

**The Utility of the Implicit Relational Assessment Procedure (IRAP):  
Measuring Self-Esteem Profiles of Children with Special Educational Needs  
and Assessing the Attitudes of Teachers to Children with Emotional  
Behavioural Difficulties in Mainstream Education**



**Geraldine Scanlon BA**

**Thesis submitted to the National University of Ireland, Maynooth in fulfilment of the  
requirements for a PhD. in Psychology**

**Supervised by Dr. Yvonne Barnes-Holmes**

**Department of Psychology**

**July 2007**

*To Jonathan, Meggan and Holly*

*Thank you for all of your support and encouragement*

*With Love*

## CONTENTS

	<b>Page</b>
<b>Chapter 1</b> General Introduction	1
<b>Chapter 2</b> Examining the Utility of the IRAP in the Study of Gender with Normally-Developing Children	24
<b>Chapter 3</b> Examining the Utility of the IRAP in the Study of Gender and Self-Concept with Normally-Developing and Dyslexic Children	49
<b>Chapter 4</b> Examining the Utility of the IRAP in the Study of Gender and Self-Concept with Normally-Developing and ADHD Children	79
<b>Chapter 5</b> Examining the Utility of the IRAP in the Study of Teachers’ Implicit Attitudes to Children with Emotional Behavioural Difficulties	95
<b>Chapter 6</b> Examining the Utility of the IRAP in the Study of Teachers’ Implicit Attitudes to Normally-Developing Children	119
<b>Chapter 7</b> Systematic Comparisons among the IRAP Studies of Childhood and Adulthood	134

<b>Chapter 8</b>	General Discussion	146
<b>References</b>		166
<b>Appendix</b>		187

## **Acknowledgements**

Now the end is finally here, there are a few people I would like to thank who have supported and encouraged me through the past three years. To all the gang, especially those involved in ACT, but in particular, Andy who shared many a cup of coffee with me and one or two anxious moments! To all of the children who participated in the studies and to the schools and teachers who kindly accommodated me in an already busy daily schedule. To Anne Lodge in the Education Department who offered advice and encouragement and Anne and Derek who always seem to know exactly what to do. Finally, to Yvonne for giving me the encouragement and guidance to believe in myself and who gave so generously her time and support.

## **Abstract**

The aim of the current thesis was threefold. The first aim was primarily methodological and attempted to determine the utility of the IRAP as a measure of children's attitude to gender. Experiment 1 was the first study that attempted to employ the IRAP as a measure of children's implicit attitudes in relation to gender. The IRAP outcomes observed were consistent with previous IRAP studies and demonstrated the utility of the IRAP with children. The second aim was to determine if children who varied along the Special Educational Needs continuum would show a profile of implicit attitude to the self and others that was consistent with developmental profiles. Within this aim, Experiments 3 and 4 investigated the self-esteem profiles of three groups of children (normally-developing, dyslexic and Attention Deficit/Hyperactivity Disorder) by employing the IRAP and an explicit measure (the Piers-Harris Self Concept Scale). Across both Experiments the IRAP highlighted the differences between Normally-Developing children and those with Dyslexia and ADHD in terms of implicit self-esteem and their attitudes to others. While all three groups of children fell within the average range of self-esteem on the Piers-Harris differences were observed on the domain scales that were consistent with both the Dyslexic and ADHD groups academic and social difficulties. The third aim of the thesis was to examine teacher's implicit attitudes to pupils with Emotional and Behavioural Difficulties in mainstream education in comparison to Normally-Developing Pupils. The attitudes of three groups of teachers who varied in their experience and the educational context in which they were employed were examined (in-Training, Primary and Post-Primary) Comparisons between Experiments 4 and 5 revealed a significant difference in attitudes for Pupil versus EBD pupil for Primary teachers only.

The current findings indicate that the negative attitudes of teachers likely exacerbate existing feelings of inadequacy and thus impact negatively on direct interactions with children who present with a profile of Emotional Behavioural Difficulties in mainstream education

## Chapter 1

### General Introduction

# Chapter 1

## Introduction

Since the 1930's, social psychologists assumed that attitudes, and even stereotypes, operated in a conscious mode. But, there is increasing reason to believe that this is not the case and that human social decisions and actions are, in fact, guided by cognitive processes of which we are often not aware (Steele & Morawski, 2002). For example, individuals cannot choose their actions as freely as they or others expect (e.g. Latane & Darly, 1968; Milgram, 1963). Indeed, researchers in this area have argued that unconscious sources play a key role in perceptions, beliefs, preferences and actions. In order to encompass the full impact and range of these influences, Banaji, Blair and Glasser (1997) adopted the term '*microenvironments*' and argued that "attention to microenvironments means attending to the subtle and ongoing influences that shape preference and desires, knowledge and beliefs, and motives towards or away from other social objects" (p.67). According to these authors, lack of attention to microenvironments is likely to give rise to faulty theories of the self and others, because an individual cannot be consciously aware of the powerful implicit influences.

### *Defining Implicit Attitudes*

In order to identify the nature of these unconscious influences, social psychologists began to unpack the traditional concept of the '*attitude*' primarily because it failed to distinguish between conscious and unconscious processes (Greenwald & Banaji, 1995). However, other important features of what are now referred to as '*implicit attitudes*' have also been identified. Specifically, Eagly and

Chaiken (1998) proposed that attitudes encompass *evaluations* that reflect some degree of favour or disfavour. In addition, Schacter (1987) suggested that attitudes must, at least on occasion, involve some aspect of *implicit memory* because performance on a task does not require conscious or intentional recollection of previous experiences. Combining these two features, Greenwald and Banaji (1995) proposed the now widely accepted definition of an implicit attitude as: “introspectively unidentified (or inaccurately identified) traces of past experience that mediate favourable or unfavourable feeling, thought or action toward social objects” (p.8).

Other researchers have offered several additional features of implicit attitudes, including their unknown origin (i.e. people are unaware of the basis of their evaluations) and the fact that they are activated automatically (Wilson, Lindsey, & Schooler, 2000). Perhaps more critically, however, Wilson et al., (2000) also argued that implicit attitudes exert an influence over overt action, which individuals frequently do not attempt to control because they do not view these reactions as an expression of their attitudes. Put simply, the influence of implicit attitudes may be largely unavailable to self-report or introspection (Greenwald, 1990).

### *Explicit versus Implicit Measures*

The new focus on the implicit nature of attitudes, not surprisingly, suggested weaknesses in traditional means of measuring attitudes (de Jong, Pasman, Kindt, & van den Hout, 2001; Dovidio & Fazio, 1992; Gamar, Segal, Sagrati, & Kennedy, 2001). The susceptibility of self-report methodologies to deception and self-presentational strategies appears to be particularly problematic (e.g. de Jong, 2002; Teachman, Gregg, & Woody, 2001). The difficulties with this situation were

highlighted by de Houwer (2002) with an example of explicit clinical assessment tools. Specifically, he argued that various forms of psychopathology are underpinned by dysfunctional implicit beliefs that may exert significant influence on overt action (see also Beck, 1976). Yet, traditional clinical measures of psychopathology do not target implicit attitudes. As a result, it may be almost impossible for clinicians using traditional measures to assess or remediate problematic implicit processes associated with the disorder in question.

The focus on the implicit nature of attitudes has also created possibilities for new measures that might tap more directly into implicit processes. One such methodology is *priming*, which (as the term implies) involves a brief presentation of an object (e.g. photos of black or white faces) referred to as the '*prime*', followed immediately by an evaluative adjective (e.g. PLEASANT or AWFUL). Participants respond by rapidly selecting the evaluation most readily associated with the preceding prime (e.g. Black face-AWFUL and White face-PLEASANT). These associations are thereafter interpreted as indicators of implicit evaluations of the targeted objects (e.g. implicit attitudes to race (Fazio, Jackson, Dunton, & Williams, 1995).

The existing array of recently developed measures of implicit attitudes include the Go/No-go Association Task (GNAT: Nosek & Banaji, 2001) and the Extrinsic Affective Simon Task (EAST: de Houwer, 2003). Less well-established implicit procedures also include word fragment completion tasks (Gilbert & Hixon 1991; Tulving, Schacter, & Stark 1982; Warrington & Weiskrants, 1968); name-letter preference tasks (Koole, Dijksterhuis, & van Knippenberg 2001; Nuttin, 1985); and the original spatial Simon task (Simon, 1990). However, the Implicit Association Test (IAT) is by far the most commonly known and widely used implicit

methodology, with over 250 published studies of its use (Greenwald, McGhee, & Schwartz, 1998).

### *The Implicit Association Test (IAT)*

Consistent with the previous definition of implicit attitudes as unconscious evaluations, the IAT rests on the basic assumption that it is easier to map congruent than incongruent concepts (de Houwer, Geldof, & de Bruycker, 2005). Consider the following IAT trial in which participants are presented with Flower Names (e.g. TULIP), Insect Names (e.g. SPIDER), Positive Words (e.g. LOVE) and Negative Words (e.g. UGLY). Because of existing *associations* in memory between the concept Flowers and the attribute Positive and between the concept Insect and the attribute Negative, it is reasonable to assume that response options congruent with these associations (referred to as *consistent*) will be faster than those that are incongruent (referred to as *inconsistent* -- if Flowers was paired with Negative and Insects paired with Positive). The IAT, therefore, is a latency-based measure of the strength of the target concept-attribute associations or pairings. Hence, *the IAT effect* derives from the difference in mean latency between consistent and inconsistent associations and predicts superiority (i.e. shorter latencies) of the former over the latter (Greenwald et al., 1998).

The depth and range of supportive empirical evidence involving the IAT is more than impressive and spans a wide spread of phenomena across the discipline of psychology, including: clinical psychology (de Jong et al., 2001; Teachman et al., 2001); developmental psychology (Baron & Banaji, 2006; Dunham, Baron, & Banaji, 2006); neuroscience (Cunningham, et al., 2004); health psychology (Teachman, Gapsinski, Brownell, Rawlins, & Jeyaram, 2003); and even market research (Maison, Greenwald, & Bruin, 2001).

### *The Psychometric Properties of the IAT*

While the breadth of IAT effects is a core advantage to the utility of the procedure, there are other more methodological criteria that must be met for a methodology to be deemed acceptable, including evidence of: split-half reliability, internal consistency and test-retest reliability. Indeed, IAT researchers have devoted considerable attention to these psychometric properties. In terms of split-half reliability, the IAT appears to significantly outperform the EAST with scores that range between .83 and .95 and between .2 and .63, respectively (Teige, Schabel, Banse, & Asendorpf, 2004). Internal consistency estimates for the IAT are also promising and range from .7 to .9 compared to -.05 to .28 for other latency-based measures (Greenwald & Nosek, 2001). Finally, test-retest reliability outcomes are reported to be in the region of .56 with little variation in retest interval (Schmukle & Egloff, 2004).

Two other critical features that attest to the utility of implicit measures concern the malleability and durability of the effects. Although investigations of the fakeability of the IAT across multiple content domains have demonstrated considerably lower susceptibility to manipulation than self-report measures, there remains considerable controversy about the susceptibility of the procedure in its own right. On the one hand, findings from several studies suggest that the IAT is relatively insensitive to faking or manipulation. For example, Banse, Seise and Zerbes (2001) presented two groups of heterosexual participants with an IAT involving attitudes to homosexuality. One group was asked to pretend that they held really positive attitudes to homosexuality, while the other was simply instructed to respond as normal. However, both groups behaved in equally negative ways towards the homosexual stimuli. In a similar study, Fiedler and Bluemke (2003) exposed

three groups of participants to an IAT involving Turkish and German groups combined with positive and negative traits. The Control Group were simply asked to respond as normal. In contrast, both of the Faking Groups were explicitly instructed to try to fake their responses in a more socially desirable way, but only one group was permitted to slow down their response latencies in order to help them do so. Both Faking Groups were also provided with a pre-test in order to increase their levels of experience with the procedure and thus aid their ability to fake it. The results recorded across several experiments indicated that neither the pre-test exposure nor the explicit faking instructions per se exerted significant influence over the IAT effect (see also Asendorf, Banse, & Mucke, 2002; Egloff & Schmuckle, 2002).

Other studies, however, have reported less positive evidence for the IAT effect in the context of faking. For example, Fiedler and Bluemke (2003) demonstrated that instructing participants to slow down responding *combined with* faking instructions did significantly reduce the size of the effect. These results were consistent with those reported by Kim (2003) who also found that participants could fake their IAT responses, but only when instructed to respond more slowly on a subset of trial-types. In summary, therefore, no firm conclusions can yet be drawn about the susceptibility of the IAT to faking (Nosek, Greenwald, & Banaji, 2007).

A number of IAT studies have examined the durability of IAT effects across repeated exposures. Using the web-based data (on which participants indicate the number of previous exposures to the procedure), Nosek, Banaji and Greenwald (2002a) created a conservative subset that included only participants with one or no previous exposures. Comparisons between this and the full data set indicated strong similarities overall that suggested that level of exposure did not influence the IAT effect significantly. Furthermore, Nosek et al. (2007) have argued that the use of the

*D* scoring algorithm has reduced susceptibility of the effects in this regard in more recent studies. However, the data reported by Nosek et al., (2007) is inconsistent with existing evidence by Greenwald, Nosek and Banaji (2003). In this study, the researchers used the same web-based datasets to evaluate alternative scoring procedures, but distinguished on this occasion only between zero (i.e. no previous exposure) and one plus exposure(s) (i.e. the actual number of previous exposures more than one were ignored). The results indicated that the magnitude of the IAT effect decreased with repeated exposure because all correlations with prior experience were uniformly negative. Taken together, therefore, there is as yet no clear consensus regarding the impact of repeated exposures on the IAT effect.

#### *Concordance between Implicit and Explicit Measures*

There is considerable debate within the IAT literature on concordance between implicit and explicit measures (Nosek, Banaji & Greenwald, 2002b), with equally many studies reporting the presence or absence of concordance (Greenwald & Banaji, 1995). In general, earlier studies were characterised by absent or weak correlations with self-report measures (Greenwald et al., 1998), while later studies reported stronger correlations with same (e.g. Greenwald et al., 2003; Nosek, 2005). For example, in a meta-analysis of correlations between the IAT and self-reports, Hoffman, Gawronski, Gschwender, Be, & Schmitt, (2005) reported a mean effect size of .24 based on a collection of 126 studies that involved at least one IAT and one explicit self-report measure. In one study, Nosek (2005) for example measured implicit and explicit preferences for a variety of evaluative target domains from a range of topics of social significance. The average implicit-explicit correlation obtained in this study was .36 with 52 out of the 57 object pairs.

A related issue concerns the relationship between implicit attitudes and overt behaviour (which itself may or may not be correlated with explicit outcomes). One study by Asendorpf et al., (2002) demonstrated a correspondence between the IAT and spontaneous overt behaviour in the domain of shyness. Specifically, participants were observed in a shyness-inducing laboratory situation and completed an explicit self-report measure of shyness. A moderate correlation was recorded between this measure and the IAT ( $r = .66$ ). Interestingly, the IAT predicted spontaneous (but not controlled) shy behaviour, whereas the explicit measure predicted the reverse (controlled but not spontaneous behaviour). Thus, a double dissociation was found between the implicit and explicit trait of shy behaviour, in spite of the presence of a weak positive correlation.

A combination of other studies has demonstrated a similar relationship between implicit attitudes to anxiety and behavioural indices. For example, Egloff and Schmukle (2002) reported that the internal consistency of the Anxiety-IAT (i.e. assessing associations of self versus other with anxiety-related versus calmness related words) was .77. However, the anxiety effect did not correlate with explicit anxiety. Interestingly, the IAT predicted changes in experimenter-rated anxiety, changes in performance after stress, and other behavioural indicators of anxiety, but the explicit measures did not.

According to Greenwald et al. (2003), explicit-implicit correlations are greatly influenced by factors such as participants, types of measures and stimulus type. The MODE model reported by Dovidio and Fazio (1992) proposed that the magnitude of the relation between an implicit and an explicit measure depends in particular upon participants' motivation and opportunity to deliberate. According to this view, if either motivation or opportunity is relatively low at the time that the explicit response

is being considered, then a greater correlation will be observed with the implicit measures. However, when both motivation and opportunity are relatively high, the two types of measure are less likely to correlate. One would predict, therefore, that material of greater sensitivity or social significance will enhance motivation and therein increase discordance (Fazio & Olsen, 2003).

### *Conceptual Issues with the IAT*

According to Greenwald et al. (2003), the relationship between explicit and implicit measures may also be used as an index of the validity of what is being measured, such that the higher the concordance, the greater the construct validity. This argument, however, assumes some degree of overlap between the concepts targeted by the two types of measure. Indeed, there has been considerable debate within the broader field of implicit cognition about what the IAT is actually measuring. It is generally proposed that the IAT effect is a measure of individual differences in terms of the strength of associations between concepts in memory (Greenwald et al., 1998). However, de Houwer (2002) has argued that implicit measures should target beliefs, which should comprise of much more than associations between concepts:

Greenwald et al. (1998) designed the IAT to assess the strength of associations between concepts in memory. One can argue that beliefs involve more than just associations between concepts. First, beliefs reflect qualified associations. For instance, the belief “I am a bad person” implies *a special type of association* between the concept “self” and the concept “bad”, namely *a directional association* which specifies that “bad” is a property or characteristic of “self”. IAT effects do not reflect the *nature or directionality of an association between concepts*, they can reflect only strength of association. Second, many beliefs involve several associations and several concepts. For instance, conditional beliefs such as “if I do not perform well on a task, then I am an inferior person” involve rather complex structures of qualified associations between several concepts. The IAT cannot be used to directly capture such complex conditional beliefs (also see de Jong et al., 2001, p. 111) . . . In sum, the IAT does not provide a measure of beliefs, nor was it designed to do so. It can only provide an index of associations that are assumed to be involved in certain beliefs and thus indirect evidence for the presence of certain beliefs (pp. 117-118, emphases added).

For de Houwer, therefore, IAT effects do not reflect the nature or directionality of associations between concepts, only the strength of associations.

Although part of the difficulty for the IAT in terms of simply offering a measure of the strength of associations results from the physical format presented within the procedure, researchers have also raised broader concerns about the utility of the basic concept of the association. For example, Barnes-Holmes, Barnes-Holmes Hayden, Milne, and Power (2006) have argued that the IAT's reliance upon the concept of the association is not problematic if one assumes that associative processes underpin attitudes, beliefs and indeed all aspects of higher cognition. Difficulties arise, however, if this assumption is not held and one assumes alternatively that associations are but one aspect of human cognition. This latter perspective characterises the approach offered by Barnes-Holmes et al., which relies heavily on *relations* rather than associations. A summary of this approach is provided in the sections that follow.

### *Studying Stimulus Relations*

Relational Frame Theory (RFT) is as a modern behavioural account that views human language and cognition as primarily relational acts (Hayes, Barnes-Holmes, & Roche, 2001). Thus, RFT proposes that stimulus relations rather than associations are central to the understanding of human psychological events (Hayes & Barnes-Holmes, 2004). Within the array of empirical evidence in support of the theory, a number of experimental methodologies have emerged that appear to tap into the range and sophistication of stimulus relations and the processes by which these are established and maintained.

*Matching-to-Sample (MTS)*. Derived stimulus relations and relations of coordination in particular (often referred to as *stimulus equivalence*) were traditionally examined in the context of the matching-to-sample (MTS) paradigm (Sidman, 1971). In a typical MTS task, participants are presented with a computer screen containing three or more stimuli (e.g. nonsense syllables). One stimulus (i.e. the *sample*) is usually presented in the middle of the screen, with the remaining stimuli (i.e. the *comparisons*) located side by side below the sample. In order to establish the target conditional discriminations between the sample and comparison stimuli, differential reinforcement is provided for selecting a specific comparison stimulus in the presence of a specific sample. That is, for example, reinforcement may be provided for selecting VEK (rather than JOM) in the presence of CUG, but JOM (rather than VEK) in the presence of DAX. In this way, the conditional discriminations CUG-VEK and DAX-JOM would be established. Having established target CUG-VEK and DAX-JOM relations, participants may then be trained to establish VEK-PAV and JOM-ROP relations, where the previously correct comparison stimuli now function as samples (VEK and JOM) and two new nonsense syllables (PAV and ROP) become the correct comparisons, respectively. Altogether then, the relations CUG-VEK, VEK-PAV, DAX-JOM and JOM-ROP are directly trained. Because the nonsense syllables have no inherent meaning and the relations among them are not based on any consistent formal dimension, the latter are described as *arbitrary* (because the Researcher arbitrarily assigns the comparison stimulus that is ‘correct’ or ‘incorrect’ in the presence of a particular sample).

