Antecedent influences on challenging behaviour: a preliminary assessment of the reliability, generalisability and validity of the Ecological Interview

Nick Barratt[1], Peter McGill[2] and Carl Hughes[3]

[1] Dimensions

[2] Tizard Centre, University of Kent, Canterbury

[3] University of Wales, Bangor

Abstract

Background: People with intellectual disabilities who display challenging behaviour are often exposed to a range of negative outcomes, including social and material deprivation, abusive practices and disruption to family life. Several studies have linked specific antecedent events to the occurrence of challenging behaviour.

Method and materials: The present study describes the further development of an indirect method of identifying antecedent influences on challenging behaviour by assessing the reliability, generalisability and validity of the Ecological Interview (EI). Twenty care workers providing direct support to individuals with intellectual disabilities and challenging behaviour were interviewed using the EI.

Results: The EI has good test–retest reliability and generalisability, and moderate validity.

Conclusions: The results of the study are discussed in terms of their implications for clinical practice and future research, and recommendations are made regarding alternative methods of assessing validity.

Keywords: Antecedents, motivating operations, establishing operations, challenging behaviour, functional behaviour assessment.

Introduction

Approximately 10–15 per cent of people with intellectual disabilities display behaviours that could be described as challenging (Clare & McGill, 2000). Engaging in challenging behaviour is associated with a range of negative outcomes, including injury to the person and others, excessive use of psychoactive medication, abusive practices, and social and material deprivation (Durand, 1990; Emerson et al, 1999; Robertson et al, 2005). Furthermore, behaviours such as self-injury, aggression and tantrums are often cited as barriers to community inclusion for people with intellectual disabilities, and can interfere with family life, employment and education (Durand, 1990; McAtee, Carr & Schulte, 2004). Over three decades of applied behaviour analytic research have generated a substantial literature on challenging behaviour. Much is now known about the environmental determinants of behaviour, and how they can be manipulated to produce desirable behaviour change. Historically, behaviour analysts emphasized the three-term contingency, and how consequences affect behaviour and produce stimulus discrimination and control (Cooper, Heron & Heward, 2007; Luiselli, 2006; Smith & Iwata, 1997). This changed following the publication of Michael's seminal paper on establishing operations (Michael, 1982) which, together with subsequent elaborations (e.g. McGill, 1999; Michael, 1993), facilitated a conceptual alignment between the basic principles of behaviour and antecedents other than discriminative stimuli (Cooper et al, 2007).

In this paper the term *antecedent,* which is used interchangeably with *antecedent variables, antecedent events* and *antecedent influences*, refers to discriminative

stimuli (S^Ds) and motivating operations (MOs). The S^D is a stimulus in the presence of which a response has been reinforced and in the absence of which it has not. As a result, the S^D increases the frequency of the response (Cooper et al, 2007). The MO is an event or stimulus condition which momentarily changes (a) the reinforcing or punishing effectiveness of environmental consequences and (b) the frequency of responses associated with such consequences. The term *motivating operation* has come to replace that of *establishing operation* (Laraway et al, 2003).

Three categories of antecedents have been identified in the literature on challenging behaviour: social, biological and physical (Carr et al, 2003). Social antecedents include communication difficulties (Carr et al, 1996; McGill et al, 2003), criticism (Mayer, 1995), a failure to match task requirements to level of ability (Mayer, 1995), non-preferred activities (Koegel et al, 1987), lack of choice (Dunlap et al, 1991), denial of a planned activity (Horner at al, 1996), the presence of particular people (Touchette et al, 1985), and a lack of predictability (Horner et al, 1996). Biological antecedents include illness (McAtee et al, 2004), pain (Horner et al, 1996), and fatigue (O'Reilly, 1995). Physical antecedents include temperature (Carr et al, 2003), noise (O'Reilly Lacey and Lancioni, 2000), and the absence of stimulating objects (Horner, 1980).

Information on antecedent influences can be gathered through experimental functional analysis, direct observation and the use of interviews and checklists (McAtee et al, 2004). O'Reilly (1997) used experimental methods to investigate a hypothesised biological antecedent – otitis media – for self-injurious behaviour in a young girl. Six experimental conditions were presented, three of which took place when the girl was diagnosed with otitis media, and three when otitis media was absent. The study demonstrated a conditional functional relation between otitis media and self-injury. Assessing antecedents through experimental functional analysis is the only way to demonstrate experimental control, the most convincing evidence for causality (Cooper et al, 2007; McAtee et al, 2004; O'Neill et al, 1997). However, the design, execution and analysis involved takes time and effort, and requires a high level of professional expertise (Miltenberger, 2004; but see Iwata & Dozier, 2008, for methods of experimental functional analysis that are more practical in clinical settings). Furthermore, there are times when potential antecedents cannot be manipulated for ethical reasons (McAtee et al, 2004)

Repp et al (1991) conducted direct observations in order to investigate antecedent events associated with both challenging and adaptive behaviour. Observers used microcomputers to collect real-time, continuous data on activities thought to be acting as antecedents. Results indicated that different activities functioned as antecedents for stereotypy and appropriate task-related behaviours. Direct observation does not involve the manipulation of environmental stimuli associated with challenging behaviour, and therefore avoids the kind of ethical issues that can arise when conducting experimental functional analyses (McAtee et al, 2004). However, direct observations are time-consuming, yield only correlation data, and require a degree of proficiency on the part of the observer (McAtee et al, 2004; Miltenberger, 2004).

