment for oppositional behaviour*. New York, NY: Guilford Press.
- Reid, J. B., Patterson, G. R., & Snyder, J. J. (2002). *Antisocial behaviour in children and adolescents: A developmental analysis and model for intervention*. Washington, DC: American Psychological Association.
- Sanders, M. R., & Dadds, M. R. (1982). The effects of planned activities and child management training: An analysis of setting generality. *Behavior Therapy, 13*, 452–461.
- Sanders, M. R., Markie-Dadds, C., Tully, L. A., & Bor, W. (2000). The Triple P – Positive Parenting Program: A Comparison of enhanced, standard and self directed behavioral family intervention for parents of children with early onset conduct problems. *Journal of Consulting and Clinical Psychology, 64*, 624–640.
- Sanders, M. R., Waugh, L., Tully, L., & Hynes, K. (1996). *The Revised Family Observation Schedule* (3rd ed.). Brisbane, Queensland, Australia: Parenting and Family Support Centre.
- Shelton, K. K., Frick, P. J., & Wootton, J. M. (1996). Assessment of parenting practices in families of elementary school-age children. *Journal of Clinical Child Psychology, 25*, 317–329.
- Snyder, J., & Stoolmiller, M. (2002). Reinforcement and coercive mechanisms in the development of antisocial behavior. The family. In J. Reid, G. Patterson, & J. Snyder (Eds.), *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention* (pp. 65–100). Washington, DC: American Psychological Association.
- Webster-Stratton, C., & Hammond, M. (1997). Treatment children with early onset conduct problems: A comparison of child and parent-training interventions. *Journal of Consulting and Clinical Psychology, 65*, 93–109.