For RFT, it is not the emergence of explicitly trained conditional discriminations that renders stimulus relations such as these relevant to human language and cognition, but the *derived* or non-trained relations. In the previous

example, given VEK as a sample, verbally competent humans readily select CUG (rather than DAX) even though only the reverse relation (CUG-VEK) was explicitly trained. According to RFT, this novel untrained performance is described as a mutually entailed bi-directional relation of co-ordination between CUG and VEK, in which explicit training of one direction (CUG-VEK) mutually entails the other direction (VEK-CUG) without additional training. The ability to derive these simple relations appears to develop very early in the human lifespan (Lipkens, Hayes, & Hayes, 1993).

Even more complex performances have been observed using the MTS paradigm. Continuing with the previous example, participants also readily demonstrate combinatorially entailed relations when they select PAV in the presence of CUG and vice versa. In this performance the mutually entailed CUG-VEK and VEK-PAV relations generate the combinatorially entailed CUG-PAV and PAV-CUG relations, again without additional training. As well as considerable empirical evidence in support of these derived performances, there is additional support for RFT predictions that: (1) stimulus functions transfer and transform in accordance with the relations (Barnes-Holmes, Hayes, Dymond, & O’Hora, 2001); (2) more complex types of relations (e.g. distinction, opposition, hierarchy and even perspective-taking) are readily observed (Hayes, Fox, Gifford, Wilson et al., 2001) and (3) similar performances are recorded with a broad array of socially and personally meaningful terms (Leslie, Tierney, Robinson, Keenan, & Watt, 1993; Watt, Keenan, Barnes, & Cairns, 1991).

Although many studies of co-ordination and other relations successfully employed the MTS format, behavioural researchers have questioned its utility, especially in the context of multiple (i.e. non-co-ordinated) stimulus relations and

meaningful stimuli (e.g. Barnes-Holmes, Staunton, Barnes-Holmes, McHugh, et al., 2004; Leader & Barnes-Holmes, 2004). For example, although verbally competent adults would undoubtedly be capable of deriving relations of co-ordination, individuals commonly fail to complete the conditional discrimination training or to demonstrate the target derived performances on MTS trials. In response to this, researchers argued that the MTS environment may not contain adequate or appropriate contextual cues for the target relational responses (e.g. humans derive almost all of their stimulus relations in a free verbal environment and not in the context of an automated procedure with nonsense syllables).

*The Relational Evaluation Procedure (REP).* In response to the problems encountered with MTS, RFT researchers developed the Relational Evaluation Procedure (REP), which allowed participants to evaluate, or report on, the stimulus relation presented on a given trial (Barnes-Holmes, Healy, & Hayes, 2000; Hayes & Barnes, 1997). For example, a participant may be presented with a contextual cue for DIFFERENT and two arbitrary stimuli (e.g. geometric shapes) that participate in a difference relation (controlled by the cue). The participant is required to choose one of two arbitrary stimuli, each of which has the previously established functions of TRUE and FALSE. Hence, in this case a correct response would involve the participant selecting the shape that had the function TRUE (i.e. it is true that the cue for DIFFERENT controls a difference relation) because the DIFFERENT cue and the arbitrary stimulus correspond relationally. Hence, in this case if the cue is SAME the participant should choose FALSE. Several studies demonstrated the utility of the REP in analyses of relational responding in adult participants (Cullinan, Barnes, & Smeets 1998; Cullinan, Barnes-Holmes, & Smeets, 2000, 2001).

Although both MTS and the REP had considerable success in demonstrating and facilitating derived relations, RFT researchers have also argued that if derived stimulus relations are relevant to verbal behaviour, then alternative methodologies germane to other schools of psychology might be equally useful as an experimental forum for generating and investigating derived relations, especially with meaningful stimuli. In response to this suggestion, a group of RFT researchers began to investigate the IAT and developed a new experimental methodology that offered a hybrid between the IAT and the REP that came to be known as the Implicit Relational Assessment Procedure (IRAP). What is particularly important about the IRAP is that its' focus on targeting stimulus relations directly not only offers a good test of RFT predictions, but may well permit a closer examination of the relative utility of the concepts of the association and the relations, as well as sound methodological comparisons with the IAT.

#### *The Implicit Relational Assessment Procedure (IRAP)*

The IRAP is an automated latency-based task with basic performance assumptions similar to those of the IAT. That is, participants are required to respond quickly and accurately in a manner that is either consistent or inconsistent with predicted implicit attitudes. Specifically, response latencies should be shorter on consistent trials relative to inconsistent trials (Barnes-Holmes et al., 2006).

Although aspects of the visual presentation of the IRAP are similar to the IAT, the primary difference is that the former contains only one basic type of task, rather than a range of tasks that increase in complexity across consecutive blocks. All IRAP trials involve a single sample stimulus, a single comparison stimulus and two relational terms that match response options (so that the properties of the relations

between the target stimuli can be directly assessed). Consider a trial from Experiment 1 of the first published IRAP study reported by Barnes-Holmes et al. (2006). This IRAP involved the presentation of the attribute words PLEASANT or UNPLEASANT as the samples, with a range of target words that would readily be categorised as consistent or inconsistent with either sample. Specifically, the target words consistent with PLEASANT comprised common words normally associated with positive evaluations (CARESS, FREEDOM, HEALTH, LOVE, PEACE and CHEER). In contrast, the target words consistent with UNPLEASANT comprised common words normally associated with negative evaluations (ABUSE, CRASH, FILTH, MURDER, SICKNESS and ACCIDENT).

Although only one sample and one comparison stimulus appear on each trial, both relational terms (e.g. SIMILAR and OPPOSITE) are presented simultaneously. Thus, on each trial participants respond to the relations between the sample and the target in terms of one or other of the relational terms (e.g. PLEASANT-CARESS-SIMILAR). Because all IRAP trials are identical in format, they are differentiated only in terms of the type of responding that is deemed consistent or inconsistent with predictions. On consistent trials presented during the example above, responses are designated as correct when participants select PLEASANT-Positive target-SIMILAR; PLEASANT-Negative target-OPPOSITE; UNPLEASANT-Positive target-OPPOSITE; and UNPLEASANT-Negative target-SIMILAR. In other words, these responses match those normally attributed to the general population (i.e. positive words as pleasant and negative words as unpleasant). In contrast, correct responding during inconsistent trials would involve participants selecting PLEASANT-Positive target-OPPOSITE; PLEASANT-Negative target-SIMILAR; UNPLEASANT-Positive target-SIMILAR; and UNPLEASANT-Negative target-OPPOSITE. As expected,

these responses are counter to those normally attributed to the general population (i.e. positive words as unpleasant and negative words as pleasant).

Although the IRAP contains a similar number of blocks of trials to the IAT, it comprises of a set sequence of alternating consistent and inconsistent trials that are identical in format but across which participants are required to switch between consistent and inconsistent responding. In Experiment 1 of the research by Barnes-Holmes et al. (2006), the IRAP sequence comprised of a total of eight blocks – two practice blocks (one consistent and one inconsistent), followed by six test blocks (in alternating sequence between consistent and inconsistent responding). Each block was comprised of 24 trials, with each of the six target words presented once with each sample (i.e. six PLEASANT-Positive; six UNPLEASANT-Positive; six PLEASANT-Negative; and six UNPLEASANT-Negative trials).

The results of Experiment 1 were entirely consistent with experimental predictions (Barnes-Holmes et al., 2006). That is, mean response latencies were significantly shorter on consistent than inconsistent trials. In other words, on consistent trials participants more readily responded PLEASANT-Positive-SIMILAR (than OPPOSITE); PLEASANT-Negative-OPPOSITE (than SIMILAR); UNPLEASANT-Positive-OPPOSITE (than SIMILAR); and UNPLEASANT-Negative-SIMILAR (than OPPOSITE) on consistent trials. On inconsistent trials, they more readily responded, as expected, with PLEASANT-Positive-OPPOSITE (than SIMILAR); PLEASANT-Negative-SIMILAR (than OPPOSITE); UNPLEASANT-Positive-SIMILAR (than OPPOSITE); and UNPLEASANT-Negative-OPPOSITE (than SIMILAR).

Although Experiment 1 of the original research provided some evidence of predicted IRAP effects, it was reasonable to argue that the simplicity of the target

distinction between PLEASANT and UNPLEASANT and the positive versus negative target words could be delivered by even a relatively insensitive procedure. Hence, it remained to be seen whether the IRAP could tap into more complex and more socially sensitive relations. In Experiment 2 of the original work, the attitudes of three groups who differed in their range of experience working with children with Autistic Spectrum Disorder (ASD) were measured using the IRAP in conjunction with an explicit measure of attitudes to this population (the Challenging Behaviours Attribution Scale – CHABA -- Hastings, 1997). Specifically, one group had some years of experience of children with this disability, another groups had months of experience and the third group had no experience. During this IRAP, the phrases AUTISTIC SPECTRUM DISORDER and NORMALLY-DEVELOPING were presented as samples; positive words (HAPPY, GOOD, EASY, CALM, POSITIVE and PLEASANT) and negative words (SAD, BAD, DIFFICULT, ANGRY, NEGATIVE and UNPLEASANT) were targets, and SIMILAR and OPPOSITE were the relational terms. In this case, it was predicted that on consistent trials participants would more readily pair normal with positive and ASD with negative, with reverse performances on inconsistent trials.

Once again, the results of Experiment 2 were consistent with predictions, with all three groups showing a bias in favour of normally-developing children and against children with ASD. That is, the mean response latencies were significantly longer on inconsistent tasks relative to consistent tasks for each of the three groups. Although no significant differences were recorded among the groups (all were equally biased), they did differ significantly on the explicit measure. Specifically, the most experienced group showed significantly more positive attitudes to autism than the least experienced group. This implicit-explicit discordance suggested the possibility

that the IRAP was tapping into implicit attitudes of the groups to this socially sensitive material that could not be manipulated, whereas the explicit measure outcomes were more politically correct and influenced predictably by experience.

Although published IRAP evidence thus far remains limited, there is good reason to suggest the promise of the methodology and good support for the utility of the concept of the stimulus relation, at least as part of the broader concept of the implicit attitude. The simplicity and yet the relational focus of the new procedure appear to offer considerable advantages over both MTS and the REP. Furthermore, its similarity (in some respects) with the IAT suggests its potential as an implicit measure and offers the possibility of comparisons with the original procedure.

One area of investigation that offered a potentially good test of the IRAP but which had received only little IAT attention was the utility of the procedure with children. Research in this area would offer a number of advantages: (1) comparisons could be drawn between the IRAP and relevant IAT studies; (2) the development of target relations might be investigated; (3) the simplicity of the procedure would be tested by a sample of participants of lower verbal sophistication; (4) the social or personal applicability of the procedure could be targeted (e.g. in the context of implicit attitudes to gender or self-esteem). Existing IAT research conducted with children is reviewed in the section that follows.

### *The Development of Implicit Cognition*

With well-established outcomes with adult samples, researchers have begun to employ the IAT with children, although the number of studies in this area remains limited. In one of the earliest studies, Skowronski and Lawrence (2001) compared the implicit and explicit attitudes of eleven-year old children and college students towards

men and women and towards male soldiers and female soldiers. In line with previous research, the authors predicted that explicit attitudes towards women would be more positive than men, but argued that the implicit attitudes would likely be influenced by both the gender and age of the participants (see also Eagly, Mladinic, & Otto, 1991). The IAT results indicated that female college students demonstrated pro-female and pro-female soldier preferences. In contrast, male college students showed a pro-male soldier preference only (i.e. no pro-male preference). Interestingly, the eleven-year olds showed no preference in either direction on the gender or soldier categories. The authors argued that the difference between the two groups of participants was accounted for by the fact that the children had not fully assimilated standard sex-typed gender roles.

Although the Skowronski and Lawrence (2001) study indicated that eleven-year olds employed the IAT with ease, subsequent research with six-year olds reported considerable difficulties with this younger sample and the procedure (Rutland, Cameron, Milne, & McGeorge, 2005). In response to these problems, the researchers then developed a more pictorially-based procedure (referred to as the *Ch-IAT*) in which the attribute words PLEASANT and UNPLEASANT were replaced with simple line drawings of Happy and Sad cartoon faces, respectively (Rutland, et al., 2005). Instead of responding with keyboard presses, the children were also ALLOWED to move the mouse either towards or away from themselves.

The researchers then employed the Ch-IAT to investigate the susceptibility of children's explicit and implicit racial bias to public awareness. Specifically, in a Public Condition, the children were openly videotaped when completing both the explicit and implicit measures of race. Alternatively, in the Private Condition, all aspects of the procedure were identical, except that the children were explicitly

instructed that they were not being videotaped. During the Ch-IAT, they were presented with unfamiliar neutral Black and White faces as concepts, with the Happy and Sad cartoon faces as the attributes. The results of the study indicated that the six-year olds in the Public Condition (but not in Private) attempted to inhibit their in-group bias on the explicit measure. The opposite pattern emerged with the ten-year olds who attempted to inhibit their explicit bias in the Private rather than Public Condition. In an attempt to account for these developmental differences, the researchers argued that the younger group were externally motivated to control their implicit biases, whereas the older group were internally motivated to do so.

The Ch-IAT was subsequently modified again for a later study by Baron and Banaji (2006) to investigate the developmental profile of implicit racial attitudes with samples of children aged six and ten years, as well as with a group of adult participants. This modification involved the standard use of racial pictures (e.g. Black and White faces), but the printed attribute words were replaced by audio recordings (to control for variability in the children's reading levels) of the words GOOD and BAD. The response options were modified again to enable the children to respond differentially with simple button presses, with one button press for GOOD and another for BAD. Furthermore, all of the instructions were automated and presented as audio recordings and a Researcher was present throughout the procedure.

In their first study, Baron and Banaji (2006) simply piloted the modified Ch-IAT with the standard Insect/Flower-GOOD/BAD pairings and reported strong and predicted IAT effects. In the second study, the racial attitudes of White American six and ten-year olds were investigated. The results of this study demonstrated that the children had well-established and predicted implicit pro-White/anti-Black biases that were consistent with the adults and with standard IAT data with numerous other

White American adult samples. These data were further supported by findings from a very similar study with American and Japanese children six- and ten-year olds (Dunham, Baron, & Banaji, 2006).

### *The Current Thesis*

Part 1 of the current thesis was primarily methodological and attempted to determine the utility of the IRAP with children. The first empirical chapter targeted the children's attitudes to gender, with the child's own gender label in combination with Positive and Negative evaluative terms. The effects observed in Experiment 1 represented the first IRAP outcomes recorded with young children. However, it might be argued that presenting gender associations to children is not a particularly good test of the IRAP.

Experiment 2 in Chapter 3 attempted to extend the experimental aims of the previous study by targeting the more amorphous concept of self-esteem and involved comparisons between Normally-Developing children and a small sample who presented with Dyslexia. The potential correlations between implicit attitudes to self and the Piers-Harris Self-concept Scale was also examined. The IRAP highlighted some differences between the two groups of children in terms of their implicit attitudes to others, however, their attitudes to self on both the implicit and explicit measures indicated that both groups of children had normally high levels of implicit and explicit self-regard. No correlations between the implicit and explicit measures were observed.

It might be reasonably argued that Dyslexia as a specific learning difficulty does not pose sufficient intellectual, emotional or educational difficulties to significantly differentiate this group of children from their normally-developing peers

with regard to self-esteem. Chapter 4 in the current thesis extended the previous studies by examining the self-esteem profile of children with Attention Deficit/Hyperactivity Disorder (ADHD). It is well established that this group of children have more difficulties in education than children with Dyslexia, and their Normally- Developing peers. The IRAP outcomes revealed considerable differences between the two groups and indicated that the ADHD children had less positive implicit attitude to self relative to the Normally-Developing group. Differences were also observed on the domain scales of the Piers-Harris 2, specifically on the Behavioural Adjustment and Popularity sub-scales.

Part 2 of the current thesis employed the IRAP to examine what may be viewed as the other side of attitudes to Special Educational Needs (SEN) in terms of teachers' attitudes. Chapter 5 employed the IRAP as a measure of three groups of teachers' implicit attitudes to children with Emotional Behavioural Difficulties (EBD) in mainstream education. These groups were differentiated by their experience in teaching and the educational contexts in which they were employed (i.e. primary and post-primary schools). The IRAP effects observed in Experiment 4 indicated that the three groups of teachers showed strong implicit positive bias towards teachers and negative (albeit weaker) attitudes to EBD Pupil. Participants were also asked to complete an explicit measure, the Opinions Relative to Mainstreaming scale (ORM scale), which indicated that while all groups were unanimously in favour of including children with emotional behavioural difficulties in mainstream education they were more reticent when asked about the placement of this population within their own classrooms.

An issue that emerged from Chapter 5 concerned the possibility that the IRAP was tapping into the differences between the teachers' implicit attitudes to teachers

versus all pupils. Hence, Chapter 6 examined the implicit attitudes of two groups of teachers (primary and post-primary) to Normally Developing children. The IRAP results indicated that both groups showed implicit bias towards teachers with the effect being much stronger for Primary over Post-Primary teachers. Their attitudes towards pupils were mixed, (positive and negative) and both groups of teachers were similar in this regard. Both groups also completed the ORM scale and the results were similar to those reported in chapter 4.

While the studies within the current theses were conducted separately a broader view of the work was examined by drawing comparisons among the relevant experiments in Chapter 7. Experiments 2 and 3 are compared in an analysis of the childhood studies of gender and self-esteem with results on the IRAP indicating differences between the Normally-Developing group and those groups of children with SEN. Differences were also observed on the specific domain scales scores for the Dyslexic and ADHD groups of children in comparison to their Normally-Developing peers. Comparisons between Experiments 4 and 5 in teacher's attitudes to Normally-Developing versus EBD children revealed a significant difference for the Primary group only, in their implicit attitudes to Pupil versus EBD Pupil.

In the eight and final chapter the findings of all 5 Experiments are discussed. The research reported in the current thesis suggests that the IRAP was tapping into important issues that comprise one's attitudes to oneself and others. The benefits of integrating different types of questions and different methodologies has also been highlighted and suggests that such integration maybe useful for the solving of future research questions, which will need to direct attention towards the relationship for example, between implicit attitudes and overt behaviour.

## Chapter 2

### Examining the Utility of the IRAP in the Study of Gender with Normally-Developing Children

## Chapter 2

# Examining the Utility of the IRAP in the Study of Gender with Normally-Developing Children

### INTRODUCTION

Only a limited number of studies to date have attempted to employ the IAT with children (in a procedure commonly known as the Ch-IAT – see Chapter 1). Indeed, the only areas of children’s implicit attitudes that have received empirical attention in this context are gender preferences (Eagly et al., 1991) and racial bias (Baron & Banaji, 2006; Rutland et al., 2005). Perhaps of greatest importance across both areas was the fact that the IAT *was successfully* employed as a measure of implicit attitudes in children, although Baron and Banaji had reported the need for some procedural modifications (e.g. replacing printed attribute words with audio recordings).

The IAT outcomes for race were consistent with experimental predictions and previous results obtained with adults in demonstrating that both six- and ten-year olds had well-established implicit pro-White/anti-Black biases. Perhaps surprisingly, however, the IAT outcomes for gender indicated that eleven-year olds showed no distinct pattern of gender preferences, although adults did in the same experiment. Indeed, in an attempt to account for the children’s outcomes, Baron and Banaji (2006) argued that children of this age group had not fully assimilated standard sex-typed gender roles. Taken together, it is difficult to imagine that children, particularly at age ten, did not have established gender stereotypes, or that these are preceded in the normal course of

development by racial stereotypes. Furthermore, these findings are largely inconsistent with a considerable wealth of developmental evidence to suggest that gender stereotypes should be well established by age ten.

### *The Development of Gender Identities*

Attitudes to gender are generally conceptualised as a core feature of the self-concept (Sroufe, 1990a), the latter of which appears to be well-established as early as 30 months (with spontaneous pointing to the self in pictures and use of own name -- see Butterworth, 1990). Indeed, some researchers have argued that gender distinctions begin to emerge as early as nine or twelve months when males and females may be distinguished on the basis of external features, such as hair length (Fagot & Leinbach, 1993). By 24 to 30 months, toddlers can already accurately identify pictures of girls and boys and categorise their own gender appropriately (Wenar & Kerig, 2006). Nonetheless, these distinctions do appear to be based largely on external (size, clothing and hairstyle) rather than internal cues (Golombok & Fivush, 1994). As a result, children at this age are believed to be able to distinguish one gender from another, but do not possess a coherent understanding of gender as a concept.

The under-developed gender concept of young children encompasses excessively rigid gender responses that characterise pre-schoolers from older children (and are also consistent with generic patterns of pre-operational thinking). For example, a young child may believe that 'if a young boy wears a dress he is a girl' because only girls wear dresses. In fact, it is only around the age of six or seven that children understand that

gender is permanent and immutable and that genitals are the crucial factor in the determination of one's gender (Rathus, 2006).