Information on antecedent influences can also be gathered using interviews or checklists (McAtee et al, 2004). The *Functional Assessment Interview* (O'Neill et al, 1997) includes a section for identifying possible antecedents. Respondents are asked to discuss environmental variables such as sleep patterns, choice-making opportunities and staffing ratios. The *Setting Event Checklist* (Gardner et al, 1986) is a tool for identifying potential antecedents that occurred in the recent past; for example, the previous evening. Indirect methods have the disadvantage of relying on respondents' memories of events and, like direct observation, generate only correlational data (Miltenberger, 2004). However, interviews and checklists are less time-consuming than direct and experimental methods, and require less proficiency on the part of the practitioner (McAtee et al, 2004). Another advantage is that interviews and checklists provide clinicians with a relatively quick and simple method of identifying, from a range of possible environmental variables, antecedents that may be relevant to the challenging behaviour of a particular individual (McAtee et al, 2004). The resulting data could either provide the starting point for more targeted and rigorous methods of assessment (e.g. experimental functional analysis) or lead directly to intervention (e.g. modifying an individual's environment to eliminate identified antecedents) (McAtee et al, 2004; McGill et al, 2003).

Two indirect methods which facilitate a comprehensive assessment of potential antecedent events are the *Contextual Assessment Inventory* (CAI) (McAtee et al, 2004) and the *Ecological Interview* (EI) (McGill et al, 2003, McGill et al, 2012). The purpose of both measures

is to identify antecedent variables which may be associated with challenging behaviour. The CAI contains 80 items grouped into four categories: Social/cultural, Nature of task or activity, Physical and Biological. Items are scored on a five-point Likert-type scale from 1 (*never*) to 5 (*always*) according to the likelihood of their association with incidents of challenging behaviour. McAtee et al (2004) administered the CAI to 40 care staff providing support to 20 individuals with intellectual disabilities. Results indicated that the CAI has good test-retest reliability (Pearson product-moment yielded a correlation coefficient of 0.74) and modest inter-rater reliability (the average intraclass correlation coefficient was 0.28). In a later study, Carr, Ladd, and Schulte (2008) assessed the convergent and predictive validity of the CAI by comparing CAI ratings with, respectively, behaviour log entries and the results of direct observation. Using Cohen's *d* (Cohen, 1988), the effect sizes for convergent and predictive validity were, respectively, 0.76 (large) and 0.85 (large).

Items in the EI are grouped into ten categories of antecedent: Physical setting (12 items), Time of day (9 items), Day of week (7 items), Holidays/Seasons (6 items), Weather conditions (6 items), Activity (13 items), Other clients (number of items is equivalent to number of clients in setting), Staff/Carers (number of items is equivalent to number of staff in setting), Social context (9 items), and Personal context (14 items). Respondents are asked whether, in the presence of a potential antecedent, challenging behaviour is more or less likely or whether the antecedent makes no difference. McGill et al. (2003) administered the EI to 66 care staff providing support to 22 individuals with intellectual disabilities. In order to evaluate test-retest reliability, 20 staff underwent a second interview and overall agreement was 79 per cent (range across categories: 66% –86 per cent).

Results of the study by McGill et al (2003) suggest that a wide range of antecedent variables are associated with challenging behaviour, including difficult tasks, lack of engagement, difficulties with expressive and receptive communication, disrupted sleep, and negative emotional states, such as anxiety. However, the EI was administered to a single small sample, which raises questions about the extent to which results would generalise to other people and situations. Furthermore, test-retest reliability was assessed using the Smaller/Larger index (Suen & Ary, 1989), a method that did not allow for comparisons at the level of individual respondents' ratings. Of note was that no attempt was made to assess validity.

The purpose of the present study is to evaluate the utility of the EI. The study has three specific aims: (1) assess the test–retest reliability of the EI using a method that compares individual respondents' ratings over time; (2) investigate generalisability by comparing data from EIs administered in the present study with data from EIs administered in the study by McGill et al (2003); and (3) assess convergent validity by comparing ratings made using the EI with those made using the CAI.

Method

Participants

Service managers working for a care services provider were asked to identify 20 individuals living in residential services who met the following criteria: (1) they had an intellectual disability; (2) they displayed challenging behaviour; and (3) they were not undergoing a major life change, such as a change of residence or a new treatment. The resulting list of 20 participants consisted of 16 men and four women, ranging in age from 21–69 years (mean: 41.7). Nine participants were diagnosed with a moderate intellectual disability, seven with a severe intellectual disability, and four with a profound intellectual disability. Participants' details are shown in *Table 1*.