According to Bem's 'gender-schema theory', a more sophisticated gender schema comprises a cluster of gender-related attributes, including physical features, behaviour and personality traits. For example, on the dimension of strength versus weakness, children learn to associate strength with males and weakness with females, and these attributes are then weighted relative to others in becoming part of the gender stereotype (Rathus, 2006). For instance, boys attribute more value to sports achievement (because it is associated with physical strength) in categorising males than in categorising females.

Gender-identities and the relevant features thereof form a critical part of parental and peer influences and thus not surprisingly are central to the development of the self-concept (Santrock, 2004). For example, play that involves gender-appropriate activities receives higher levels of reinforcement by peers than cross-gender activities (Buhrmester, 1993; Macoby, 1998) and this pattern continues well into adolescence (Macoby & Jacklin, 1987).

In summary, therefore, there is a wealth of research evidence that suggests that gender preferences, particularly in terms of one's associations with one's own gender, are well-established by age ten. This calls into question the accuracy of the IAT findings reported by Skowronski and Lawrence (2001) who found no evidence of implicit gender preferences in this age group using the original (unmodified) IAT.

### *The Current Study*

The primary aim of Experiment 1 was to simply examine the utility of the IRAP (i.e. used previously only with adults) with a sample of children. Because it would seem to be the case that gender concept is established by age eight, and certainly that children of this age can readily identify their own gender, the secondary aim was to examine children's implicit attitudes to gender. Because of the intricate developmental relationship between gender and the self, we targeted the children's attitudes to gender in terms of positive versus negative associations regarding their own gender and the gender of others. Put simply, the experimental question sought to determine whether the children would more readily associate their own gender with positive terms and the other gender with negative terms in the context of the IRAP?

## METHOD

### *Participants*

Eight children, aged between 11 and 15 years old (mean age of 11 yrs 9 months) participated in Experiment 1. The children were recruited through direct contact with their parents, hence parental consent for participation was obtained directly. All participants were categorised as within their age-appropriate developmental ranges and had no prior evidence of learning difficulties. None of the children had been previously exposed to an IRAP, nor had any knowledge of the procedure. Participants were initially divided into two groups according to gender (i.e. four males and four females).

### *Setting*

For all participants, Experiment 1 was conducted in a quiet room in each child's home. As much as possible, the room remained free from noise and distraction throughout all aspects of the study. All participation was conducted on an individual basis. The Researcher remained present in the room at all times, seated beside the participant. A parent/guardian of each child was not present in the room during the experiment, but was available nearby if requested by the child or if deemed necessary by the Researcher (this did not occur).

### *Apparatus and Materials*

Participants completed all experimental trials on the same DELL laptop computer with a Pentium 4 processor. The IRAP was delivered via a program written in Visual

Basic (Version 6.0.), which controlled all aspects of stimulus presentation and the recording of all participant responses.

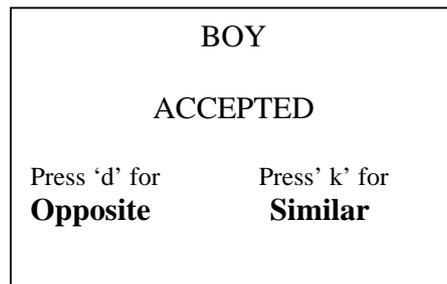
The stimuli presented during the IRAP comprised of three groups of words (sample stimuli, target stimuli and response options). The *sample stimuli* contained only the two gender words BOY and GIRL. These were presented to participants as the word that was consistent with their ‘own gender’ (i.e. GIRL for female participants and BOY for male participants) or the word that was consistent with the ‘other gender’ (i.e. GIRL for male participants and BOY for female participants). The target stimuli contained six evaluative terms – three positive (ACCEPTED, POPULAR and PERFECT) and three negative (FAULTY, BROKEN and USELESS). The sample and target words were generated from preliminary work carried out with children who did not subsequently participate in the study. The response options contained only the two relational terms SIMILAR and OPPOSITE. The stimulus arrangements employed in Experiment 1 are presented in Table 1.

TABLE 1  
The Stimulus Arrangements Employed in Experiment 1.

<b>Sample Stimulus 1</b> OWN GENDER (BOY/GIRL)	<b>Sample Stimulus 2</b> OTHER GENDER (BOY/GIRL)
<b>Response Option consistent with Sample 1</b> SIMILAR	<b>Response Option consistent with Sample 2</b> OPPOSITE
<b>Targets Stimuli consistent with Sample 1</b>	<b>Targets Stimuli consistent with Sample 2</b>
ACCEPTED POPULAR PERFECT	FAULTY BROKEN USELESS

In order to ensure that the children received adequate instructions regarding the procedure, a series of visual aids were developed that comprised of small schematic representations of each of the four trial-types. This yielded a total of four pictures – OWN GENDER (e.g. BOY) presented with a positive target (see Figure 1); OWN

GENDER with a negative target; OTHER GENDER (e.g. GIRL) with a positive target; and OTHER GENDER with a negative target.



---

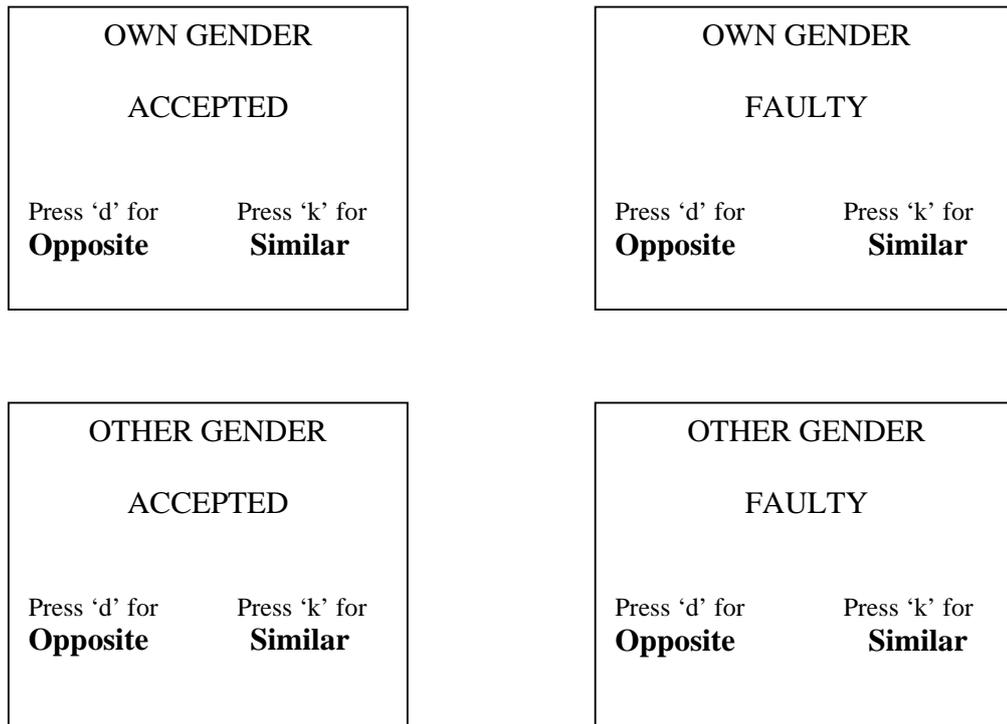
Figure 1: Visual aid containing a schematic representation of an IRAP trial.

---

### *Trial-types*

The four visual aids presented to the children during the instructional phase were derived from the four basic trial-types contained within the IRAP. These are presented in Figure 2. The four trial-types were constructed from the configurations of the sample stimuli (i.e. OWN GENDER and OTHER GENDER) and the two types of target stimuli (i.e. positive and negative evaluations). Hence, for experimental purposes, the four trial-types were referred to as OWN GENDER/Positive (top left of Figure 2); OWN GENDER/Negative (top right); OTHER GENDER/Positive (bottom left); and OTHER GENDER/Negative (bottom right). Each block of IRAP trials contained six exposures to each of the four trial-types randomly presented across the block (i.e. a total of 24 trials in each block). On each of the six exposures to a trial-type, each of the three target stimuli appeared twice with the same sample stimulus. For example, each participant was

presented with six OWN GENDER/Positive trials (i.e. two ACCEPTED, two POPULAR and two PERFECT).



---

Figure 2: Schematic representation of the four basic IRAP trial-types.

---

### *Consistent and Inconsistent Responding*

The recording of a response on any trial as correct or incorrect depended upon whether the trial (for experimental purposes) had been categorised as consistent or inconsistent. Trials were referred to as *consistent* when the relations among the sample

and target stimuli were consistent with the views believed to be held by the experimental participants. That is, it was predicted that the children here would more readily associate the target positive evaluations than the negative evaluations with their own gender, and the negative (rather than positive) evaluations with the other gender. Consider the consistent trials presented in Figure 3. If a female participant was presented with GIRL and ACCEPTED, a correct response involved selecting SIMILAR (i.e. GIRL SIMILAR to ACCEPTED). An incorrect response in this context would involve selecting OPPOSITE (i.e. GIRL OPPOSITE to ACCEPTED).

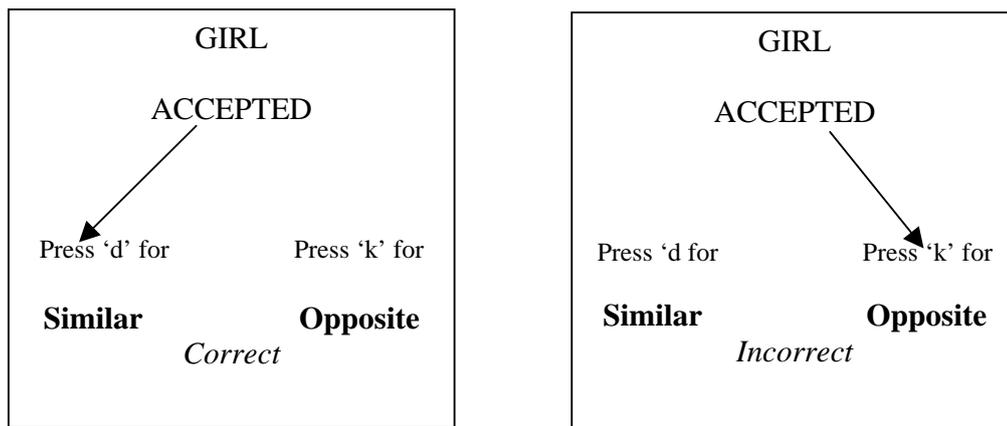


Figure 3: Schematic examples of consistent trials for females.

Trials were referred to as *inconsistent* when the relations among the sample and target stimuli were *not* consistent with views attributed to the participants. Consider an inconsistent trial in which a female participant was presented with GIRL and ACCEPTED. A correct response in this case involved selecting OPPOSITE (i.e. GIRL OPPOSITE to ACCEPTED), whereas an incorrect response involved selecting SIMILAR

(i.e. GIRL SIMILAR to ACCEPTED). Both consistent and inconsistent trials, therefore, involved the selection of SIMILAR *and* OPPOSITE response options, but the accuracy of the selection depended upon whether the participant was male or female and whether the trial had been designated as consistent or inconsistent.

### *The IRAP Sequence*

Each complete IRAP comprised of a total of eight blocks of trials – two practice blocks and six test blocks always presented in that order. The block sequence was divided into consistent trials (four blocks) and inconsistent trials (four blocks). Hence, one practice block and three test blocks presented consistent trials only, while the other practice block and the other three test blocks presented inconsistent trials only (i.e. consistent and inconsistent trials never appeared in the same block). The IRAP sequence was always presented as alternating blocks of consistent and inconsistent trials. As a result, participants were required to switch their patterns of correct responding across blocks (i.e. the contingencies were reversed). In order to control for potential order effects, the sequencing of the blocks was counterbalanced across participants. That is, half of the participants were presented with a consistent practice block first, followed by the inconsistent practice block, followed by a consistent test block, and so on. In contrast, the other half were presented with an inconsistent practice block first, and so on.

### *General Procedure*

All participants completed the study in a single experimental session that lasted between 30 and 40mins. in total. Although made available, none of the participants opted for short breaks throughout the study.

### *Procedure*

Prior to exposure to the IRAP, all participants were presented with a lengthy instructional phase that: (1) reassured them that they would not be required to do anything they did not understand; (2) ensured that they clearly understood all of the words contained within the IRAP; and (3) ensured that they understood *what* they would be required to do. The participants were first informed that they were going to help to test a new game that would appear on a laptop computer.

The first set of instructions to which participants were exposed were explicitly designed to ensure that they understood the meanings of all of the experimental words. For example, the Researcher verbally presented the following instructions, to a male participant

This game involves you having to match certain words together. So, if I show you the word BOY (participant is shown the two sample stimuli) because you are a boy, then you might have to match that word with one of these words (participant is shown the target stimuli), such as ACCEPTED. But I need to make sure that you know what all these words mean. So, if you had to match BOY with ACCEPTED, what do you think ACCEPTED means?

Participants were given ample time to respond and were asked to qualify each answer with an example. Consider the following transcript:

Child: People accept me because they like me.

Researcher: Yes, it's a bit like somebody in your class who might annoy you sometimes, because they have a hard time sitting down all the time. But that's ok because

they are a nice person and you have fun with them. So you accept them even if they are a little silly sometimes.

The same procedure was then followed for each of the six target stimuli.

Once the correct meanings of the target stimuli had been determined, the participants were presented with specific instructions regarding the matching of the sample and target stimuli. For example, the following instructions were presented to a female participant:

Sometimes the computer will want you to match the word GIRL with this set of words (participant is shown the three positive target stimuli), and sometimes it will want you to match the word BOY to this set of words (participant is shown the three negative target stimuli). So, you might have to match GIRL with ACCEPTED and BOY with FAULTY.

Then the computer will change its mind and it will want you to match the word GIRL with this set of words (participant is shown the three negative target stimuli) and the word BOY with this set of words (participant is shown the three positive target stimuli). So, now you might have to match GIRL with FAULTY and BOY with ACCEPTED.

In order to match the words in this game you have to press either the 'd' or 'k' button on the keyboard. Can you show me where they are? They are the only two keys you have to press. You do not have to press anything else. Ok?

Now, there are two other words involved in this game -- OPPOSITE and SIMILAR. So if I tell you that I want you to match the word GIRL to ACCEPTED, POPULAR and PERFECT and you see this coming up on the screen (child is presented with schematic representation of a consistent trial with GIRL) which key will you press? Will you press 'd' or 'k'?

The Child is given ample time to reply to indicate their response and the

Researcher continues:

Ok, so what happens if BOY and ACCEPTED come up on the screen but I have told you that I want you to match GIRL to ACCEPTED? What key would you press? Would you press 'd' or 'k'?

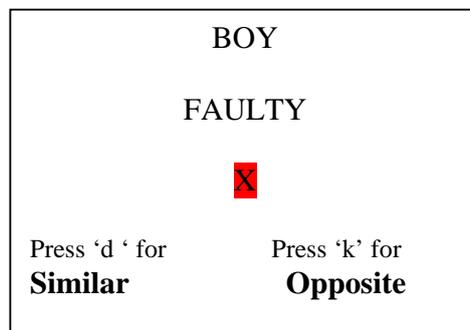
Again the child indicates their response and the Researcher continues:

Now, like all games there is a little trick involved between the words OPPOSITE and SIMILAR. Sometimes the word OPPOSITE is on this side of the screen (child is shown schematic representation of OPPOSITE on the left-hand side of the screen) and sometimes it is over on the other side of the screen (child is shown schematic representation of OPPOSITE on the right hand side of the screen). It's also the same for the word SIMILAR. Sometimes it is on this side of the screen (pointing to left) and then it changes to the other side of the screen (pointing to right). So the trick is that you have to keep your eye on which side of the screen OPPOSITE and SIMILAR are on, because

remember, you have to press 'd' or 'k' to match the words. So your job in this game is to keep your eye on the word at the top (i.e. GIRL or BOY), the word in the middle (e.g. ACCEPTED) and the places in which OPPOSITE and SIMILAR appear on the screen. Do you understand?

### *Feedback*

Although, the children had been instructed in all aspects of the IRAP procedure at this point, they had not completed any actual trials and thus had not yet been exposed to the feedback mechanism contained within the program. Although the IRAP is referred to as a test, all incorrect trials (but not correct trials) are consequted with written automatic feedback that indicates that a wrong response has been emitted. This feedback involves the presentation of a red X that appeared automatically in the middle of the screen (see Figure 4). The feedback remained on the screen until the correct response is emitted.



---

Figure 4: Visual aid containing a schematic representation of feedback for an incorrect response on an IRAP trial.

---

In order to instruct the children effectively on the feedback, a visual representation of feedback was presented along with the following instructions:

Ok. Like all games, you will sometimes match the words wrong and that's ok. This game is not about getting it right all of the time. Although it's important to get as many words

right as you can, the trick is that you have to match the words as fast as you can while trying to get them right. Ok?

If you get one wrong, the computer will show you this (child is presented with the schematic representation of feedback) and the only way you can go forward in the game is to press the correct key. So the computer will actually tell you when you have made a mistake, ok?

*Practice Trials.* Although the format of all eight blocks was identical, the children were given specific instructions initially concerning whether a block was a practice block or whether it was a test block and the instructions explicitly depended upon whether the first block of practice trials was consistent or inconsistent. The instructions provided during the practice blocks were specifically targeted towards teaching the child: (1) how to press the appropriate keys and (2) how to match the sample and target stimuli appropriately. The instructions provided to a male child exposed to consistent trials first, for example, were as follows:

In order for you to learn how to play the game, we are going to do some practice. For this first practice I want you to match BOY to ACCEPTED, PERFECT and POPLULAR, and GIRL to FAULTY, BROKEN and USELESS. I will help you through the first few matching tasks. Ok?

Can you sit in close to the computer and make sure you are comfortable? Good, now show me where you are going to put your fingers. Remember for this game you only need to press the 'd' and 'k' keys on the keypad. When you are ready, just press the space bar (Researcher indicated location of spacebar) and the game will start.

The first exposure within the first practice block was then presented. During the first few practice trials, the Researcher prompted the child by indicating the correct response option and the correct key associated with it. Prompting for correct responding continued until each child was familiar with correct responding (this rarely exceeded five trials). If a child emitted an incorrect response during the initial practice trials, she/he was instructed as follows:

That's ok. Remember that I told you it's ok to get some wrong. All you have to do now is to press the correct key in order to go forward with the game. So, which key will you press?

Throughout the first practice block, the Researcher also intermittently provided verbal feedback on correct and incorrect responding.

At the end of each block of trials, the IRAP presented participants with automated feedback on the percentage of trials correct and the median response time (in ms.) achieved during that block. Because of the participants' ages, the Researcher provided further clarification on what the automated scores meant. For example, if a child's median reaction time on the first practice block exceeded 3000ms. s/he was instructed as follows:

Ok ..... (Child's name). That was very good. You seem to be getting the hang of this game. There is one other thing I need to tell you. In order to get to the next level of the game you need to try and keep your score under 3000ms. Ok? And I still want you to try to get as many right as you can. Ok?

Hence, irrespective of performance, all children then proceeded immediately with the second practice block.

Because the IRAP systematically presented blocks that alternated between consistent and inconsistent trials, the children required explicit instructions after the first practice block regarding the change in trials in the subsequent block. For instance, in the examples provided above the instructions concerned consistent responding and thus the children then required further instructions regarding the inconsistent trials presented in the subsequent practice block. Hence, the children were explicitly instructed that the second practice block of trials would involve changing their responses as follows:

Ok, now do you remember that I told you that the computer sometimes changes its mind? This time it wants you to match BOY to FAULTY, BROKEN and USELESS and GIRL to ACCEPTED, PERFECT and POPULAR. This is still only a practice just like the last time. But I want you to try and go as fast as you can, while still trying to get as many right as you can. Ok?

When you are comfortable and ready to go press the space bar. The first set of words will come up on the screen.

The instructional and presentation formats that accompanied the second block of trials were identical to the first, except that the children were instructed to emit alternative response patterns.

### *Test Trials*

After the presentation of the automated feedback following the second practice block of trials, the children were invited to take a short break prior to the first block of test trials. The first test block then commenced with the following verbal instructions:

Ok, it's obvious that you have caught onto this game really quickly. So I think you are ready to go on to the next stage. We are going to do this six more times, but as you can see you get through the stages really quickly and we will count each task together when you have finished. Ok?

Now, this time I am not going to tell you which key to press. Ok? So, that means that the practice is now over and the computer is going to test you to see how much you remember. Ok? Before each matching task I will just check with you to see if you remember which words you are matching to BOY and which words you are matching to GIRL. Is that ok?

Now, for the next ones can you tell me which words you have to match BOY to? (Researcher used the word lists to prompt the child and waited for a response). Now can you tell me which words you have to match GIRL to? (Again, Researcher waited for the child's response).

So, are you comfortable and ready to go? Make sure you have your fingers on the right keys. Remember -- try and go as fast as you can but still trying to get them right. Ok, off you go and press the space bar to begin.

Although the children were provided with these specific instructions at the beginning of each block regarding the type of responding appropriate for that block (i.e. inconsistent or consistent – but these actual words were not employed), they were not prompted throughout the block as had been the case with the practice trials. However, the Researcher did remain seated beside each child throughout all trials and continued to

provide positive verbal encouragement regarding appropriate on-task behaviour (e.g. “you are doing well and working really fast”), but not for accurate or inaccurate responding.

At the end of the sixth and final test block, the Researcher indicated to the children that the game was over and thanked them for their participation. Each child was then given a packet of gel pens as a reward for taking part in the study.

## RESULTS

The IRAP software records levels of accuracy and response latency for each participant on every trial. *Accuracy* is defined as the first response emitted on each trial. Hence, even if a subsequent accurate response is emitted on the same trial (because every trial incorporates a correction procedure for incorrect responding), the trial is recorded automatically as incorrect. *Response latency* is defined as the time (in ms.) between the onset of the trial and the emission of a correct response. The accuracy and response latency data are described separately below.

### *Accuracy Data*

In previous IRAP research (Barnes-Holmes et al., 2006), the accuracy data have not been incorporated into the analyses, but are used as a screening procedure to remove participants whose overall accuracy levels in each block of trials are lower than 70% correct. This criterion was applied in the current research. In the current study all participants' accuracy levels were above 70%.

### *Response Latency Data*

Because there is no fixed time to respond on IRAP trials, considerable variation both within and across participants is to be expected. In order to control for this variability, Greenwald et al.'s (2004) C4 scoring algorithm was employed (see also Barnes-Holmes et al., 2006). Specifically, latencies over 3000ms. are adjusted by the program and recorded as 3000ms. and latencies less than 300ms. are recorded as 300ms.

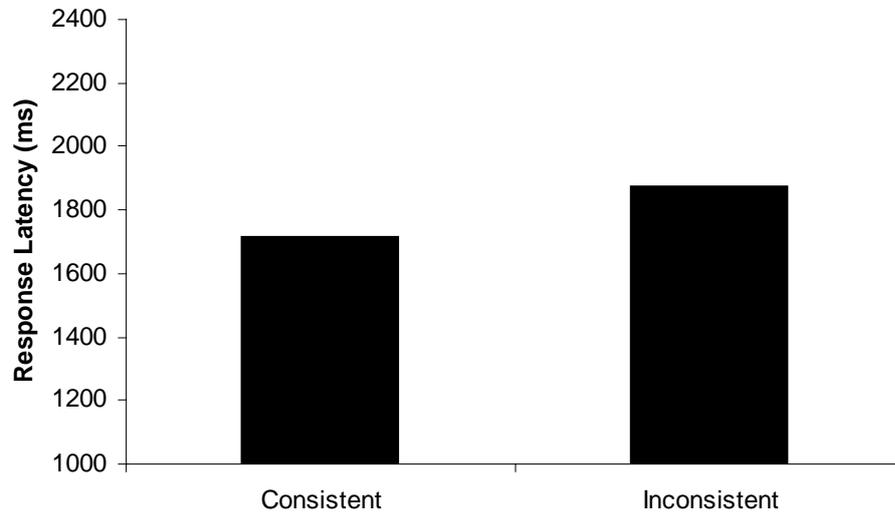
For the purposes of analysis, mean response latencies were calculated (after adjustments) for each participant on each of the six test blocks. This calculation generated three mean latencies for consistent responding and three for inconsistent responding. Table 2 presents the mean latency data (for consistent and inconsistent responding across the six test blocks). The table indicates that all participants responded considerably faster on consistent trials relative to inconsistent trials, thus suggesting predicted IRAP effects.

TABLE 2  
Mean Response Latencies (in ms.) for Consistent and Inconsistent Responding across Test Blocks.

Blocks 1 and 2		Blocks 3 and 4		Blocks 5 and 6	
Consistent	Inconsistent	Consistent	Inconsistent	Consistent	Inconsistent
1738	1937	1706	1844	1708	1841

In order to determine if any of the observed differences were statistically significant, a 2x3x2x2 mixed repeated measures analysis of variance (ANOVA) was conducted, with IRAP condition (consistent versus inconsistent trials) and test sequence (Blocks 1, 2 and 3) as within participant variables, and order (consistent-first versus inconsistent-first) and gender (males versus females) as between participant variables. The results of the analysis revealed a significant main effect for IRAP condition [ $F(1,4) = 12.34, p = .0270, \eta_p^2 = .193$ ]. For illustrative purposes, the IRAP effects are presented in Figure 5, which clearly shows the predicted superiority of consistent over inconsistent responding. The analysis also revealed that the main effects for order ( $p =$

.18), test sequence ( $p = .08$ ) and gender were non-significant. ( $p = .63$ ), as were all interaction effects (all  $p$ 's  $> .19$ ).



---

Figure 5: Overall mean response latencies for consistent and inconsistent responding.

---

### *Trial-type Analyses*

As noted previously, the IRAP may be conceptualised as four trial-types (OWN GENDER/Positive [e.g. ACCEPTED]; OWN GENDER/Negative [e.g. FAULTY]; OTHER GENDER/Positive; and OTHER GENDER/Negative). Subsequent analyses were conducted to explore the impact of trial-types on responding and to determine if there were any interactions among trial-types and the critical IRAP effects. The mean response latencies for all participants on consistent and inconsistent trials for each of the four trial-types are presented in Table 3.

TABLE 3  
The Mean Response Latencies (+SE) in ms. For Consistent and Inconsistent  
Responding on Each Trial-type.

	POSITIVE		NEGATIVE	
	Consistent	Inconsistent	Consistent	Inconsistent
<b>OWN GENDER</b>	1906 (720)	1991 (567)	1905 (617)	2261 (617)
<b>OTHER GENDER</b>	1866 (569)	2015 (633)	2036 (653)	2076 (455)

The table indicates few differences in responding across trial-types, with no notable differences between responding to OWN GENDER vs. OTHER GENDER or between Positive vs. Negative. Although it seemed unlikely, a 2x2x2 mixed repeated measures ANOVA was conducted with sample, target-type and IRAP condition as repeated measures to determine whether there were any significant trial-type differences. As expected, all of the variables were non-significant -- target ( $p = .85$ ), sample ( $p = .19$ ) and IRAP condition ( $p = .09$ ), and there were no significant interaction effects (all  $p$ 's > .26).

In summary, the results from the mean response latencies recorded in the current study showed a significant overall IRAP effect in the predicted direction, indicating that the children showed a bias towards their own gender and against the other gender. However, there were little or no differences among the four trial-types and no differences between males and females.

#### *The $D_{IRAP}$ Algorithm*

An adaptation of the D-algorithm developed by Greenwald et al, (2003) was used to transform the individual response latencies for each participant. The version of the D-

algorithm employed in the current study (referred to here as the D-IRAP algorithm) transforms the raw latency data for each participant using the following steps: (1) only use response-latency data from test blocks; (2) eliminate latencies above 10,000ms. from the dataset; (3) eliminate the data from a participant for whom more than 10% of test block trials have latencies less than 300ms.; (4) compute twelve standard deviations for the four trial-types -- four for the response-latencies from Blocks 1 and 2, four from the latencies from Blocks 3 and 4, and a further four from Blocks 5 and 6; (5) compute the 24 mean latencies for the four trial-types in each test block; (6) compute difference scores for each of the four trial-types for each pair of test blocks by subtracting the mean latency of the consistent block from the mean latency of the corresponding inconsistent block; (7) divide each difference score by its corresponding standard deviation from step 4, yielding 12 D-IRAP scores -- one score for each trial-type for each pair of test blocks (8) calculate four overall trial-type D-IRAP scores by averaging the three scores for each trial-type across the three pairs of test blocks; (9) calculate a single overall D-IRAP score by averaging the 12 trial-type D-IRAP scores.

The  $D_{IRAP}$ -algorithm can be used without regard to trial-type to calculate a single  $D_{IRAP}$  statistic for each participant as above (1-9), or alternatively it can be used with data segregated according to trial-type to produce four  $D_{IRAP}$  statistics for each participant ( $D_{IRAP-(trial-type)}$ ) -- one for each trial-type (1-7). The first analyses conducted on the D-Scores from Experiment 1 employed the  $D_{IRAP-(trial-type)}$  data. The  $D_{IRAP-(trial-type)}$  scores are presented in Figure 6. The figure first indicates that responding on all trial-types was in the predicted direction. The D scores for the OWN GENDER/Positive, OWN GENDER/Negative and OTHER GENDER/Positive trial-types were all large, whereas

the D score for OTHER GENDER/Negative was considerably smaller. These findings indicated strong positivity (and disagreement with negativity) towards own gender, with some disagreement that OTHER GENDER was positive (but not necessarily negative).

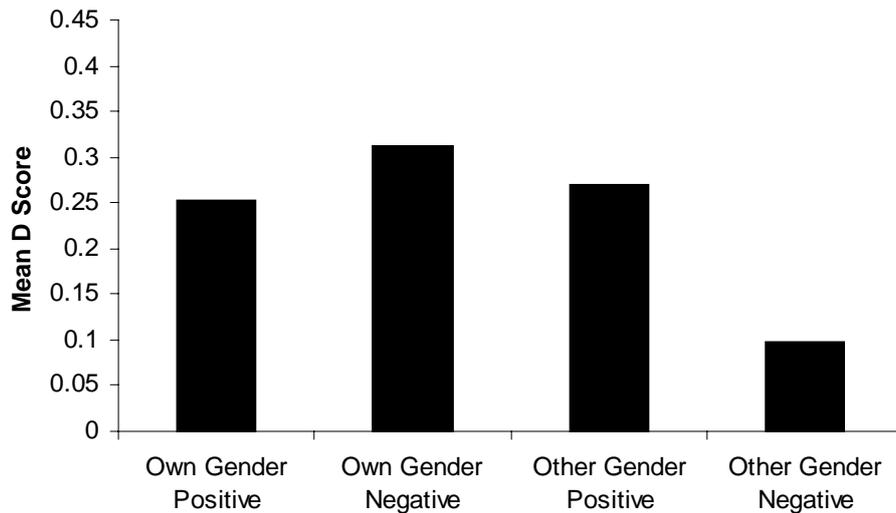


Figure 6: Mean D-IRAP trial-type scores.

A 2x4 repeated measures ANOVA, with IRAP condition and trial-type as within participant variables, revealed no significant main effects for either variable, and no interaction effects (all  $p$ 's > .11). Hence, although there appeared to be some differences in responding to the various trial-types, trial-type did not have a significant influence on the performances and the main IRAP effect was not retained in analyses at this level.

Four separate one-group t-tests were then conducted to determine whether each of the D-IRAP scores for the four trial-types differed significantly from 0. These analyses revealed that only OWN GENDER/Negative was significant [ $t(23) = 2.52$   $p < .0190$ ], although OWN GENDER/Positive [ $t(23) = 1.71$   $p > .0998$ ] and OTHER

GENDER/Positive [ $t(23) = 2.03$   $p > .0535$ ] were approaching significance. As expected, OTHER GENDER/Negative [ $t(23) = .839$   $p > .4103$ ] was not significant.

A single D IRAP statistic was then calculated for each participant and these were subjected to a one-group t-test in order to determine whether the overall D scores (rather than the  $D_{\text{IRAP-(trial-type)}}$ ) differed significantly from 0. Indeed the analysis revealed a significant difference from 0 ( $p < .0274$ ) that supported the original IRAP effect.

## DISCUSSION

The IRAP effects observed in Experiment 1 were consistent with previous IRAP studies (Barnes-Holmes, et al., 2006), but represented the first IRAP outcomes recorded with young children. The similarity of the findings with existing IRAP work with adults indicates the broad utility of the procedure in terms of the simplicity of the task. Indeed, none of the interactions the children in this study had with any aspects of the procedure were negative and their approach to the program could be summarised overall as positive and challenging. Interestingly, there remain little or no IAT studies involving children, so it is difficult to compare the IAT and the IRAP outcomes in this regard.

Although Experiment 1 was primarily concerned with the general utility of the procedure with children, this effect was conducted using the children's positivity or negativity regarding their own gender and the gender of others. As expected, the IRAP indicated that the children overall showed a positivity bias towards their own gender, with males and females almost identical in this regard. Furthermore this own gender positivity was bolstered by predicted negativity towards the opposite gender for both males and females. This implicit gender outcome is entirely consistent with the

developmental literature, which has repeatedly demonstrated that most children have acquired a strong gender identity as early as age three (Santrock, 2001). It is not surprising, therefore, that the children who participated here and who ranged in age between 10 and 14 years old had well-established biases towards their own gender.

Although the utility of the IRAP with children was established by the outcomes in Experiment 1, it might be argued that presenting gender associations even to children of this age is not a particularly good test of the IRAP. For example, one might argue that even a largely insensitive procedure might tap into gender differences because children's gender identity is so well-established. If this is the case, then the next step in a research program involving the IRAP and young children would be to examine a much more subtle or amorphous concept, such as self-esteem. Furthermore, investigations in this latter area might permit an exploration of the relationship between the IRAP and explicit measures (see Chapter 1). Hence, more might be said of the IRAP, especially with children, if the methodology could be used successfully to measure self-esteem in children, to correlate with explicit self-esteem measures, and/or to differentiate children who vary along this dimension. This issue was the focus of Experiment 2.

## Chapter 3

### Examining the Utility of the IRAP in the Study of Gender and Self-Concept with Normally-Developing and Dyslexic Children

## **Chapter 3: Experiment 2**

### **Examining the Utility of the IRAP in the Study of Gender and Self-Concept with Normally-Developing and Dyslexic Children**

#### **INTRODUCTION**

The IRAP effects observed in Experiment 1 were consistent with previous studies (Barnes-Holmes et al., 2006), but represented the first IRAP outcomes recorded with children. The similarity of the findings with existing IRAP work with adults indicated the broad utility of the procedure both in terms of the simplicity of the task and as a means of targeting implicit relations. As expected, the children overall showed a positivity bias towards their own gender and negativity towards the opposite gender that is consistent with their current developmental stage (Santrock, 2001). As noted previously, however, gender is a critical feature of one's self-concept and indeed the design of Experiment 1 explicitly targeted gender attitudes in the context of positive and negative evaluations towards the self and others. In other words, the findings may have reflected the children's implicit attitudes to themselves as much as their attitudes to gender. Put simply, the children in that study were demonstrating positivity not only to their own gender, but to their self-concept, while demonstrating generic negativity towards others. Because of this possibility, Experiment 2 attempted to compare two groups of children who might differ in terms of self-concept based on the strength or weakness of their academic attainment (a critical feature of self-concept). In order to contextualise the current study, the

literature describing explicit and implicit outcomes of self-concept is reviewed in the sections that follow.

### *The Self*

A range of terms have been employed by psychologists in their attempts to understand the 'self', including; 'self-concept', 'self-esteem', 'self-image', 'self-perceptions' and 'self-worth' (Humphrey, 2002). In a nutshell, one's understanding of oneself appears to comprise two key components, namely the *content* and *valence* of one's knowledge. The content is frequently referred to as *self-concept* and includes largely descriptive questions such as "What am I like?" *Valence*, on the other hand (commonly referred to self-esteem), assigns affective and/or evaluative appraisals to the self with questions such as "Do I like who I am?" (Baumeister, 1993, 1998; Wenar & Kerig, 2005). Indeed, there is general consensus that individuals aspire towards positive self-evaluations and show considerable reticence towards self-attributions that are perceived as negative (Deraedt, Schhacht, Frack, & de Houwer, in press). Not surprisingly, some researchers have highlighted the combination and/or interaction of these components in the composition of the self (Blascovich & Tomaka, 1991). For example, James (1890) argued that both the content and valence of one's self-perceptions are guided by comparisons between the actual self and an ideal self, as well as between the self and others. In cases where one's attribution to the actual self, ideal self and others are largely positive and similar, then the underlying implication is that one will be accepted rather than rejected (Heinonen, Raikkonen, & Keltikangas-Jarvinen, 2005). However, where the difference is considerable with stronger negativity towards the self than the ideal self or others, the underlying

implication is that one will be rejected rather than accepted (Heinonen, Raikkonen, & Keltikangas-Jarvinen, 2005).

### *The Development of the Self*

The development of the self appears to run consistently in parallel with an understanding of others (i.e. understanding what is and is not the self). The earliest signs of *a sense of self* appear to emerge in infancy, when toddlers begin to recognise themselves visually and can readily differentiate themselves from others (Rathus, 2006). As knowledge and experience of the self and others becomes more extensive, evaluations emerge. These are initially characterised in early to middle childhood by unrealistic positive self-regard (e.g. “I know other kids are bad at some things but not me”). However, these grandiose self-perceptions are more indicative of developmental patterns of inflexible thinking that are equally attributable to opinions of others. For example, a young girl whose learning disability results in low academic achievement in a particular area will conclude that she is dumb and would make a similar conclusion about another child in a similar predicament (Wenar & Kerig, 2005).

Although a more composite sense of self and others emerges around eight to eleven years of age, these attributions are primarily of an external, rather than internal, nature. Children at this age, for example, are popular because they are extroverted or academically talented. In contrast, attributions regarding psychological traits do not emerge until late childhood or early teens when personality characteristics begin to demonstrate both stability and integration (Wenar & Kerig, 2005). At this developmental stage, a global sense of self comprises an array of competencies that are academic, physical, social and even psychological (Harter, 1990). As attributions

become increasingly based on psychological traits, it becomes more likely that one's sense of self incorporates weaknesses and negativity than was previously the case (Wiegfield & Karpathian, 1991). These changes are consistent with changes to parenting style (i.e. greater parental willingness to discipline and criticise) and increased academic experiences (that will involve weaknesses relative to others in one area or another). The considerable fluctuations in self-esteem across the teenage years, therefore, are not surprising (Harter, 1990b).

### *The Relationship between the Self and Academic Attainment*

Although one's sense of self becomes increasingly integrated and psychological in the teenage years, academic attainment continues to be a critical feature in evaluations of the self and others. Indeed, researchers in this area have argued that the relationship between self-esteem and attainment is bi-directional (Lawrence, 1996). That is, perceived academic weakness results in the attribution of negative self-evaluations, while self-esteem in turn exerts control over academic attainment. It is not surprising, therefore, that educational difficulties are associated with maladaptive self-referential styles and low self-esteem (Elbaum & Vaughan, 2001; Gurney, 1988; Humphrey, 2001).

For children, the varied challenges of the educational environment require some ego resilience, especially in the mastery of new concepts and tasks (Wenar & Kerig, 2006). This is, of course, amplified by the inherently social nature of education, where peers can accurately identify who is best or worst in the class at particular subjects. For normally-developing children, however, the challenges are met with relative ease because weaknesses in one area are frequently offset by strengths in another (e.g. a young boy weak at maths may be the best soccer player in

the playground). As a result, the self-esteem turbulences of normally-developing teenagers end in the reasonably integrated and positive self-concept that characterises the majority of young adults (Erikson, 1968).

Where academic deficits are greater and cannot be compensated by additional skills, the challenges may be simply too great, with only the most unusual child possessing the necessary ego resilience to guard off low self-esteem in these circumstances. Such an outcome, of course, typifies the experiences of children diagnosed under the broad spectrum of Special Educational Needs (SEN), whose educational and self trajectory frequently do not match the normal developmental sequence, thus likely impacting negatively on self-esteem.

### *Special Educational Needs (SEN) and Self-esteem*

Definitions of SEN are based on individual deviation from normal expectation. Thus, children are deemed to have ‘special educational needs’ if they require substantial additional education provision in comparison to peers (Frederickson & Cline, 2002). The label SEN, in practical terms, is of greater educational than intellectual significance because it is designed to speak directly to the necessary educational provision that will allow an individual child to reach full potential. Nonetheless, the term, almost by definition, implies some level of cognitive impairment. Hence, the broad category of SEN comprises an array of ‘diagnostic categories’, including physical disabilities (e.g. cerebral palsy), general or specific learning disabilities (e.g. dyslexia) and emotional and behavioural disturbance (e.g. Attention Deficit and Hyperactivity Disorder: ADHD). For educational reasons, these are frequently further sub-divided into mild, moderate and severe, where a child’s

diagnosis dictates the place on a continuum of necessary resources (Griffin & Shevlin, 2007).

There is general consensus that children with SEN develop maladaptive self-referential styles, which in turn manifest altered self-concept and low self-esteem, with considerable evidence that peer rejection plays a pivotal role (Gurney, 1988; Humphrey, 2001). For example, in a large-scale study children with SEN explicitly reported that other children differentiated them on the basis of their disabilities and disliked them accordingly (Wade & Moore, 1997). In support of these perceptions, Euade (1999) reported that children with SEN are more likely to be overtly rejected by school peers. Interestingly, a recent study by Scanlon and McGilloway (2006) reported that although teaching professionals clearly observed the importance of peer acceptance to children with SEN in mainstream education, there was a general consensus that peer rejection and intolerance was more extreme with older children with SEN and was also dependent on the type of difficulty presented. For example, children with ADHD (characterised as emotionally and behaviourally disturbed) were reported to be more socially isolated than children with general or specific learning difficulties, such as dyslexia (see also Department of Education and Science, 2001).