Managers were asked to nominate staff members to act as respondents for each of the 20 individuals with intellectual disabilities taking part in the study. Nomination criteria consisted of the requirements that staff members were currently providing direct support to one of the individuals in question, and had been supporting that individual for at least three months. The latter criterion reflected the finding that the length of time a staff member supported a particular individual was positively correlated with the number of antecedents they were able to identify for that individual's challenging behaviour (McAtee et al, 2004). Ten men and 10 women acted as respondents, ranging in age from 21–64 (mean age: 34.5). On average, they had supported the people whose behaviour they were rating for 32.35 months (range: 5–120). Respondents were provided with information on the study and asked to agree in writing that they were willing to participate. The researcher had no direct contact with the individuals with learning disabilities. Furthermore, respondents were asked to use pseudonyms when referring to the individuals they supported, thus ensuring that the 20 individuals with learning disabilities participating in the study remained unknown to the researcher. In addition, when completing the EI and CAI, respondents provided answers based

Table 1: Characteristics of participants with intellectual disabilities

Participant	Gender	Level of ID*	Additional Diagnoses	Challenging behaviour
1	Female	Moderate	Physical aggression	
2	Female	Profound	ASD**	Screaming
3	Male	Severe	ASD	Physical aggression
4	Male	Severe	Epilepsy	Physical aggression
5	Female	Severe	ASD	Physical aggression
6	Male	Severe	Hydrocephalus	Screaming
7	Male	Moderate	Epilepsy	Physical aggression
8	Male	Severe	Self-injury	
9	Male	Moderate	Hearing impairment	Screaming
10	Male	Moderate	Screaming	
11	Male	Severe	Cornelia de Lange	Self-injury
12	Male	Profound	Self-injury	
13	Male	Profound	Bi-polar disorder	Physical aggression
14	Male	Moderate	Physical aggression	
15	Male	Moderate	ASD	Screaming
16	Female	Profound	Incontinence	
17	Male	Moderate	ASD	Screaming
18	Male	Moderate	Physical aggression	
19	Male	Moderate	Physical aggression	
20	Male	Severe	Physical aggression	

* Where possible, information on level of intellectual disability was obtained from formal diagnoses recorded in participants' case files. However, in some instances, the researcher relied on respondents' descriptions.
**Autism Spectrum Disorder.

on their perceptions of individuals, not by interacting with, or observing, them. Therefore, neither the researchers nor the ethics committee deemed it necessary to obtain consent from the individuals with learning disabilities. Ethical approval was obtained prior to the commencement of the study.

Setting
Data were gathered in 14 community homes for people with intellectual disabilities. All residential services were located in ordinary community settings – for example, a semi-detached house in a residential street – and staffed on a 24-hours-a-day basis.

Measures
Respondents were interviewed using a revised version of the EI (McGill et al, 2003) and the original CAI (McAtee et al, 2004). Items were excluded from the revised version if they had been rated by more than 50 per cent of respondents in the study by McGill et al (2003) as making no difference to the occurrence of challenging behaviour. Three categories of items met this criterion: Days of the week, Holidays/Seasons and Weather conditions. However, only Days of the week and Weather conditions were omitted, as two items in Holidays/ Seasons ('Holidays' and 'At Christmas Time') did not meet the criterion for exclusion. The revised version of

the EI used in the present study, then, excluded Days of the week and Weather conditions.

Design
At Time 1, respondents were interviewed using the EI with regard to an individual they supported. On the same day, in order to facilitate the assessment of convergent validity, respondents rated the same individual using the CAI. The presentation of the EI and CAI was counterbalanced across participants to avoid order effects. For the purpose of assessing test-retest reliability, respondents were interviewed a second time using only the EI (Time 2). Time 1 and Time 2 were separated by a period of approximately 18 days (range: 13–28). One respondent was unable to complete the EI at Time 2 (therefore n was 19 for the assessment of test–retest reliability). For the purpose of assessing generalisability, summary data from EIs administered at Time 1 in the present study were compared with summary data from EIs administered in the study by McGill et al (2003).

Scoring reliability
During six interviews, a second researcher independently scored respondents' answers. Percentage agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements, and

then multiplying by 100. Mean agreement was calculated as 100 per cent.