Although a key educational initiative has attempted to integrate normally-developing and SEN children in the same educational environments, there are mixed findings about whether this intervention aids or abets the popularity and self-esteem of the latter group. For example, there is some evidence to suggest that children with SEN in integrated classes spend less time working alone and more time with normally-developing peers (Hunt, Alwell, Farron-Davis, & Goets, 1994). Furthermore, integration does appear to foster greater tolerance by normally-developing children of their SEN classmates (Helmstette, Peck, & Giangreco, 1994).

However, integration per se may hamper these positive outcomes as children with SEN continue to be removed periodically from classes for additional support. As a result, this frequent intrusion may in fact serve as a constant reminder of the differences between the two groups of children (Scanlon & McGiloway, 2006).

### *Measuring Self-esteem*

The difficulties inherent in defining the self have contributed in no small way to problems in measurement (Farnham, Greenham, & Banaji, 1999). The traditional and more common means of assessing the self involves the use of explicit methodologies, in which respondents either indicate how they feel about themselves globally (Rosenberg, 1989) or rate themselves on a variety of specific traits (Marsh, 1986; Pelham & Swann, 1989).

The *Rosenberg Self-Esteem Scale* (SES; Rosenberg, 1965) was originally designed to measure adolescents' global feelings of self-worth or self-acceptance. The scale consists of 10 items in which respondents report feelings about the self directly (e.g. "I feel that I have a number of good qualities"). The measure is scored in terms of respondents' overall level of agreement or disagreement with the statements, resulting in a scale range of 10-40. Higher scores are reported to represent higher self-esteem, with lower scores representing lower self-esteem (Rosenberg, 1989).

The *Self-esteem Inventory* is another explicit measure of self-esteem developed explicitly for use with children (Coppersmith, 1967). This tool consists of items designed to assess self-regard in four specific areas: peers, parents, school and personal interests (e.g. "I'm pretty sure of myself"). Respondents are instructed to indicate the degree of applicability in terms of whether the statement is "like me" or

“unlike me”. Similarity overall generates a high score, while dissimilarity generates a low score, with an overall range from 0 to 50. Again, higher scores reflect strong self-regard and lower scores reflect weaker self-regard.

*The Self-Attributes Questionnaire* (SAQ; Pelham & Swann, 1989) has also been used widely as an explicit measure of self in children. It consists of ten valenced attributes such as intellectual capability, physical attractiveness, leadership ability and emotional stability. For each attribute respondents rate themselves from 1 to 10 in relation to peers. Once again, higher scores reflect higher self-esteem.

As noted previously however (see Chapter 1), explicit measures show strong susceptibility to demand characteristics, social desirability and self-presentation and it is not surprising that many of these variables apply particularly when individuals are asked to identify how they feel about themselves. The putative influence of self-presentation, in particular, commonly yields an assessment of self-esteem that is higher than what is actually the case (Brown, 1991). According to Lindeman and Verkasalo (1995), this skewed outcome results from the influence of two problematic elements of self-presentation. Specifically, *impression management* (i.e. self-presentation to others) obstructs the accurate measurement of self-esteem, while *self-deception* (self-presentation to the self) generates inaccurately high self-assessments (Paulhus, 1986). Furthermore, researchers have also argued that explicit assessments provide an inaccurate account of the self because they fail to take into account the subtle and complex interplay between positive and negative evaluations (Sackheim & Gur, 1978). As a result, explicit outcomes generally exacerbate positive attitudes, while concealing the negative.

### *Implicit Self-esteem*

The use of implicit methodologies to circumvent susceptibilities inherent in explicit measures is a burgeoning area of research that has already been discussed in detail (see Chapter 1). Hence, it is not surprising that several authors have attempted to employ implicit tools in the study of the self (de Jong, 2002; de Jong et al., 2001; Gemar et al., 2001; Teachman, et al., 2001). Indeed, a large body of research suggests that implicit self-esteem is linked to a number of important physical and emotional outcomes (de Hart & Pelham, 2006). For example, the combination of high explicit and low implicit self-esteem has been associated with greater defensiveness (Bosson, Brwon, Zeigler-Hill, & Swann, 2003; Jordan, Spencer, Zanna, & Hoshino-Browne et al., 2003) and poor physical health in response to positive life events (Shimizu & Pelham, 2004). This combination is also believed to predict non-verbal anxiety in a self-threatening interview (Spalding & Hardin, 1999).

In perhaps the earliest attempt to capture the implicit self, Morgan and Murray (1935) employed a projective measure known as the *Thematic Apperception Test* (TAT). The TAT is based on the assumption that simply presenting a series of pictures and encouraging participants to tell stories about them reveals significant components of the personality. An adapted version of the TAT was used by Spence (1974) to measure women's "motive to avoid success", where respondents were given a verbal rather than pictorial cue. Male and female participants, for example, were each presented with one of three short stories (i.e. a successful medical student who was a single female, a single male, or a married female). Participants were then required to answer an objective questionnaire about the figure in the story with which they were provided. The results indicated highly significant differences in participants' reactions to the figures, such that they were more positive towards a

single female than the married female as a medical student. The author argued that the TAT offers a more complete set of attitudes, expectations and personality characteristics than explicit measures.

A later projective measure employed by Rogers, Kuiper and Kirker (1977) involved *Trait Self-Descriptiveness Judgements* provided under laboratory conditions. This approach was based on Craik and Lockhart's (1972) theory, which proposed that words that have been deeply coded during a task should be recalled better than words with shallow coding. This study compared self-reference to several other encoding processes in an effort to determine the relative degree of richness and embellishment that self-reference imparts to the encoding of adjectives. Using an incidental recall paradigm, participants were required to rate 40 adjectives in conjunction with four rating tasks (a structural task, a phonemic task, a semantic task and a self-reference task). The adjectives were chosen to represent a broad spectrum of possible characteristics (e.g. shy, outgoing, etc) and participants indicated whether or not any of these characteristics applied to the adjective. The results revealed that the self-reference task generated a stronger trace than the semantic task. Thus, the authors suggested that self-reference functions as a powerful coding device.

However, as well as limited empirical evidence in both cases, other authors have suggested limitations in the extent to which these tools can measure individual differences because they have been used mostly to examine aggregated effects, either in the form of pre-existing differences between groups or in the form of effects of experimental manipulations (Greenwald & Farnham, 2000).

*The Supraliminal Attitude-prime Task* is another projective measure that has been used to study the implicit self (Hetts, Sakkuma, & Pelham, 1999). The target outcome of this tool is described as the 'automatic attitude activation effect', which is

based on the assumption that when an attitude object is encountered, the evaluation (e.g. good) associated with it is activated spontaneously and without conscious effort (Bargh, Chaiken, Grewer, & Pratto, 1992). As a result, participants are primed with self-relevant words (e.g. I and WE) and the ease with which they can identify positive versus negative stimuli is then assessed.

*The Name Letter Preference Task (NLPT)* is a more recent projective task for assessing self-esteem. This tool is based on the assumption that individuals assign value to objects that are closely associated with the self and which will automatically activate attitudes that provide an indirect assessment of the self (Bosson, Swann, & Pennebaker, 2000). One study by indicated that participants over-evaluate the letters contained within their own names, thus indicating high self-regard (Koole, Dijksterhuis, & van Knippenber, 2001).

*Stimulus Equivalence.* Although the stimulus equivalence paradigm is not traditionally categorised as an implicit methodology, the emergence of derived stimulus relations may be viewed as an indirect measure of the target concepts and their relationship (see Chapter 1). One relevant study within this tradition and using the MTS methodology was reported by Barnes, Lawlor and Smeets (1996). These researchers investigated the academic self-concept of normally-developing children versus those with SEN. Specifically, the children were required to derive relations of co-ordination (or equivalence) among their own names and words associated with academic self-concept (i.e. SLOW and ABLE). The results of the study indicated that the children with SEN were less likely than their normally-developing counterparts to derive the OWN NAME-ABLE relation and more likely to derive OWN NAME-SLOW.

### *The IAT as a Measure of Implicit Self-esteem*

Only a limited number of studies have reported the use of the IAT as a measure of implicit self and none of these involved young participants. In one study, participants were asked to imagine that they were a member of one of two fictitious four-person groups (Farnham, 1999). On the IAT, participants were presented with names associated with each of the four groups and pleasant versus unpleasant attributes. They were also presented with another IAT as a measure of group identification, in which the self was associated with some of the group names. The results of the study indicated that the fictitious membership group had implicitly become an in-group that was more closely associated with the self than alternative groups. Furthermore, weak positive correlations were reported between the IAT and explicit self-esteem measures (average  $r = .18$ ). In a similar study, Haines (1999) reported that women assigned to a powerful role in a simulation game increased the association between the self and the attribute of dominance. Greenwald and Farnham (2000) also investigated associations of the self versus others in the context of Pleasant and Unpleasant attributes. Consistent with predictions, the IAT outcomes indicated strong self-positivity that was positively (but weakly) correlated with the RSEQ and SAQ (average  $r = .17$ ).

Bosson, Swann and Pennebaker (2000) have reported the most complete data set for implicit self-esteem on the IAT. This study incorporated no less than seven implicit measures, including the Stroop Colour Naming Task and the Ambiguous Statements Task. The explicit measures used included the Rosenberg Self-Esteem Scale (1965), the Self-Liking and Competence Scale (Taforodi and Swann 1995), the Self-Attribute Questionnaire (Pelham & Swann 1989), and the Writing Task (Pennebaker, 1997). Unexpectedly, the seven measures of implicit self-esteem were

not positively correlated with one another, and they correlated only weakly with the explicit measures of self-esteem. In particular they found that the IAT weakly correlated with explicit self-esteem ( $r = .21$ ) and also had the highest test-retest reliability ( $r = .69$ ).

According to Karpinski (2004), however, the IAT is more than a measure of self-esteem because it also targets associations involving the Self in direct comparison with associations involving Others. Hence, this latter feature may be critical to interpretations of the outcomes. In order to investigate the role of Other in IAT effects, Karpinski conducted two studies in which Other was and was not specified. Specifically, in one study Other was unspecified, while in the other study Other was specified as “close” (e.g. a best friend, boyfriend, or girlfriend). The researcher predicted that participants’ performances regarding the implicit Self associations would differ across these two contexts and indeed this was the case. Specifically, the overall IAT self-esteem scores were significantly higher in the context of ‘close Other’ relative to ‘unspecified Other’. Karpinsky accounted for these differential findings by suggesting that the IAT with ‘close Other’ was a better reflection of the self, than when close others are not specified. The researcher also argued that the data highlighted the contextual sensitivities of the methodology (see also Blair, Ma, & Lenton, 2001; Dasgupta & Greenwald, 2001).

The complexity of the whole concept of self-esteem and the inherent difficulties in its measurement have been highlighted by Greenwald and Farnham (2000) who argued that self-esteem and self-concept are inextricably bound and their measurement is almost inseparable from distortions in self-presentation (Greenwald & Farnham, 2000). Hence, low correlations observed between the IAT and explicit

measures of self- will frequently fail to correlate because of the potential breadth of what can be measured.

### *The Current Research*

Experiment 2 attempted to extend the preliminary IRAP research reported in the previous study by enhancing the associations between gender and self. That is, a new sample of children here were presented with an IRAP that contained their own names juxtaposed with a name readily associated with the opposite gender (rather than the words BOY and GIRL). The name of this modification was primarily to enable the IRAP to tap more directly into implicit attitudes to self, rather than gender. Furthermore, a sub-set of children diagnosed with SEN (and specifically with Dyslexia) were included in order to facilitate comparisons with the normally-developing counterparts, with regard to self. The study also examined potential correlations between implicit attitudes to self and the Piers-Harris Self-concept Scale (2).

## METHOD

### *Participants*

Twenty children were recruited for participation in Experiment 2. They ranged in age from 9 to 14 years (mean age was 10 years and 2 months). The participants were recruited by the Researcher through direct contacts with local Primary and Post-Primary schools in the Kildare region and parental consent was obtained in all cases. The participants were divided into two groups that were differentiated in terms of their levels of intellectual development. The ‘Normally-Developing’ group comprised of ten participants (9 females and 1 male) who were within the ‘normal’ range of intellectual development and thus had no prior evidence of a learning disability. The ‘Dyslexic’ Group (6 females and 4 males) had a formal diagnosis of Dyslexia by an independent agent. Dyslexia maybe defined as

“a continuum of specific learning difficulties related to the acquisition of basic skills in reading, spelling and/or writing, such difficulties being unexpected in relation to an individual’s other abilities and educational experiences”

(Report of the Task Force on Dyslexia, p. 7, 2002)

However, it is important to note that Dyslexia is not normally associated with generic deficits in IQ.

### *Setting*

All aspects of the setting employed in Experiment 2 were identical to Experiment 1.

### *Apparatus and Materials*

Experiment 2 involved two basic sets of materials/apparatus: the *explicit* measure which comprised of The Piers-Harris Children’s Self-Concept Scale 2

(referred to as The Piers-Harris 2) and the *implicit* measure that comprised of the IRAP.

*The Piers-Harris 2.* The Piers-Harris 2 (see Appendix A) is a self-report instrument for the assessment of self-concept in children. The measure comprises a 60-item questionnaire entitled “*The Way I Feel About Myself*” (Piers & Herzberg, 2003) that is sub-divided in terms of six domains of self-concept. These include: Behavioural Adjustment (BEH: 14 questions); Intellectual and School Status (INT: 16 questions); Physical Appearance and Attributes (PHY: 11 questions); Freedom from Anxiety (FRE: 14 questions); Popularity (POP: 12 questions); and Happiness and Satisfaction (HAP: 10 questions). Responding to the Piers-Harris 2 simply involved circling a YES/NO answer in terms of whether each statement applied to the respondent. These answers generate an overall TOT (T) score for self-concept, which can be sub-divided according to the six domain scores. In general, a higher TOT score indicates a higher self-concept and the categories are arranged as follows: a T-score range of 30-39 is defined as LOW self-concept; 45-55 is AVERAGE; and 60-69 is HIGH.

The validity of responding within each domain is also assessed in terms of Inconsistent Responding (INC) and Response Bias (RES). Inconsistent Responding is defined as a lack of correspondence among items that are similar in nature (e.g. “My classmates make fun of me” and “People pick on me”). Response Bias refers to a participant’s tendency to agree or disagree with test items irrespective of content (i.e. the number of YES responses recorded by the participant overall, thereby suggesting that participants persist with a particular answer that does not match specific questions). Mild to moderate levels of Response Bias in either direction typically do not interfere with the interpretation of the Piers-Harris 2. However, if the level of

Response Bias is 40 or above ( $T \geq 70$ ), or 18 or below ( $T \leq 30$ ), further clarification with respondents may be needed to ensure that they fully understood the items. Thus, the number of Inconsistent Responses as well as the Response Bias overall are used to determine the degree of caution needed in evaluating the TOT score (higher levels of Inconsistent Responding reduce the reliability of the TOT).

*The IRAP.* The IRAP employed in Experiment 2 was identical in format to the previous study. The sample stimuli in the current study attempted to target the self, as well as gender, so each participant was presented with his or her own name and a name that characterised a person of the opposite gender. For example, if the participant was a young boy called Martin, the samples were MARTIN and LOLA (where Lola was clearly identifiable as a girl’s name). Similarly, if the participant was a girl called Mary, the samples were MARY and PETER (where Peter was clearly identifiable as a boy’s name). In all cases, the actual names LOLA and PETER were used to identify female and male, respectively.

The target stimuli were identical to Experiment 1 (i.e. three positive terms: ACCEPTED, POPULAR and PERFECT and three negative: FAULTY, BROKEN and USELESS). The same relational terms SIMILAR and OPPOSITE were also employed as response options. The stimulus arrangements are presented in Table 4.

TABLE 4  
The Stimulus Arrangements Employed in Experiment 2.

<b>Sample 1</b> (Participant’s) OWN NAME	<b>Sample 2</b> PETER/LOLA (OPPOSITE GENDER NAME)
<b>Response Option 1</b> SIMILAR	<b>Response Option 2</b> OPPOSITE
<b>Targets deemed consistent with Sample 1</b>	<b>Targets deemed consistent with Sample 2</b>
ACCEPTED POPULAR PERFECT	FAULTY BROKEN USELESS

The screen shots employed as visual aids in the previous experiment were adapted for current purposes in order to familiarise the children particularly with the use of the name that characterised the opposite gender.

### *Procedure*

All participants completed the study in a single experimental session that lasted between 30 and 40mins. in total. Although made available to them, none of the children opted for short breaks throughout the study. All participants were presented with the same experimental sequence that involved exposure to the explicit measure followed after a short break by the IRAP.

*Piers-Harris 2.* At the beginning of the study, each participant was invited to sit beside the Researcher, answer some questions about him/herself, and to help test a new game on a laptop computer. Participants were provided with the following verbal instructions regarding completion of the Piers-Harris 2.

This paper contains a lot of questions to find out how children really feel about themselves, so it gives you the chance to say how you feel about different things. This is not a test and when you answer the questions, I would like you to think of how you really are, not how you think you should be. So, there are no right or wrong answers. Ok?

Here are some sentences that tell how children may feel about themselves. You have to read each sentence and decide whether or not it describes the way you feel about yourself. If you like, I can read them along with you. If the sentence is **TRUE** or **MOSTLY TRUE** for you I want you to circle the word **YES** next to the statement. If it is **FALSE** or **MOSTLY FALSE** for you, circle the word **NO**.

Ok, I need you to answer every question, even if some are hard to decide on. Do not circle both **YES** and **NO** for the same sentence. If you want to change your answer, cross it out with an **X** and circle your new answer. Remember that there are no right or wrong answers. Only you can tell me how you feel about yourself. So I hope you will mark each sentence the way you really fell inside. Ok?

At this point, the Researcher ensured that participants clearly understood what was expected of them. Some children needed further clarification about particular statements and how they should be responded to. One such interaction with a normally-developing nine-year old male participant was as follows:

Child read the statement: 'I am nervous' and responded: "Sometimes I am nervous and sometimes I am not".

Researcher: Yes, there are times when I am nervous too. I remember I was really nervous when I had to do a big test in my school. But once I got started I wasn't nervous anymore. Can you tell me what makes you nervous?

Child: Yes, sometimes when I am going to play a football match.

Researcher: Would you be so nervous that it would stop you from playing in the match?

Child: No, I usually forget about it once I am on the pitch.

Researcher: So do you think you are nervous all of the time, or just when you have to do specific things?

Child: No I am not nervous all of the time.

The Researcher remained seated beside each participant throughout all aspects of completing the questionnaire, after which each participant was thanked and informed that it was now time to test the new game.

*The IRAP.* All procedural aspects of the IRAP employed in Experiment 2 were identical to the previous study. Similar to Experiment 1, the children needed extensive instructions regarding their understanding of the IRAP trials and how they should be responded to. Because the instructions of Experiment 2 were almost identical to the previous study (and the fact that the trial-types were very similar), they are not presented in the current Procedure section, but in the interests of clarity they are provided in full in Appendix B.

The four basic trial-types presented here for a male participant were referred to as OWN NAME/POSITIVE (top left of Figure 7); OWN NAME/NEGATIVE (top right); OPPOSITE GENDER NAME/POSITIVE (bottom left); and OPPOSITE GENDER NAME/NEGATIVE (bottom right). Each block of IRAP trials comprised of six exposures to each of the four trial-types which contained two exposures to each of the six target words (e.g. ACCEPTED appeared twice with Own Name).

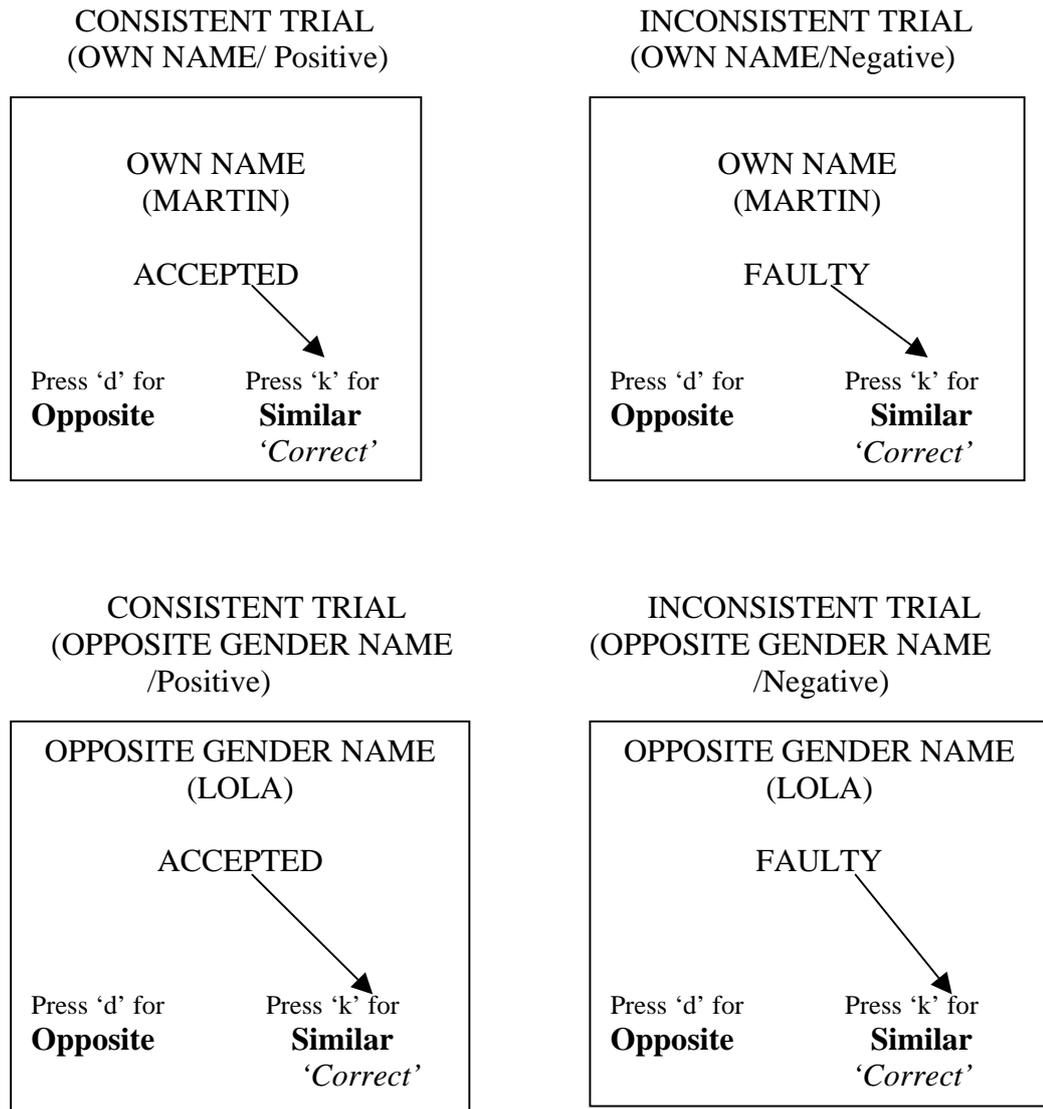


Figure 7: Schematic representation of the four basic trial-types in Experiment 2.

Correct responses on consistent trials involved positivity towards one’s own name and negativity towards the opposite gender name. In contrast, correct responses on inconsistent trials involved negativity towards one’s own name and positivity towards the opposite gender name.

At the end of the sixth and final test block, the Researcher indicated to the children that the game was over and thanked them for their participation. Each child was then given a packet of gel pens as a reward for taking part in the study.

## RESULTS

### *Response Latency Data*

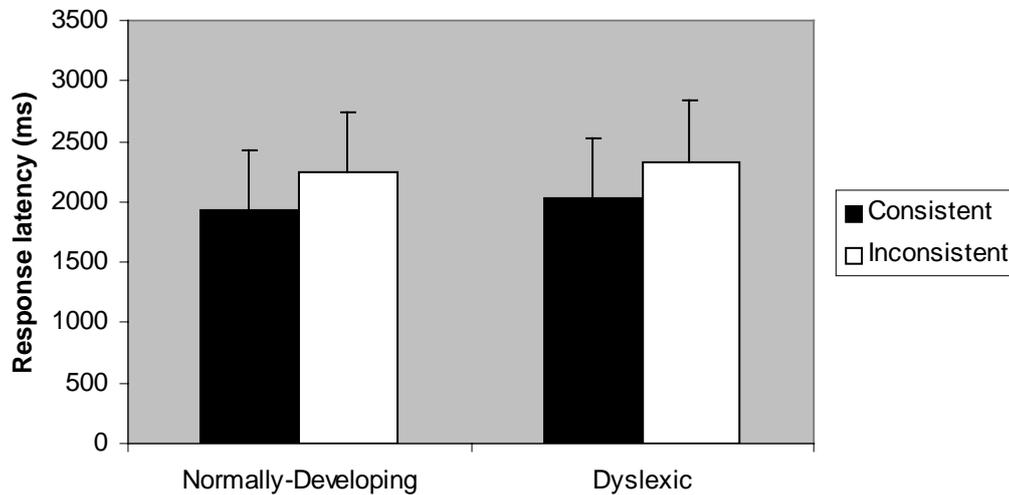
Mean response latencies (three consistent and three inconsistent) were calculated (after adjustments) for each participant on the six test blocks and were then collated by group -- see Table 5. Participants in both groups responded more quickly on consistent than inconsistent trials across the three sets of blocks, thus suggesting an IRAP effect. However, the Normally-Developing children appeared to produce shorter mean latencies overall (on both types of trials), relative to the Dyslexic group.

TABLE 5  
Mean Response Latencies for the Two Groups on Consistent and Inconsistent Responding

	Blocks 1 and 2		Blocks 3 and 4		Blocks 5 and 6	
	Con	Incon	Con	Incon	Con	Incon
Normally-Developing	2031	2064	1905	2136	1842	1891
Dyslexic	2290	2472	2268	2315	2171	2208

A 2x3x2 mixed repeated ANOVA was conducted on the latency data, with IRAP condition and blocks as within participant variables and order (consistent-first versus inconsistent-first) as the between participant variable. Both groups of participants showed significant main effects for IRAP condition -- Normally-Developing [ $F(2, 8) = 23.68, p = .0012, \eta_p^2 = .56$ ] and Dyslexic [ $F(2, 8) = 6.20, p = .0374, \eta_p^2 = .233$ ], and the IRAP effects for both groups are illustrated in Figure 8. No main effects were recorded for order for either group (all  $p$ 's  $> .36$ ). Blocks were non-significant for the Normally-Developing group ( $p > .13$ ), but significant for the Dyslexic group, [ $F(2, 8) = 11.218, p = .0009, \eta_p^2 = .72$ ]. Post hoc tests (Scheffe)

indicated that the significant difference lay between Blocks 1 and 3, probably suggesting practice effects that were particularly strong this group. No significant interaction effects were recorded (all  $p$ 's > .71).



---

Figure 8: Mean Response latencies for consistent and inconsistent responding for the two groups.

---

### *Preliminary Trial Analyses*

As noted previously, the IRAP may be conceptualised in terms of the four trial-types (OWN NAME/POSITIVE; OWN NAME/NEGATIVE; OTHER GENDER NAME/POSITIVE; and OTHER GENDER NAME/NEGATIVE). Trial-type analyses were conducted on the latency data to explore the impact of trial-types on responding and to determine if there were any interactions among trial-type and IRAP effect.

Overall mean response latencies for each participant on consistent and inconsistent trials for each of the four trial-types were calculated and these are presented in Table 6.

TABLE 6  
The Mean Response Latencies for Consistent and Inconsistent Responding By Trial-Type.

Group	OWN NAME /POSITIVE		OWNNAME /NEGATIVE		OTHER GENDER NAME/POSITIVE		OTHER GENDER NAME/NEGATIVE	
	Con	Incon	Con	Incon	Con	Incon	Con	Incon
Normally-Developing	1750	2094	2038	2023	2183	2157	2188	2329
Dyslexic	2342	2851	2306	2692	2473	2548	2686	2626

The table indicates that participants in the Normally-Developing group responded more quickly on all four trial-types relative to the Dyslexic group. Although, the Normally-Developing group responded somewhat faster on OWN NAME/POSITIVE relative to the other three trial-types, both groups showed little or no other distinctions across the latter. For each group, a 2x2x2x2 mixed repeated measures ANOVA was conducted with sample (OWN NAME/OTHER GENDER NAME), target-type (POSITIVE/NEGATIVE), order and IRAP effect as the repeated measures variables. This analysis indicated that only the Dyslexic group showed a significant IRAP effect [ $F(1, 8) = 10.794, p = .0111, \eta_p^2 = .497$ ], the Normally-Developing group did not ( $p > .07$ ). Sample was non-significant for both groups (all  $p$ 's  $> .14$ ). However, target-type was significant for the Normally-Developing group [ $F(1, 8) = 10.36, p = .0122, \eta_p^2 = .109$ ], but not for the Dyslexics ( $p = .75$ ), suggesting differential responding for the former only on POSITIVE versus NEGATIVE trial-types. Order was non-significant for both groups (all  $p$ 's  $> .54$ ). All interactions were non-significant (all  $p$ 's  $> 1.47$ ).

*The D-IRAP Algorithm and Between-Group Comparisons*

Consistent with the analyses conducted on the D scores in Experiment 1, the  $D_{\text{IRAP-(trial-type)}}$  scores were calculated for each participant in Experiment 2. For the purposes of this study both types of  $D_{\text{IRAP}}$ -algorithm analyses were employed. The  $D_{\text{IRAP-(trial-type)}}$  scores are presented in Table 7 and indicate that all scores were in the predicted direction. For both groups, the largest  $D_{\text{IRAP-(trial-type)}}$  score was recorded for OWN NAME/POSITIVE, with the lowest for OTHER GENDER NAME/POSITIVE.

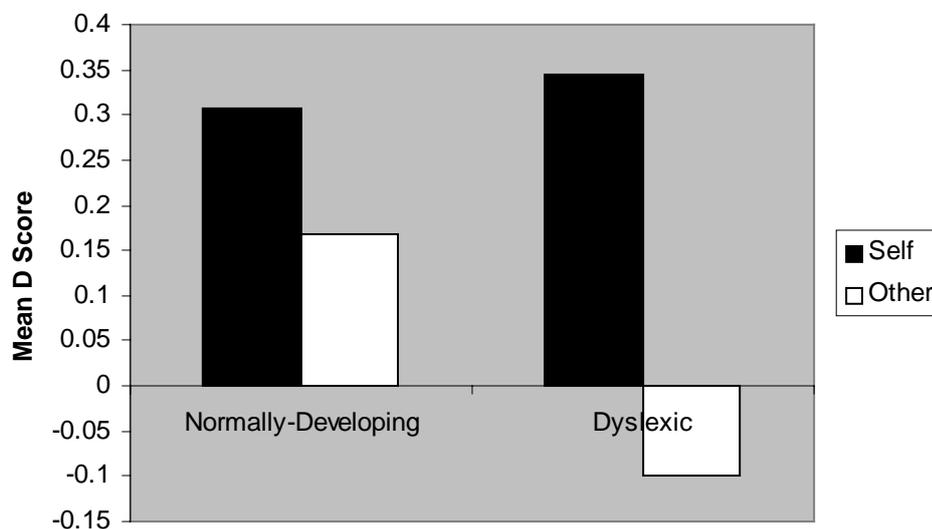
TABLE 7  
The  $D_{\text{IRAP-(trial-type)}}$  Scores (+SE) for the Four Trial-types.

<b>Group</b>	<b>OWNNAME /POSITIVE</b>	<b>OWNNAME /NEGATIVE</b>	<b>OTHER GENDER NAME/POSITIVE</b>	<b>OTHER GENDER NAME/NEGATIVE</b>
<i>Normally-Developing</i>	.381 (.893)	.235 (.636)	.040 (.560)	.297 (.508)
<i>Dyslexic</i>	.442 (.582)	.245 (.626)	.072 (.670)	-.270 (1.691)

The  $D_{\text{IRAP-(trial type)}}$  scores recorded for both groups of participants indicated that responding was in the predicted direction for OWN NAME/POSITIVE, OWN NAME/NEGATIVE and OTHER GENDER NAME/POSITIVE. That is, both groups responded strongly to OWN NAME/POSITIVE as SIMILAR but not OPPOSITE and to OWN NAME/NEGATIVE as OPPOSITE but not SIMILAR, as expected, suggesting positivity towards the self and the absence of negativity. Their attitudes to OTHER GENDER NAME/POSITIVE, although less strong, were still in the predicted direction (OTHER GENDER NAME/POSITIVE/OPPOSITE over SIMILAR), showing negativity towards others. However, the Dyslexic group emitted a pattern of non-predicted responding on OTHER GENDER NAME/NEGATIVE, while the Normally-Developing group did not. In other words, the Normally-

Developing children more readily selected OTHER GENDER NAME/NEGATIVE/SIMILAR over OPPOSITE, while the Dyslexics more readily selected OTHER GENDER NAME/NEGATIVE/OPPOSITE. The Normally-Developing children, therefore, demonstrated an implicit negativity to others that was not demonstrated by those with Dyslexia.

For illustrative purposes, the trial-types were grouped by sample, in order to highlight the differences in responding to the self versus responding to others -- see Figure 8. The figure indicates that both groups of children had similar strong and positive attitudes to the self, but they differed somewhat in their attitudes to others. While the Normally-Developing children demonstrated a lack of positivity combined with greater negativity towards others, the Dyslexic children's attitudes to others were more mixed, and were generally in the non-predicted direction (i.e. lack of predicted negativity).



---

Figure 8: Group mean D scores for sample.

---

The  $D_{IRAP-(trial-type)}$  scores were entered into a 2x3x4 mixed repeated measures ANOVA, with group as the between participant variable and blocks and trial-type as the within participant variables. The analysis revealed no significant main effects for: group ( $p = .35$ ) or blocks ( $p = .57$ ); although trial-type was significant [ $F(3, 18) = 2.99, p = .0388, \eta_p^2 = .365$ ]. No interaction effects were recorded (all  $p$ 's  $> .12$ ).

In order to examine potential differences between the two groups on trial-type, two separate one way 3x4 repeated measures ANOVA's were conducted for each group, with blocks and trial-type as the within participant variables. The analyses revealed no significant main effect for blocks for either group (all  $p$ 's  $> .359$ ) and no main effect for trial-type (Normally-Developing:  $p = .20$ ), although the Dyslexic group was approaching significance ( $p = .0547$ ).

Four separate one-group t-tests were then conducted to determine whether each of the  $D_{IRAP-(trial-type)}$  scores differed significantly from 0 -- see Table 8. The D scores for OWN NAME/POSITIVE and OWN NAME/NEGATIVE were significant or approaching significance for both groups of children. OTHER GENDER NAME/POSITIVE was non-significant for both. However, OTHER GENDER NAME/NEGATIVE was significant only for the Normally-Developing, but not for the Dyslexic group. These outcomes support the previous findings and indicate that the children had strong and positive attitudes particularly to themselves. The Normally-Developing children (but not the Dyslexics) had strong views of others as negative.

TABLE 8  
The T-test Results By Trial-type for Both Groups.

	<b>OWN NAME /POSITIVE</b>	<b>OWN NAME /NEGATIVE</b>	<b>OTHER GENDER NAME/POSITIVE</b>	<b>OTHER GENDER NAME/NEGATIVE</b>
<i>Normally-Developing</i>	$p < .0266$	$p > .0517$	$p = .6956$	$p < .0033$
<i>Dyslexic</i>	$p < .0003$	$p < .0405$	$p > .3890$	$p > .2675$

*Explicit Measure: Piers-Harris 2*

Prior to exposure to the IRAP, all children had been exposed to the Piers-Harris 2 as an explicit measure of self-esteem. Only two children (both normally-developing) produced Response Bias scores that were problematic (i.e. <30). The TOT scores for individuals in each group are presented in Table 9.

TABLE 9  
Piers-Harris 2 TOT Scores for All Participants Grouped by Self-Esteem Range.

Self-Esteem Range	Normally-Developing				Dyslexic				
<i>Low</i>	-	-	-	-	33	37	-	-	-
<i>Low-Average</i>	41	-	-	-	44	44	-	-	-
<i>Average</i>	49	53	55	50	47	55	53	46	48
<i>High-Average</i>	58	58	-	-	-	-	-	-	-
<i>High</i>	61	60	60	-	66	-	-	-	-

Of the Normally-Developing group (n=10) one child (10%) scored within the low-average range; four (40%) scored as average; two (20%) as high-average, and three (30%) as high. Within the Dyslexic group (n=10), two children (20%) scored as low; one (10%) as low-average; five (50%) as average and one (10%) as high. The similarities across the two groups were also reflected in the group means, with the self-esteem mean for the Dyslexic children at 47.30 and 55.00 for the Normally-Developing children. Taken together, the TOT scores indicated that the majority of children were within the average to high range of self-esteem and that there were little or no differences in generic self-esteem between children who did and did not have Dyslexia. The more specific domain scores of all children were analysed closely (see Appendix C), and indicated the Normally-Developing group scored within the

Average range on all the domain scales in comparison to the Dyslexic group who scored in the Low Average range on Intellectual and School Status and Physical Appearance.

### *Correlational Analyses of Implicit and Explicit Measures*

One of the primary aims of the current study was to assess any potential correlation between the implicit and explicit measures. A Pearson's  $r$  was calculated between the explicit total TOT score and the implicit overall D IRAP score for each participant, but the correlation did not reach significance ( $r = .255$ ,  $p = .26$ ).

## DISCUSSION

The primary aim of Experiment 2 was to extend the preliminary IRAP work reported in Experiment 1 by examining implicit self-esteem in a group of children with a formal diagnosis of Dyslexia, relative to normally-developing counterparts. Once again, a critical feature of the research was to determine the utility of the procedure with the group of children who presented with a specific learning difficulty.

The IRAP revealed that both groups of children had strong positive attitudes to the self and readily disconfirmed that they were negative. Although both groups demonstrated weaker implicit attitudes to the other gender, only the Normally-Developing children were consistently negative towards others, the Dyslexic children were not (i.e. they did not agree that they were positive but they did disagree that they were negative).

The current study was the first to employ the IRAP with a sample of children whose diagnosis would be subsumed under the label SEN and thus considerably extends the research with the normally-developing children in Experiment 1. The current study, of course, also involved normally-developing children, whose implicit

positivity towards themselves was consistent with the findings from the previous study. While it was perhaps unexpected that the Dyslexic children showed equally strong positive implicit self-regard, the IRAP did detect differences between the groups in their implicit attitudes to others, with the Dyslexic children showing more mixed attitudes and less negativity towards others relative to the Normally-Developing sample here and in Experiment 1.

The IRAP outcomes were generally consistent with the results of the Piers-Harris 2 in which both groups of children scored within the average range for self-esteem, with only minor superiority for the Normally-Developing group in the overall TOT score and no differences on specific domain scores. Unexpectedly, however, statistical analyses failed to find a correlation between the two measures.

Although the IRAP successfully highlighted some differences between the Normally-Developing and Dyslexic children in terms of their implicit attitudes particularly to others, their attitudes to the self on both the implicit and explicit measures indicated that these children had normally high levels of implicit and explicit self-regard. It might reasonably be argued, therefore, that Dyslexia as a specific learning difficulty does not pose sufficient intellectual, emotional or educational difficulties to significantly differentiate this group of children from their normally-developing counterparts, at least with regard to self-esteem. Furthermore, children with this diagnosis, where the problem is identified early, frequently have access to support services that commonly include an emphasis on building self-esteem (Dyslexia Association of Ireland, 2007). Taken together, these factors suggest that the self-esteem differences between normally-developing and Dyslexic children might be less than was initially predicted. Furthermore, the frequent access to support services might indeed account, at least in part, for the absence of negativity on the

part of the Dyslexic children towards others. As a result, a more challenging form of disability, such as Emotional Behavioural Disturbance (EBD), might be necessary to facilitate greater distinctions from normal development within the IRAP. This issue was addressed in Experiment 3.

## Chapter 4

### Examining the Utility of the IRAP in the Study of Gender and Self-Concept with Normally-Developing and ADHD Children

## **Chapter 4: Experiment 3**

### **Examining the Utility of the IRAP in the Study of Gender and Self-Concept with Normally-Developing and ADHD Children**

#### **INTRODUCTION**

The results from Experiment 2 indicated that the IRAP revealed similar positive implicit attitudes to the self for both Normally-Developing and Dyslexic children. However, differences were observed between the groups in relation to their implicit perceptions of others. Although self-differences between the two groups may have been predicted but did not emerge, the specific nature of Dyslexia (i.e. it is not deemed to be a form of general learning difficulty) and the children's access to support services may have limited the differences in terms of self-esteem between these and children described as normally-developing. Put simply, the self-esteem of children with Dyslexia may not, on the whole, differ greatly from normally-developing children because they are not noticeably different socially, emotionally or behaviourally. As a result, it would seem reasonable to suggest that perhaps a group of children with more explicit behavioural difficulties (e.g. ADHD) would present with a more differentiated self-esteem profile that could be identified using the IRAP. Put simply, the fact that children with ADHD have emotional and behavioural disturbances that render them even more identifiable to peers suggests that they are more likely to be rejected, and are thus more likely to suffer from altered self-esteem.