Results

Test–retest reliability was assessed using percentage agreement and weighted kappa (Cohen, 1968). Weighted kappa corrects for chance agreement, and could therefore be considered a more accurate method of assessing test-retest reliability than percentage agreement. Weighted kappa is preferable to standard kappa as it takes into account the degree of disagreement between ratings (Bakeman, Deckner & Quera, 2005). For example, with regard to an item on the EI, a Time 1 to Time 2 disagreement of 'less likely' to 'makes no difference' would be considered less important than a Time 1 to Time 2 disagreement of 'less likely' to 'more likely'. Due to a lack of variability between ratings, weighted kappa would have been unable to compare all 19 respondents' ratings of a particular item at Time 1 and Time 2. It was therefore decided to calculate agreement at the level of pairs of questionnaires – that is, questionnaires completed by each respondent at Time 1 and Time 2. For example, a respondent's rating for an item at Time 1 was compared with his or her rating for the same item at Time 2, and so on for every item in the EI. Overall test-retest reliability for the EI was obtained by calculating the sum of percentage agreements and weighted kappas for pairs of questionnaires, then dividing by 19 (the number of respondents who completed EIs at Time 1 and Time 2). The recommendations made by Fleiss (1981) were followed in order to determine whether weighted kappa values were excellent (0.75), good (between 0.60 and 0.75) or fair (between 0.40 and 0.60). Overall percentage agreement and weighted kappa were, respectively, 75.37 per cent (range: 61.29 per cent to 96.61 per cent) and 0.64 (range: 0.43–0.93). *Table 2* shows percentage agreement and weighted kappa values for EI sections,

Table 2: Test–retest reliability of EI sections

EI Sections	Percentage agreement	Weighted kappa
Physical setting	76.06	0.67
Time of day	64.91	0.52
Holidays/seasons	84.07	0.68
Activity	68.07	0.55
Other clients	77.50	0.71
Staff/carers	93.33	0.90
Social context – other	77.98	0.71
Personal context	83.43	0.68

which ranged from, respectively, 64.91 per cent to 93.33 per cent and 0.52–0.9.

Convergent validity was assessed by comparing ratings made using the EI with those made using the CAI. The first step was to ascertain which items in the EI had corresponding items in the CAI. Fourteen comparable items were identified, seven of which refer to biological antecedents, and seven to social antecedents. Next, Spearman's rho was used to calculate correlation coefficients. Of the 14 correlations, six were significant at the 0.01 level (2-tailed), and 3 at the 0.05 level (2-tailed). *Tables 3 and 4* show, respectively, comparable social and biological antecedents, their correlation coefficients (*rho*), and the number of respondents for each item. It should be noted that there were only four respondents for 'Around the time of the menstrual period', and only 1 for 'When short of cigarettes'.

The mean validity coefficients for social and biological antecedents were, respectively, 0.42 (range: 0.15–0.71) and 0.63 (range: 0.42–1.00). The overall validity coefficient for the 14 comparable EI items, obtained by calculating the sum of the correlation coefficients and dividing by 14, was 0.53 (range: 0.15–1.00).

Validity was also assessed by ascertaining which of the 14 EI items, and their comparable CAI counterparts, were rated by more than 50 per cent of respondents as associated with a high probability of challenging behaviour (indicated by a rating of *More likely* in the EI and *Always* – a 4 or 5 – in the CAI). Items meeting this criterion, and the percentage of respondents endorsing them, are shown in *Table 5* (ranked from highest to lowest according to the percentage score of EI items). Of the seven EI items in *Table 5*, only two yielded significant correlations when compared with corresponding CAI items using Spearman's rho, namely: 'When having difficulty understanding others' and 'Doing nothing' (see *Tables 3 and 4*). Of the remaining five non-significant EI items, four had CAI counterparts which were also rated by more than 50 per cent of respondents as likely to be associated with challenging behaviour, namely: 'When having difficulty making oneself understood', 'In a crowded room', 'Doing tasks which they find difficult' and 'Waiting for an activity'.

Generalisability was assessed by comparing the percentage of respondents in the current and previous (McGill et al, 2003) studies who rated each item as more likely to be associated with challenging behaviour.

Table 3: Comparable social antecedents and their correlation coefficients

EI (n)	CAI (n)	rho
Car/minibus (20)	Being in van or car (20)	0.32
Holidays (19)	Anxiety over holiday activities (19)	0.71**
Doing nothing (20)	Boredom/inactivity (20)	0.62**
Waiting for an activity (20)	Waiting (in line, waiting rooms) (20)	0.15
Just after an activity ends (20)	Favourite activity ends (20)	0.51*
Doing tasks which they find difficult (20)	Difficult tasks (20)	0.44
In a crowded room (20)	Too many people around the person (20)	0.22

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Table 4: Comparable biological antecedents and their correlation coefficients

EI (n)	CAI (n)	rho
When medication has been changed (13)	Changes in medication (16)	0.52*
When ill (18)	Acute illness/pain (20)	0.45*
When on a diet (9)	Being on a diet (11)	0.74**
Around the time of menstrual period (4)	Menstrual discomfort (4)	0.72**
When short of cigarettes (1)	Long periods between cigarettes (1)	1.00**
When having difficulty making oneself understood (20)	Ongoing difficulty communicating wants or needs (20)	0.42
When having difficulty understanding others (20)	Difficulty understanding staff directions (20)	0.57**

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Spearman's rho generated a coefficient of 0.79, significant at the 0.01 level (2-tailed). Generalisability was also assessed by comparing items rated by at least 50 per cent of respondents in both studies as more or less likely to be associated with challenging behaviour. *Tables 6 and 7* show, respectively, items in the previous and current studies which met both this criterion and the requirement that a minimum number of respondents made the rating (*n* was at least 20 in the previous study and at least 15 in the current study, reflecting differences in sample size). Of the 14 items in the previous study that qualified, 13 were also rated by at least 50 per cent of respondents in the current study as more or less likely to be associated with challenging behaviour ('One-to-one with staff/caregiver' was not rated by more than 50 per cent of respondents in the present study as associated

with a lower likelihood of challenging behaviour). Stated differently, 92.86 per cent of items rated as high/low probability antecedents in the previous study received similar ratings from respondents in the current study. However, 15 items (three low probability antecedents, 12 high probability antecedents) which met the 50 per cent inclusion criteria in the current study did not do so in the study by McGill et al (2003).

Discussion

Test–retest reliability

In the study by McGill et al (2003), test-retest reliability was assessed using a method which, as stated previously, does not provide information on individual agreement. In order to assess test-retest reliability at the level of individual agreement, the present study employed

Table 5: Comparable antecedents rated using the EI and CAI by at least 50 per cent of respondents as associated with an increased likelihood of challenging behaviour

EI ratings of 'More likely' (% endorsing)	CAI ratings of 'Always' (% endorsing)
When having difficulty making oneself understood (90%)	Ongoing difficulty communicating wants or needs (65%)
When having difficulty understanding others (80%)	Difficulty understanding staff directions (50%)
In a crowded room (80%)	Too many people around the person (50%)
Doing tasks which they find difficult (75%)	Difficult tasks (55%)
Waiting for an activity (60%)	Waiting (65%)
Car/minibus (55%)	
Doing nothing (50%)	

Table 6: Antecedents rated using the EI in the previous study (McGill et al, 2003) by at least 50 per cent of respondents as more or less likely to be associated with challenging behaviour

Less likely (% endorsing)	More likely (% endorsing)
During the night (67%)	When tense or anxious (89%)
One-on-one with staff/caregiver (54%)	When in a bad mood (82%)
	In a crowded room (69%)
	When depressed or sad (66%)
	Doing tasks which they find difficult (64%)
	When there is a lot of noise (58%)
	When having difficulty making oneself understood (58%)
	Doing nothing (54%)
	Waiting for an activity (54%)
	Breakfast time (53%)
	When sleep has been disturbed (52%)
	When having difficulty understanding others (50%)

Table 7: Antecedents rated using the EI in the present study by at least 50 per cent of respondents as more or less likely to be associated with challenging behaviour

Less likely (% endorsing)	More likely (% endorsing)
Alone (70%)	When tense or anxious (100%)
Doing tasks they find easy (60%)	When in a bad mood (95%)
Last thing (55%)	
During the night (55%)	When having difficulty making oneself understood (90%)
	When having difficulty understanding others (80%)
	In a crowded room (80%)
	Doing tasks which they find difficult (75%)
	When depressed or sad (73%)
	When there is a lot of noise (70%)
	Day centre (69%)
	Kitchen (65%)
	In a large group of staff, clients (more than 6) (65%)
	Dining room (63%)
	Other public places (61%)
	Living room (60%)
	Waiting for an activity (60%)
	When sleep has been disturbed (59%)
	Car/minibus (55%)
	During the afternoon (55%)
	Doing household chores (53%)
	Shops (50%)
	Breakfast time (50%)
	Evening meal (50%)
	Doing nothing (50%)
	When ill (50%)

percentage agreement and weighted kappa, which yielded overall test-retest reliability values of, respectively, 75.37 per cent and 0.64. A percentage agreement of 75.37 per cent corresponds with reliabilities reported in other studies on the environmental determinants of behaviour in people with intellectual disabilities (McAtee et al, 2004). However, in view of the fact that percentage agreement can overestimate concordance, a weighted kappa value of 0.64 (good) is perhaps the more significant result.

With regard to specific sections of the EI, some variations in test-retest reliability could be explained in terms of the salience of particular antecedents. For example, consider *Staff/Carers,* which had the highest test–retest reliability, and *Time of day*, which had the lowest. It could be argued that it would be easier for a respondent to associate the tangible presence of a staff member with a higher or lower likelihood of challenging behaviour than it would the time of day. Furthermore, it is likely that respondents would be able to draw upon examples of

different situations in which challenging behaviour either did or did not occur in the presence of a particular member of staff. However, if an increase or decrease in challenging behaviour was often observed first thing in the morning, it is possible that a respondent would be at least as likely to attribute this to other, more obvious, antecedents (e.g. whatever activity the individual happened to be engaged in at the time).