### *Attention Deficit/Hyperactivity Disorder (ADHD)*

Attention Deficit/Hyperactivity Disorder is characterised by observable behavioural features that include; a short attention span, distractibility, impulsivity and overactivity (Mercugliano, 1999). Secondary clinical features also associated with the disorder include poor peer relations, aggression, learning problems, academic under-achievement and depressive symptomatology (Barkely, 1998). These behavioural characteristics have been observed across settings and are reported to cause notable levels of functional impairment. ADHD is three times more prevalent in boys than girls (Biederman, Mick, Farone, Braten, & Doyle, 2002) and rates overall range between 2 and 7%, depending upon co-morbidity with other clinical phenomena (Sachill & Schwab-Stone, 2000). In particular, ADHD demonstrates high co-morbidity (50-70%) with Learning Disability, Oppositional Defiant Disorder and Conduct Disorder (American Psychiatric Association, 1994).

Children with ADHD experience considerable difficulties in social functioning (Biederman, Farone, & Chen, 1999; Gentschel & McLaughlin, 2000), with social impairment as a significant predictor of adverse long-term outcomes of the presence of the condition in adolescence (Greene, et al., 1997; 1999). Specifically, recurrent socially inappropriate behaviours are seen as core features of the disorder (Greene, Biederman, Faraone, Ouelleter, & Garcia-Jetton, et al., 1996). These may be expressed as frequent conversational shifts, poor listening skills, inappropriate initiating of conversations, and continual interruptions or intrusions on others (American Psychiatric Association, 1994). It is not surprising, therefore, that a large proportion of children with ADHD are rated as being less popular than normally-developing peers (Wheeler & Carlson, 1994) and thus recipients of more frequent social rejection (Barkley, 1998; Hinshaw & Melnick, 1995).

The categorisation of ADHD is strategically located under EBD, rather than as a specific or general learning difficulty, primarily because this group of individuals generally present with normal levels of intelligence. Nonetheless, the nature of the condition presents serious educational challenges that frequently result in aversive learning experiences and poor academic attainment (Hinshaw, 1992, Semrud-Clikeman, et al., 1992). Specifically, children with ADHD experience increased school failure, greater need for additional tutoring, grade repetitions and placements in special education classes (Barkley, Anastopoulos, Guevremont, & Fletcher, 1991). Not surprisingly, this results in lower levels of schooling overall and poor occupational attainment (Manuzza, Klein, Bessler, Malloy, & LaPadula (1993). In fact, although teachers appear to be generally in favour of inclusive practices for children with SEN (including ADHD) their attitudes in this regard are invariably contingent upon the provision of appropriate support services that match the children's needs (Butler & Shevlin, 2001). Teachers in Ireland, for example, report that they feel ill-equipped to manage SEN children effectively within the mainstream context (Scanlon & McGilloyay, 2006) and perceive children with EBD amongst the most difficult to integrate (Ntinias, et al., 2006).

### *ADHD and Self-esteem*

While it is generally assumed that children with ADHD present with lower levels of self-esteem than normally-developing peers (Selikowitz, 2002), empirical evidence in this regard is equivocal. The majority of studies in this area report that self-esteem is lower than average. For example, Slomkowski, Klein and Mannuzza (1995) reported a long-term relationship between hyperactivity in childhood, poor functioning and low self-esteem into adolescence and adulthood. Specifically,

follow-up assessments at ages 18 and 26 indicated lower self-esteem, lower full scale IQ scores, lower interviewer rated psychosocial adjustment, lower educational achievement and the presence of a greater number of ADHD symptoms.

In a related study, Bussing, Zima and Perwien (2000) reported that although the self-esteem levels of children with ADHD were within the normal range, popularity was significantly below normal. Indeed, these researchers concluded that independent predictors of low self-esteem included being white, having a high functional impairment, having co-morbid internalising conditions, and the absence of medication. Other researchers have also highlighted the importance of aggression in perceptions of children with ADHD. For example, Treuting, Treuting and Hinshaw (2001) compared ADHD males aged between seven and twelve years old, who did and did not present with aggression. The results indicated that those with an aggressive status showed higher rates of depressive symptomatology and lower levels of self-esteem, relative to the others.

A series of other studies, however, appear to suggest no differences in self-esteem between ADHD and normally-developing individuals (Hoza, Pelham, Milich, Pillow, & McBride (1993). In one study with young children, Stewart and Buggy (1994) reported no significant differences for peer rating or self-esteem scores between third graders who: were normally-developing, presented with ADD (Attention Deficit Disorder), or presented with ADHD. In fact, these researchers argued that children with ADHD might focus explicitly on the positive rather than negative reactions of peers, thus perhaps *enhancing* self-esteem outcomes. This perspective is supported elsewhere in the literature. First, Barkley (1998) argued that children with ADHD are often confused by the rejection of their peers and thus frequently do not interpret it correctly. Second, children with ADHD, particularly

those who present with aggression, have been reported to overestimate their competence, with inflated self-perceptions in social and behavioural domains.

Indeed, young low academically achieving males with ADHD have been reported to inflate scores in the scholastic domain (Hoza, Pelham, Dobbs, Owens, & Pillow, 2002).

Some researchers have accounted for the differential self-esteem outcomes in ADHD by highlighting the importance of co-morbidity, features of which are often associated with different self-esteem profiles. Indeed Hoza et al. (2002) reported that boys with ADHD who have co-morbid depression present with lower self-perceptions than those without depressive symptoms. Taken together, therefore, the existing evidence with regard to the relationship between self-esteem and ADHD is inconclusive. To date, there have little or no attempts by researchers to study the implicit attitudes of individuals with ADHD, a methodology that might shed light on the self-esteem controversy, particularly because of the fact that the majority of studies reviewed thus far have employed explicit methods.

### *The Current Research*

The primary aim of Experiment 3 was to extend the previous studies by using the IRAP to draw comparisons of the implicit self-esteem of normally-developing children and those diagnosed with ADHD. The research predicted that the difference between these two groups in this regard would be greater than that observed between the normally-developing children and those with Dyslexia. Once again, the children were also presented with the Piers-Harris 2 to enable comparisons between implicit and explicit measures of self with these samples.

## METHOD

### *Participants*

Twenty children were recruited for participation in Experiment 3. They ranged in age from 8 to 11 years (mean age was 9 years and 11 months). All were recruited by the Researcher through direct contacts with local Primary and Post-Primary schools in the Kildare region and parental consent was obtained in all cases. The participants were divided into two groups that were differentiated in terms of the presentation of specific behavioural problems. The Normally-Developing group comprised of ten participants (6 females and 4 males) who were within the ‘normal’ range of intellectual development and who had no prior history of a behavioural difficulty or evidence of a learning disability. Participants in the ADHD Group (2 females and 8 males) all had a formal diagnosis of ADHD, as well as diagnoses of other psychological conditions. The latter are identified in Table 10.

TABLE 10  
The Psychological Conditions and Learning Disabilities of Participants in Experiment 3.

<b>Psychological Condition</b>	<b>Participant Number</b>									
	1	2	3	4	5	6	7	8	9	10
<i>Oppositional Defiant Disorder</i>	√									
<i>Asperger's Syndrome</i>		√	√					√		
<i>General Learning Disability</i>		√	√		√		√			
<i>Dyslexia</i>					√	√				

### *Setting*

For both groups, all aspects of Experiment 3 were conducted in a quiet room in each of the participant's schools. In all cases, the experimental setting remained free from noise and distraction throughout all aspects of the study. All participation

was conducted on an individual basis. The Researcher and an SEN Assistant remained present in the room at all times.

### *Apparatus and Materials*

All aspects of the apparatus and materials employed in Experiment 3 were identical to Experiment 2. However, at the end of each training and testing block a feedback re-enforcer incorporating each participants name saying e.g “WELL DONE MARY” and cartoon jiffs (i.e. a frog hopping out of the water) was developed to keep the children with ADHD on task.

### *Procedure*

All aspects of the procedure employed in Experiment 3 were identical to Experiment 2, except for the visual modifications to feedback. All participants completed the study in a single experimental session that lasted between 30 and 40mins.

## RESULTS

### *Response Latency Data*

Mean response latencies (three consistent and three inconsistent) were calculated (after adjustments) for each participant on the six test blocks and were then collated by group -- see Table 11. Participants in the Normally-Developing group responded more quickly on both consistent and inconsistent blocks than the ADHD group. On all three sets of blocks, the Normally-Developing group demonstrated an IRAP effect. In contrast, the ADHD group responded more quickly to inconsistent trials over consistent trials (i.e. reverse IRAP effects) on two of the three block sets.

TABLE 11  
The Mean Response Latencies for Consistent and Inconsistent Responding  
on Test Blocks For Both Groups

	Blocks 1 and 2		Blocks 3 and 4		Blocks 5 and 6	
	Con	Incon	Con	Incon	Con	Incon
<i>Normally-Developing</i>	2123	2203	2075	2196	1906	2033
<i>ADHD</i>	2531	2401	2299	2435	2252	2233

Two 2x3x2 mixed repeated ANOVAs were conducted on the latency data (one per group), with IRAP condition and blocks as within participant variables and order as the between participant variable. Consistent with the descriptive analysis, the Normally-Developing group showed a significant main effect for IRAP condition [ $F(1, 8) = 9.65, p = .0386, \eta_p^2 = .409$ ], but the ADHD group did not ( $p = .80$ ). The IRAP effects for both groups are illustrated in Figure 9. No main effects for Order were recorded for either group (all  $p$ 's  $> .34$ ). A significant main effect for blocks was found for both groups (Normally-Developing: [ $F(2, 8) = 3.743, p = .0386, \eta_p^2 = .061$ ] and ADHD: [ $F(2, 8) = 11.15, p = .0009, \eta_p^2 = .120$ ]). Post hoc tests (Scheffe's) revealed no significant differences on Blocks for the Normally-Developing group, but found that the significant difference lay between Blocks 1 and 2 for the ADHD group, suggesting a strong practice effect. The ADHD group showed a two-way interaction effect for Block by IRAP condition [ $F(2, 8) = 5.23, P = .0178, \eta_p^2 = .036$ ]. However, Post-Hoc analyses (Scheffe) revealed no significant differences. No significant interaction effects were recorded for the Normally-Developing Group (all  $p$ 's  $> .16$ ).

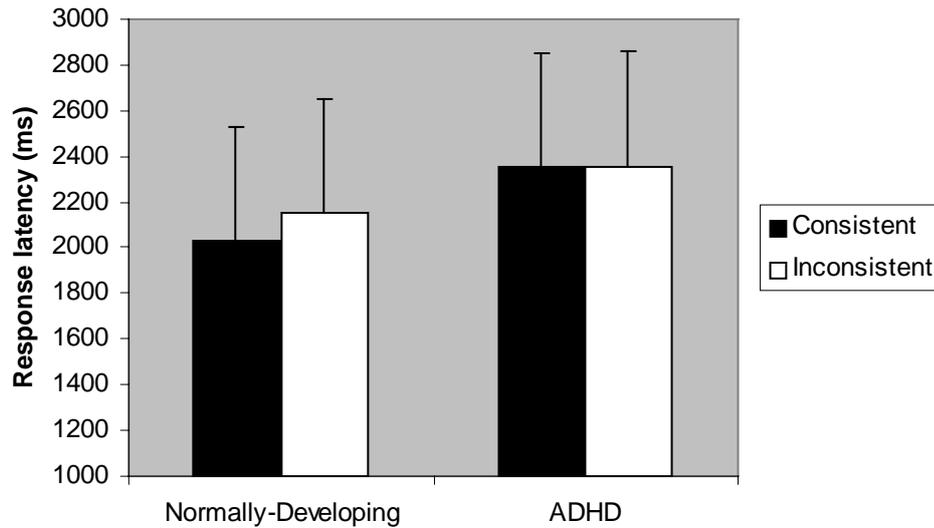


Figure 9: Mean response latencies for consistent and inconsistent responding for the two groups.

### *Preliminary Trial Analyses*

Trial-type analyses (OWN NAME/POSITIVE; OWN NAME/NEGATIVE; OPPOSITE GENDER NAME/POSITIVE; and OPPOSITE GENDER NAME/NEGATIVE) were conducted on the latency data to explore the impact of trial-types on the IRAP effects. Overall mean response latencies for each participant on consistent and inconsistent trials for each of the four trial-types were calculated and collated by group and these are presented in Table 12.

TABLE 12  
The Mean Response Latencies (+SE) for Consistent and Inconsistent Responding By Trial-Type and Group

Group	OWN NAME /POSITIVE		OWN NAME /NEGATIVE		OPPOSITE GENDER NAME/POSITIVE		OPPOSITE GENDER NAME/NEGATIVE	
	Con	Incon	Con	Incon	Con	Incon	Con	Incon
Normally -Developing	1948	2344	2088	2289	2423	2368	2514	2466
ADHD	2560	2614	2510	3018	2652	2675	3058	3078

The table indicates that participants the Normally-Developing group responded more quickly on all four trial-types relative to the ADHD group. A 2x2x2x2 mixed repeated measures ANOVA was conducted for each group with sample (OWN NAME/OPPOSITE GENDER NAME), target-type (POSITIVE/NEGATIVE), order and IRAP effect as the repeated measures variables. This analysis indicated that only the ADHD group showed a significant IRAP effect [F (1,8) = 6.03 p = .0395,  $\eta_p^2$  = .020], the Normally-Developing group did not (p > .14). Sample was significant for the ADHD group [F (1,8) = 8.68, p = .0185,  $\eta_p^2$  = .075], but not for the Normally-Developing group. Target was significant for both groups -- Normally-Developing [F (1,8) = .18.50, p = .0026,  $\eta_p^2$  = .051], and ADHD [F (1,8) = 6.32, p = .0361,  $\eta_p^2$  = .032]. Order was non-significant for both groups (all p's > .67). All interaction effects were non-significant (all p's > .19).

*The D-IRAP Algorithm and Within-Group Comparisons*

Consistent with the analyses conducted on the D scores in Experiment 2, both types of D<sub>IRAP</sub>-algorithm analyses were employed here. The D<sub>IRAP</sub>-(trial-type) scores were calculated for each participant and collated by group and these are presented in Table 13.

TABLE 13  
The D<sub>IRAP</sub>-(trial-type) Scores (+SE) for the Four Trial-types By Group.

Group	OWN NAME /POSITIVE	OWN NAME /NEGATIVE	OPPOSITE GENDER NAME/POSITIVE	OPPOSITE GENDER NAME/NEGATIVE
Normally- Developing	.424 (.539)	.292 (.671)	.020 (.654)	.092 (.623)
ADHD	.072 (.625)	.383 (.526)	.006 (.715)	-.084 (.530)

The  $D_{\text{IRAP-(trial type)}}$  scores recorded for both groups of participants indicated that responding was in the predicted direction for OWN NAME/POSITIVE (much larger D scores for Normally-Developing), OWN NAME/NEGATIVE (larger D scores for ADHD) and OPPOSITE GENDER NAME/POSITIVE (larger D scores for Normally-Developing). However, the ADHD group emitted a pattern of non-predicted responding on OPPOSITE GENDER NAME/NEGATIVE, although the Normally-Developing group did not. That is, the Normally-Developing children more readily selected OPPOSITE GENDER NAME/NEGATIVE/SIMILAR over OPPOSITE, but the ADHD group more readily selected OPPOSITE GENDER NAME/NEGATIVE/OPPOSITE over SIMILAR.

In summary, therefore the  $D_{\text{IRAP-(trial type)}}$  scores indicated that the Normally-Developing children had a stronger and more positive attitude to the self than the ADHD children. Both groups also differed in their attitudes to others. Specifically, while the Normally-Developing group demonstrated a lack of positivity and greater negativity towards others, the ADHD group had mixed attitudes in that they disagreed that others were positive but did not agree that they were not negative. For illustrative purposes, the trial-types were grouped by sample in order to highlight the differences in responding to the self versus responding to other -- see Figure 10. The Normally-Developing children demonstrated positivity to the self and negativity to others that was consistent with experimental predictions. While the self-positivity was shared by the children with ADHD, the latter overall demonstrated non-predicted attitudes that indicated that they did not agree that others were negative.

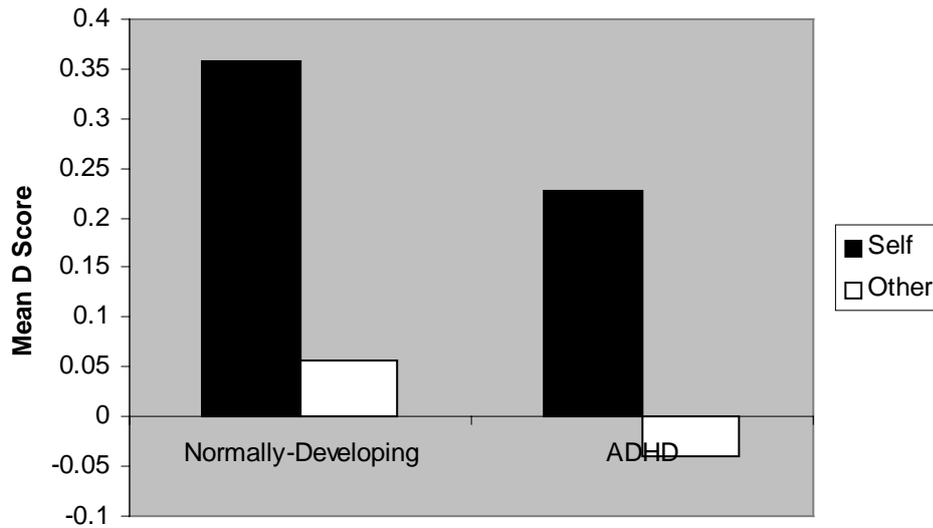


Figure 10: Group mean D scores for trial-types.

The  $D_{IRAP-(trial-type)}$  scores were entered into a 2x3x4 mixed repeated measures ANOVA, with group as the between participant variable and blocks and trial-type as the within participant variables. The analysis revealed a non-significant main effect for group ( $p = .18$ ). A significant main effect was recorded for blocks [ $F(2,18) = 8.13, p = .0012, \eta_p^2 = .744$ ] and Post-hoc tests (Scheffe) indicated that the significant difference lay between Blocks 1 and 2 and between Blocks 3 and 4. Trial-type was also significant [ $F(3, 18) = .4.66, p = .0057, \eta_p^2 = .740$ ], with Post-hoc analyses indicating that the significant difference lay between OWN NAME/POSITIVE and OPPOSITE GENDER NAME/NEGATIVE, as well as between OWN NAME/NEGATIVE and OPPOSITE GENDER NAME/POSITIVE. A two-way interaction was recorded for trial-type by group [ $F(2,18) = 4.13, p = .0242, \eta_p^2 = .377$ ].

Two separate one way 3x4 repeated measures ANOVA's were conducted (one for each group), with blocks and trial-type as the within participant variables. A

significant main effect for blocks was recorded for the ADHD group [ $F(2,9) = 9.366$ ,  $p = .0016$ ,  $\eta_p^2 = .238$ ] and Post-hoc analyses (Scheffe) indicated that the significant difference lay between Blocks 1 and 2. Blocks was non-significant for the Normally-Developing group ( $p > .2953$ ). Trial-type was also significant for the ADHD group [ $F(3, 9) = 1.235$ ,  $p = .0356$ ,  $\eta_p^2 = .458$ ] and approached significance for the Normally-Developing group [ $F(3,9) = 1.028$ ,  $p = .0557$ ,  $\eta_p^2 = .588$ ] but post-hoc analyses in both cases revealed no significant differences on any particular trial-type.

Four separate one-group t-tests were then conducted to determine whether each of the  $D_{IRAP-(trial-type)}$  scores differed significantly from 0 -- see Table 14. OWN NAME/POSITIVE was significant for the Normally-Developing group, but not for the ADHD group. OWN NAME/NEGATIVE was significant for both groups. OPPOSITE GENDER NAME/POSITIVE and OPPOSITE GENDER NAME/NEGATIVE were non-significant for both groups. Hence, the Normally-Developing children had stronger positive implicit attitudes about themselves than the ADHD children in terms of responding to OWN NAME/POSITIVE. Both groups, however, had strong views that they were not negative. Both groups failed to demonstrate strong attitudes to others as either positive or negative.

TABLE 14  
The T-Test Results By Trial-type for Both Groups.

	OWN NAME /POSITIVE	OWN NAME /NEGATIVE	OPPOSITE GENDER NAME/POSITIVE	OPPOSITE GENDER NAME/NEGATIVE
<i>Normally-Developing</i>	$p < .0002$	$p < .0237$	$p = .8696$	$p > .4247$
<i>ADHD</i>	$p > .5350$	$P < .0004$	$p = .9614$	$p > .3920$

*Explicit Measure: Piers-Harris 2*

All children were exposed to the Piers-Harris 2 prior to the IRAP. None of them produced Response Bias scores that were problematic. The TOT scores for all children separated by group are presented in Table 15.