The ability of respondents to recognise antecedent influences could also explain why *Personal context* had high test–retest reliability. The putative antecedents listed in this section are best described as biological factors; for example, 'Around the time of the menstrual period'. It is possible that such antecedents become more salient by virtue of their pervasiveness and persistence. For example, consider 'When having difficulty making oneself understood', a biological antecedent listed in *Personal Context.* A difficulty with expressive communication is likely to increase the probability of challenging behaviour in a number of different situations. It is also an antecedent which, while amenable to attenuation, is present throughout the life of the individual. Contrast this with 'Doing tasks which they find difficult', a social antecedent located in a section (*Activity)* which had low test-retest reliability. A difficult task only acts as an antecedent when the individual is actually performing the task. Furthermore, the antecedent of a difficult task reflects the interaction of a range of easily discernable environmental stimuli, such as the technical skills of the person providing support and the level of noise in the immediate environment. It is therefore possible that the effectiveness of difficult tasks as an antecedent will vary considerably over time as a function of the above-mentioned environmental variables. Finally, with regard to test-retest reliability in general, it is important to consider that disagreements across time are as likely to reflect the natural variability of behaviour as they are errors in measurement (McGill et al, 2001).

Convergent validity
The overall validity coefficient for the fourteen EI items that had comparable items on the CAI was 0.53. The following factors should be considered when interpreting this result. First, there were differences in the wording of comparable items. Of the fourteen items in the EI which had CAI counterparts, only two had highly similar wording: 'When on a diet' and 'Being on a diet'; and 'When medication has been changed' and 'Changes in medication'. The other items were worded in a way that could be described as reasonably similar but open to different interpretations; for example, 'Waiting for an activity' and 'Waiting (in line, waiting rooms)', where the former appears to be referring to activities in general and the latter to specific situations. 'Waiting for an activity' is an example of an EI item which did not correlate significantly with its CAI counterpart. However, it is also possible that items that *did* correlate significantly could have yielded even higher coefficients were it not for differences in wording; for example, 'Holidays' and 'Anxiety over holiday activities' where the former could refer to being on holiday and the latter to planning one.

Second, respondents rated items on the EI and CAI using, respectively, a relative and an absolute scale. The way in which this could lead to discrepancies between ratings for comparable items will be illustrated using the example of a participant who is more likely to engage in challenging behaviour when on a diet. When asked to rate the likelihood of the participant's challenging behaviour 'When on a diet' (EI item), the respondent answers, 'More likely'. When asked the corresponding question in the CAI – 'Being on a diet' – the respondent gives the item a score of 2 out of a possible 5. This discrepancy in ratings does not necessarily reflect inconsistent responding, for a score of 2 still indicates an increased likelihood of challenging behaviour in the presence of the antecedent (as compared to a rating of 1, which equates to 'Never'). However, a score of 2 on the CAI corresponds more closely to a rating of 'Less likely' on the EI than it does to a rating of 'More likely', thus leading to the generation of a low validity coefficient.

Third, of the five EI items used to assess validity which did not yield significant correlations, four, along with their CAI counterparts, were nevertheless rated by more than 50 per cent of participants as high probability antecedents.

Finally, that there were only 14 items in the EI with comparable counterparts in the CAI is of interest in itself, especially given that the EI and CAI are both tools for assessing antecedents. The extent of the differences between measures could indicate that, relative to our knowledge of consequences associated with challenging behaviour, we still have much to learn about antecedents.

Generalisability
A correlation coefficient of 0.79 suggests that the EI generalises across samples, settings, and time. According to the qualitative assessment of

generalisability, significant high-probability antecedents appear to be: a tense, anxious or bad mood; difficulty making oneself understood; difficulty understanding others; a crowded room; doing difficult tasks; feeling depressed or sad; a noisy environment; waiting for an activity; disturbed sleep; breakfast time; and, doing nothing. These results are consistent with the behaviour analytic literature on antecedents associated with challenging behaviour in people with learning disabilities. Many of the antecedents identified in the present study as associated with a high likelihood of challenging behaviour have been shown by other researchers to be correlated with, or functionally-related to, challenging behaviour; for example, difficult tasks (e.g. Koegel et al 1987), fatigue (e.g. O'Reilly, 1995), and low levels of stimulation (e.g. Horner, 1980). With regard to low probability antecedents, there was agreement on only one item – 'During the night' – which is probably due to the EI's focus on antecedents associated with a high likelihood of challenging behaviour (McGill et al, 2003).

Limitations of the present study and implications for future research

In common with the study by McGill et al (2003), no attempt was made to assess interrater reliability. Obtaining such data would undoubtedly facilitate a more informed assessment of the utility of the EI as a clinical instrument. The present study attempted to validate the EI using another indirect method of assessing antecedents, the CAI. Validity could be more rigorously assessed using direct or experimental methods (McAtee et al, 2004; McGill et al, 2003). For example, researchers could use the EI to generate hypotheses regarding high probability antecedents and then attempt to either obtain correlational data through direct observation or demonstrate a functional relationship through experimental functional analysis. Such methods would address a threat to validity inherent in any study that relies on indirect methods of assessing behaviour; namely, the possibility that respondents' ratings are based more on their beliefs about a given antecedent than observed instances of that antecedent influencing behaviour. It is also important to note that the assessment of convergent validity involved only 14 items in the EI, which equates to approximately 20 per cent of the total number of items.

Another obvious limitation of the present study is that it was based on data obtained from twenty participants. This will hopefully be addressed by future research on the EI involving larger sample sizes. With regard to the issue of generalisability, it is likely that a degree of homogeneity existed between samples in the present and previous studies, given that both were comprised of care workers supporting people with intellectual disabilities in residential settings. It would be interesting to see what results would be obtained if research were conducted in different settings or involved people with different characteristics, for example, in an assessment and treatment unit or with a sample comprised solely of individuals with autism.

Clinical practice

The EI has a number of clinical applications. First, it could be used as an initial screening tool which facilitates a more rigorous assessment of antecedents, such as experimental functional analysis (McAtee et al, 2004). Second, it could become one of the many instruments used by clinicians to conduct descriptive functional behaviour assessments. Third, data generated by the EI could lead directly to the implementation of antecedent interventions such as non-contingent reinforcement, 'neutralizing routines' and high-probability request sequences (Cooper et al, 2007; Horner, Day & Day, 1997, p. 601). Fourth, the EI could be used to inform policy and practice in the field of intellectual disabilities at the level of service design (McGill et al, 2003). For example, the results of the present study suggest that service providers who strive to decrease the likelihood of challenging behaviour should promote relaxed physiological states, take steps to elevate mood, avoid crowded and noisy environments, avoid difficult tasks (or teach people coping strategies), address communication difficulties (e.g. using systems of alternative and augmentative communication), avoid periods during which people are required to wait for things to happen, pay attention to eating arrangements, take steps to ensure people sleep well, and organise support so that people's lives are characterised by high levels of engagement (McGill et al, 2003).

Concluding comments

The EI was administered to twenty carers providing support to people with intellectual disabilities. Results indicate that the EI has good test-retest reliability and generalisability, and, with regard to approximately 20 per cent of its items, moderate validity. The results of the present study have a number of implications for clinical practice and research. Data from the EI could inform service provision, lead directly to the design of antecedent interventions, or compliment other data gathering methods during functional behaviour

assessment. The present study could also be viewed as representing another stage in the development of an instrument designed to assess antecedent influences on challenging behaviour in people with intellectual disabilities.

References

Bakeman, R, Deckner, DF and Quera, V (2005), 'Analysis of behavioral streams' in DM Teti (Ed.), *Handbook of research methods in developmental science* (pp. 394–420), Oxford, UK: Blackwell Publishers.

Carr EG, Ladd MV and Schulte CF (2008), 'Validation of the Contextual Assessment Inventory for problem behavior', *Journal of Positive Behavior Interventions*, issue 10, 91–104.

Carr, EG, Reeve, CE and Magito-McLaughlin, D (1996), 'Contextual influences on problem behavior in people with developmental disabilities' in LK Koegel, RL Koegel and G Dunlap (Eds.), *Positive behavioral support: Including people with difficult behavior in the community* (pp. 403–423), Baltimore: Paul H. Brookes.

Carr, EG, Smith, CE, Giacin, TA, Whelan, BM and Pancari, J (2003), 'Menstrual discomfort as a biological setting event for severe problem behavior: Assessment and intervention', *American Journal on Mental Retardation,* issue 108, pp. 117–133.

Clare, I and McGill, P (2000), 'Challenging behavior and its management' in M Davies (Ed.), *The Blackwell Encyclopaedia of Social Work* (pp. 45–47), Oxford: Blackwell Publishers.

Cohen, J (1968), 'Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit', *Psychological Bulletin,* issue 70, pp. 213–220.

Cohen, J (1988), *Statistical power analysis in the behavioral sciences* (2nd ed.), Hillsdale, NJ: Lawrence Erlbaum.

Cooper, J0, Heron, TE and Heward, WL (2007), *Applied behavior analysis (2nd ed),* Upper Saddle River, NJ: Pearson Education Inc.

Dunlap, G, Kern-Dunlap, L, Clarke, S and Robbins, FR (1991), 'Functional assessment, curricular revision, and severe behavior problems', *Journal of Applied Behavior Analysis,* issue 24, pp. 387–397.

Durand, MV (1990) *Severe behavior problems: A functional communication training approach,* New York, NY: Guilford Press.

Emerson, E, Felce, D, McGill, P and Mansell, J (1999), 'Introduction' in E Emerson, P McGill and J Mansell (Eds.), *Severe learning disabilities and challenging behaviors: designing high quality services* (pp. 3–16)*,* Cheltenham, UK: Stanley Thornes.

Fleiss, JL (1981), *Statistical methods for rates and proportions*, New York: Wiley.

Gardner, WI, Cole, CL, Davidson, DP and Karan, OC (1986), 'Reducing aggression in individuals with developmental disabilities: An expanded stimulus control, assessment, and intervention model', *Education and Training in Mental Retardation*, issue 21, pp. 3–12.

Horner, RD (1980), 'The effects of an environmental "enrichment" program on the behavior of institutionalized profoundly retarded children', *Journal of Applied Behavior Analysis*, Issue 13, pp. 473–491.

Horner, RH, Day, HM and Day, JR (1997), 'Using neutralizing routines to reduce problem behavior', *Journal of Applied Behavior Analysis,* issue 30, pp. 601–614.

Horner, RH, Vaughn, BJ, Day, HM and Ard, WR, Jr (1996), 'The relationship between setting events and problem behavior' in LK Koegel, RL Koegel and G Dunlap (Eds.), *Positive behavioral support: Including people with difficult behavior in the community* (pp. 381–402, Baltimore: Brooks.

Iwata, BA and Dozier, CL (2008), 'Clinical application of functional analysis methodology', *Behavior Analysis in Practice*, issue 1, pp. 3–9.

Koegel, RL, Dyer, K and Bell, LK (1987), 'The influence of child-preferred activities on autistic children's social behavior', *Journal of Applied Behavior Analysis*, issue 20, pp. 243–252.

Laraway, S, Snycerski, S, Michael, J and Poling, A (2003), 'Motivating operations and terms to describe them: Some further refinements', *Journal of Applied Behavior Analysis*, issue 36, pp. 407–414.

Luiselli, JK (2006), *Antecedent assessment & intervention: Supporting children & adults with developmental disabilities in community settings*, Baltimore, MD: Paul H. Brookes.

Mayer, GR (1995), 'Preventing antisocial behavior in the schools', *Journal of Applied Behavior Analysis*, issue 28, pp. 67–478.

McAtee, M, Carr, EG and Schulte, C (2004), 'A contextual assessment inventory for problem behavior: Initial development', *Journal of Positive Behavior Interventions*, issue 6, pp. 148–165.

McGill, P (1999), 'Establishing operations: Implications for the assessment, treatment and prevention of problem behavior', *Journal of Applied Behavior Analysis*, vol. 32, issue 3, pp. 393–418.

McGill, P, Teer, K, Rye, L and Hughes, D (2003), 'Staff reports of setting events associated with challenging behavior', *Behavior Modification*, issue 27, pp. 265–282.

McGill, P, Hughes, D, Teer, K and Rye, L (2012) 'Investigating naturally occurring variability in challenging behaviour – the Ecological Interview', *International Journal of Positive Behavioural Support*, vol. 2, no. 2, pp. 20–30.

Michael, J (1982), 'Distinguishing between discriminative and motivational functions of stimuli', *Journal of the Experimental Analysis of Behavior*, issue 37, pp. 149–155.

Michael, J (1993), 'Establishing operations', *The Behavior Analyst*, issue 16, pp. 191–206.

Miltenberger, RG (2004), *Behavior modification: principles and procedures*, Belmont, CA: Wadsworth/ Thompson Learning.

O'Neill, RE, Horner, RH, Albin, RW, Sprague, JR, Storey, K and Newton, JS (1997), *Functional assessment and program development for problem behavior: a practical Handbook*, Pacific Grove, CA: Brooks/Cole Publishing Company.

O'Reilly, MF (1995), 'Functional analysis and treatment of escape-maintained aggression correlated with sleep deprivation', *Journal of Applied Behavior Analysis*, issue 28, pp. 225–226.

O'Reilly, MF (1997), 'Functional analysis of episodic self-injury correlated with recurrent otitis media', *Journal of Applied Behavior Analysis*, issue 30, pp. 165–167.

O'Reilly, MF, Lacey, C and Lancioni, GE (2000), 'Assessment of the influence of background noise on escape-maintained problem behavior and pain behavior in a child with Williams syndrome', *Journal of Applied Behavior Analysis*, issue 33, pp. 511–514.

Repp, AC, Singh, NN, Karsh, KG and Deitz, DED (1991), 'Ecobehavioral analysis of stereotypic and adaptive behaviors: Activities as setting events', *Journal of Mental Deficiency Research*, issue 35, pp. 413–429.

Robertson, J, Emerson, E, Pinkney, L, Ceasar, E, Felce, D, Meek, A, Carr, D, Lowe, K, Knapp, M and Hallam, A (2005), 'Treatment and management of challenging behaviors in congregate and non-congregate community-based supported accommodation', *Journal of Intellectual Disability Research*, issue 49, pp. 63–72.

Smith, RG and Iwata, BA (1997), 'Antecedent influences on behavior disorders', *Journal of Applied Behavior Analysis*, issue 30, pp. 343–375.

Suen, HK and Ary, D (1989), *Analyzing quantitative behavioral observation data*, Hillsdale, NJ: Lawrence Erlbaum.

Touchette, PE, MacDonald, RF and Langer, SN (1985), 'A scatter plot for identifying stimulus control of problem behavior', *Journal of Applied Behavior Analysis*, issue 18, pp. 343–351.