TABLE 15  
Piers-Harris 2 TOT Scores for All Participants  
Grouped by Self-Esteem Range and Group.

Range	Normally-Developing					ADHD				
Low	38	-	-	-	-	34	37	-	-	-
Low-Average	-	-	-	-	-	40	40	-	-	-
Average	51	55	56	54	-	55	50	45	45	
High-Average	58	-	-	-	-	59	-	-	-	-
High	61	66	60	62	-	70	-	-	-	-

Of the Normally-Developing group (n = 10), one child (10%) scored within the low range; four (40%) scored as average; one (10%) as high-average, and four (40%) as high. Within the ADHD group (n = 10), two children (20%) scored as low; two (20%) as low-average; four (40%) as average; one (10%) as high-average; and one (10%) as high. The differences across the two groups were also reflected in the group means, with the self-esteem mean for the ADHD children at 48 (average) and 56 (high-average range) for the normally-developing children. Taken together the TOT scores indicated that the majority of children in the ADHD group were within the average range, whereas children in the Normally-Developing group were dispersed from average to high levels of self-esteem.

The more specific domain scores of all children were analysed closely (see Appendix D). In general, the ADHD group produced lower scores across the subscales than the Normally-Developing group. Specifically on Behavioural Adjustment,

the ADHD group scored within the low range (38), while the Normally-Developing group scored as average (54). This finding indicated that the ADHD children explicitly perceived themselves as presenting with poorer behavioural adjustment than the Normally-Developing children. Not surprisingly, the latter group perceived their behaviour to be normally adjusted. On the Popularity domain, the ADHD group also scored as low (39), compared to the Normally-Developing group who scored as average (54). This finding indicated that the ADHD children perceived themselves to be of lower than normal popularity, while the Normally-Developing children perceived themselves to be of average or normal popularity.

## DISCUSSION

The primary aim of Experiment 3 was to extend the previous research by examining implicit self-esteem in a group of children with a formal diagnosis of ADHD, relative to normally-developing counterparts. Once again, a critical feature of the research was to determine the utility of the procedure with this group of children and the findings further supported the utility of the procedure in this regard. Once again, all children proceeded rapidly and with ease through all aspects of the procedure.

The IRAP outcomes revealed considerable differences in the implicit attitudes to the self and others of the Normally-Developing and ADHD children that were consistent with predictions. Put simply, the former group demonstrated stronger positive attitudes to the self relative to the latter. Both groups demonstrated weak implicit attitudes to the others.

Although the two groups of children did not differ greatly in their overall profile on the explicit Piers Harris 2 measure, 90% of the Normally-Developing group

scored at least average compared to 63% of children with ADHD. However, the data from the specific domain scores indicated that the ADHD group had low ranges on both Behavioural Adjustment and Popularity domains, while the normally-developing group did not.

Taken together, the IRAP outcomes that revealed less positive implicit attitudes to the self by the ADHD children relative to their normally-developing peers were consistent with the modest differences between the two groups on the overall TOT score and particularly on the Behavioural Adjustment and Popularity sub-scales. Interestingly, the relative outcomes with both groups are also consistent with those recorded previously, in which normally-developing children again present with strong implicit self-esteem and the ADHD children present with greater difficulties in this regard compared to those who are normally-developing or Dyslexic. Indeed, the IRAP indicated, as predicted, that ADHD appears to present those children affected with a more significant intellectual and emotional challenge than Dyslexia that is responsible, at least in part, for lower self-esteem.

## Chapter 5

### Examining the Utility of the IRAP in the Study of Teacher's Implicit Attitudes to Children with Emotional Behavioural Difficulties

## **Chapter 5: Experiment 4**

# **Examining the Utility of the IRAP in the Study of Teachers' Implicit Attitudes to Children with EBD**

### INTRODUCTION

“The educational curriculum comprises all intended and unintended learning experiences that take place in and through the school, thus enabling all learners to manipulate information and knowledge, to understand their emotional needs and to develop their own sense of consciousness, in short, it contributes to how a learner constructs a notion of self.”

O' Brien (2000, p. 10).

Several theoretical models have attempted to address the processes by which teachers' expectations are communicated to students (Cooper, 1979; Cooper & Good, 1983). In short, teachers' expectations influence (either consciously or unconsciously) overt and covert behaviour, which in turn is communicated to students and ultimately affects academic self-efficacy and/or performance. Because self-efficacy appears to be a comparative process against which the self is compared to others, teachers' attitudes appear to influence self-efficacy by enabling the learners to appraise differential teacher behaviours that indicate high versus low achievement (Frederick, Deitz, Bryceland, & Hummel, 2000). Achievement, however, in this context is likely not restricted to academic matters, but also incorporates teachers' reactions to other matters, such as students' socialisation and personality (Weinstein, 2002).

### *Teaching Pupils with SEN*

Within education, children with EBD are perceived to be amongst the most difficult to educate because of the common presence of challenging behaviour (Bay & Bryan, 1991; Li, 1985; Ntinas, et al., 2006). For example, students with ADHD (particularly males) exhibit above average rates of gross motor activity and fidgeting, negative verbalisations and off-task behaviour (Abikoff, Jensen, Arnold, & Hoza et al., 2002). Furthermore, several studies have highlighted the relationship between problem behaviour and academic performance (Larsen, Steele, & Sailor in press; McIntosh, 2005; Tobin & Sugai, 1999), including higher rates of office discipline referrals (Irwin, Tobin, Sprague, Sugai, & Vincent, 2004).

More implicit variables exacerbate teachers' interactions with SEN children. Specifically, it is widely accepted that the success of schemes designed specially for the education of children with SEN (e.g. inclusion) depends upon teachers' attitudes (e.g. Chow & Winzer, 1992; Hastings & Oakford, 2003) and on the sophistication of their skills in managing these populations (Koegel, Harower, & Koegel, 1999). Indeed, Bekle (2004) reported that teachers respond differentially to SEN children (relative to normally-developing) by demanding less, calling upon them less frequently, praising them less and criticising them more. Research into teachers' attitudes has also indicated that these are influenced primarily by their perceptions of their own skills which are in turn influenced by levels of appropriate resources (Butler & Shevlin, 2001). However, teachers in Ireland, for example, have reported that they feel ill-equipped in managing SEN children effectively because of inadequate support (Scanlon & McGilloway, 2006). Explicit methodologies are the hallmark of research into teachers' attitudes. Indeed, at the time of writing to the author's knowledge, no

previous study has used an implicit procedure to study teachers' attitudes to children with SEN.

### *Implicit Attitudes to SEN*

While no empirical studies have specifically examined the implicit attitudes of teachers to SEN, , one study has employed the IRAP study to examine the implicit attitudes of three groups who differed in their range of experience, working with children with Autistic Spectrum Disorder (ASD) (See chapter 1). The results highlighted the discordance between implicit and explicit measures on socially sensitive complex relations. Barnes-Holmes et al., (2006) suggest that the conflict between what one thinks privately and what one might say publicly may inadvertently lead to various forms of psychological stress, which may in turn lead to professional burnout and high turnover rates.

### *The Current Research*

The primary aim of Experiment 4 was to use the IRAP to examine teacher's implicit attitudes to children with EBD. The sub-category EBD was selected over the broader category of SEN because the former represent the most difficult sub-group of children to educate (see above) and thus it seemed likely that the IRAP would be able to tap into teachers attitudes' regarding these children. In order to facilitate clear distinctions in implicit attitudes, the current study juxtaposed EBD PUPIL with TEACHER within the IRAP. Put simply, it was predicted that teachers would show a preference for TEACHER over EBD PUPIL, and thus would more readily associate TEACHER with positive attributes than negative attributes and more readily associate EBD PUPIL with negative attributes than positive. Different sub-groups of teachers with different levels of teaching experience were also identified in order to determine whether there was a relationship between implicit attitudes to TEACHER/EBD

PUPIL and level of teaching experience. Specifically, the study involved a Control Group of undergraduate students with no teaching experience; a group of Post-Primary Teachers in Training with some but limited teaching experience; experienced Primary Teachers; and experienced Post-Primary Teachers. All three groups of teachers had direct experience of teaching children with EBD in their classrooms.

## METHOD

### *Participants*

One hundred and twenty adults participated voluntarily in Experiment 4. These individuals were recruited by the Researcher through direct contact with primary and post-primary schools in the Dublin region and with the Department of Education at NUIM. The participants were divided according to four groups that were differentiated in terms of their levels and type of educational training and experience with children. The *Control Group* ( $n = 20$ ) comprised of undergraduate psychology students who had no previous experience of working directly with children with EBD in an educational setting (although these individuals would likely have had some knowledge of the difficulties associated with this condition). The *Teachers in Training* ( $n = 20$ ) were postgraduate students at NUIM pursuing a Higher Diploma in Education (i.e. teacher training) for Post Primary teaching. All of these participants were engaged in teaching practice two days per week and had direct experience of working with children with EBD in their classrooms. The *Primary Teachers* ( $n = 20$ ) were all currently employed in primary schools; had a formal qualification in teaching; had at least two years teaching experience in mainstream education; and had direct experience of working with children with EBD. The *Post-Primary Teachers* ( $n = 20$ ) were all currently employed as teachers in post-primary education; had a formal qualification in teaching; had at least two years teaching experience in mainstream education; and had direct experience of working with children with EBD. None of the three groups of teachers, however, had *formal* qualifications for working with children with SEN.

### *Setting*

For the Control Group and the Teachers in Training, all aspects of Experiment 4 were conducted in the Computer Laboratory at the Department of Psychology at NUIM. For the Primary and Post-Primary Teachers in Training, the experiment was conducted in a quiet room in their school. In all cases, the experimental setting remained free from noise and distraction throughout all aspects of the study. All participation was conducted on an individual basis. The Researcher remained present in the room at all times.

### *Apparatus and Materials*

*Opinions Relative to Mainstreaming Scale (ORM)*. One of the *explicit* measures employed in Experiment 4 comprised of a revised version of the Opinions Relative to Mainstreaming (ORM) Scale (Avramidis, Bayliss, & Burden, 2000). The original ORM Scale (Antonak & Larrivee, 1995; Larrivee, 1982) investigated attitudes to the generic concept of disabilities and was subsequently tailored by Avramidis et al., to fit an English context and to incorporate attitudes to the concept of inclusion within education for all persons with disabilities. For the purposes of the current study, the ORM Scale was further modified to fit the Irish context and incorporated specific attitudes to the inclusion of children with EBD in mainstream education. A copy of the ORM Scale employed here is presented in Appendix C. The ORM Scale comprised of four basic components. (1) The *Demographic Questionnaire* collected information primarily regarding each participant's teaching qualifications; number of years teaching experience; and additional training (if any) in special education.

(2) The *Likert scale* of attitudes to the general inclusion of children with EBD in mainstream education consisted of 12 items, such as “*The needs of students with EBD are best served through special separate classes*”. Respondents were required to indicate their level of agreement with each statement by circling: Strongly Disagree; Disagree; Undecided; Agree; or Strongly Agree.

(3) The *Emotional Reaction Scale* measured reactions to the inclusion of EBD children within the participants’ own classrooms. The scale presented a short scenario: *If a new student who was described as having severe behavioural problems was about to join your class tomorrow, you would feel . . .*” in conjunction with seven bipolar adjectives (e.g. UNCOMFORTABLE/COMFORTABLE). Participants were required to respond on a 7-point scale from 1 (negative end – UNCOMFORTABLE) to 7 (positive end – COMFORTABLE) as an indication of their reaction to the scenario on that particular dimension (i.e. level of discomfort).

(4) The final Likert scale within the ORM Scale measured participants’ willingness to adapt their own teaching practices to include children with EBD. This scale comprised of eight items, such as “*I will accept responsibility for teaching children with EBD within a whole school policy*”, against which participants once again circled their levels of agreement.

Scoring the ORM Scale does not generate an overall result, but comprises of the scores generated on each of the four sub-scales. The scores on the sub-scales are simply composites of the individual scores on each item. In all cases, a higher score indicates a more positive attitude.

*Likert Scale Directly Associated with the IRAP.* All participants were presented with a Likert scale containing the 12 target stimuli that were incorporated into the IRAP (Appendix D). These were six positive words (e.g.

ACCOMMODATING) and six negative words (e.g. NON-COMPLIANT), which participants were asked to rate for likeability. That is, they circled a number from –6 (EXTREMELY UNLIKEABLE) to 6 (EXTREMELY LIKEABLE) that best reflected the extent which they perceived the target word to be positive or negative.

*The IRAP.* All participants completed all experimental trials on DELL computers with Pentium 4 processors. The IRAP was delivered via a program written in Visual Basic (Version 6.0.), which controlled all aspects of stimulus presentation and the recording of all participant responses.

The stimuli employed in the IRAP for Experiment 4 were identical in format to Experiment 3. However, the sample stimuli in this case contained the words TEACHER and EBD PUPIL; the target stimuli contained twelve evaluative terms – six positive (ACCOMODATING, SUITABLE, CO-OPERATIVE, PLEASANT, CALM and POSITIVE) and six negative (DIFFICULT, NON-COMPLIANT, ANGRY, NEGATIVE, UNPLEASANT and INAPPROPRIATE) and the relational terms once again were SIMILAR and OPPOSITE. The stimulus arrangements for Experiment 4 are presented in Table 16.

TABLE 16  
Stimulus Arrangements Employed in Experiment 4.

<b>Sample 1</b> TEACHER	<b>Sample 2</b> EBD PUPIL
<b>Response Option 1</b> Similar	<b>Response Option 2</b> Opposite
<b>Targets deemed consistent with Sample 1</b>	<b>Targets deemed consistent with Sample 2</b>
CO-OPERATIVE SUITABLE POSITIVE CALM ACCOMODATING PLEASANT	DIFFICULT NON-COMPLIANT ANGRY NEGATIVE UNPLEASANT INAPPROPRIATE

The screen shots employed here as visual aids were identical in format to the previous experiments, but were simply modified to incorporate the new word pairs.

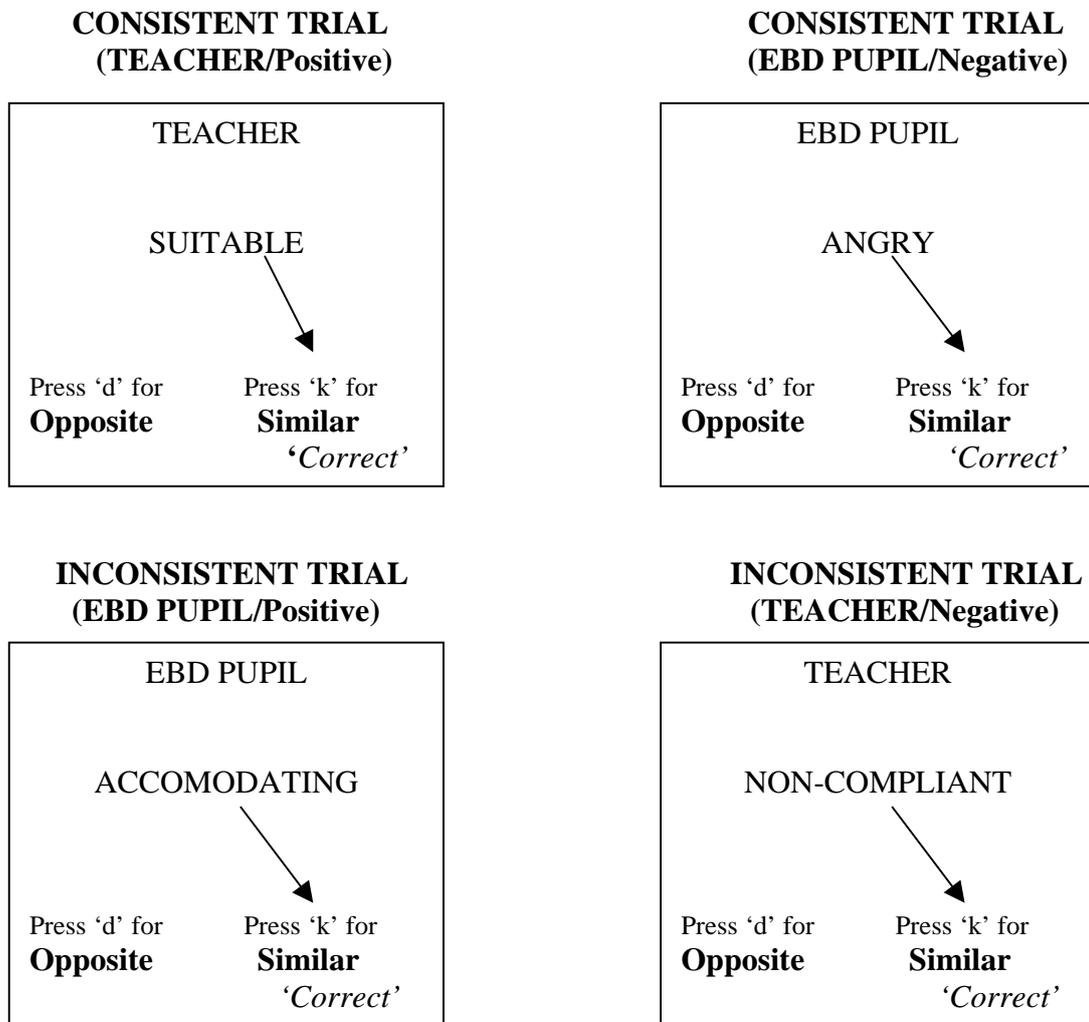
### *Procedure*

*Experimental Sequence.* All participants completed the Likert Scale Directly Associated with the IRAP immediately before exposure to the IRAP itself. However, three of the groups, excluding the Control Group, also completed the ORM Scale several days before exposure to the IRAP (and returned the completed scale to the Experimenter on that day). All aspects of the IRAP were identical in format to the previous studies. Half of the participants in each of the four groups were presented with a block of consistent trials first, while the remaining half were presented with inconsistent trials first.

*IRAP Instructions.* Once again, participants required extensive instructions for completion of the IRAP trials and the specific instructions with which they were provided is presented in Appendix G.

*Trial-types.* The four basic trial-types presented here were referred to as TEACHER/POSITIVE (top left of Figure 11); EBD PUPIL/NEGATIVE (top right); EBD PUPIL/POSITIVE (bottom left); and TEACHER/NEGATIVE (bottom right). Each block of IRAP trials comprised of twelve exposures to each of the four trial-types, which contained two exposures to each of the twelve target words (e.g. ANGRY appeared twice with TEACHER). Correct responses on consistent trials involved positivity towards TEACHER and negativity towards EBD PUPIL, whereas

correct responding on inconsistent trials was reversed (i.e. TEACHER-NEGATIVE and EBD PUPIL-POSITIVE).



---

Figure 11: Schematic representation of the four basic trial-types.

---

At the end of the sixth test block, the following message appeared automatically: *Thank You. That is the End of the Experiment. Please report to the Researcher.* Participants were thanked for their co-operation and time. They were then appropriately debriefed about the research and invited to ask any questions. All

participants completed the study in a single experimental session that lasted between 20 and 30mins. in total. Although made available to them, no participants opted for short breaks at any point

## RESULTS

### *Response Latency Data*

Mean response latencies (three consistent and three inconsistent) were calculated (after adjustments) for each participant on the six test blocks and were then collated by group -- see Table 17. Participants in all four groups responded more quickly on consistent trials than inconsistent trials across the three sets of blocks. The Control Group appeared to produce the shortest mean latencies overall on both types of trials, while the Teachers In-Training generated the longest latencies.

TABLE 17  
Mean Response Latencies For All Groups on Consistent and Inconsistent Responding

Group	Blocks 1 and 2		Blocks 3 and 4		Blocks 5 and 6	
	Con	Incon	Con	Incon	Con	Incon
<i>Control</i>	1650	1575	1610	1616	1613	1618
<i>In-Training</i>	2379	2658	2417	2816	2389	2675
<i>Primary</i>	2274	2472	2239	2272	2071	2357
<i>Post-Primary</i>	2338	2469	2247	2568	2289	2394

A 2x3x2 mixed repeated ANOVA was conducted on the latency data, with IRAP condition and blocks as within participant variables and order (consistent-first vs. inconsistent-first) as the between participant variable. All three groups of teachers showed significant main effects for IRAP condition: In-Training [F (2, 18) = 16.654,

$p = .008$ ,  $\eta_p^2 = .035$ ]; Primary [ $F(2, 18) = 3.244$ ,  $p = .0018$ ,  $\eta_p^2 = .047$ ]; and Post Primary [ $F(2, 18) = 9.639$ ,  $p = .0061$ ,  $\eta_p^2 = .057$ ], but the Control Group did not ( $p = .64$ ). These effects are illustrated in Figure 12.

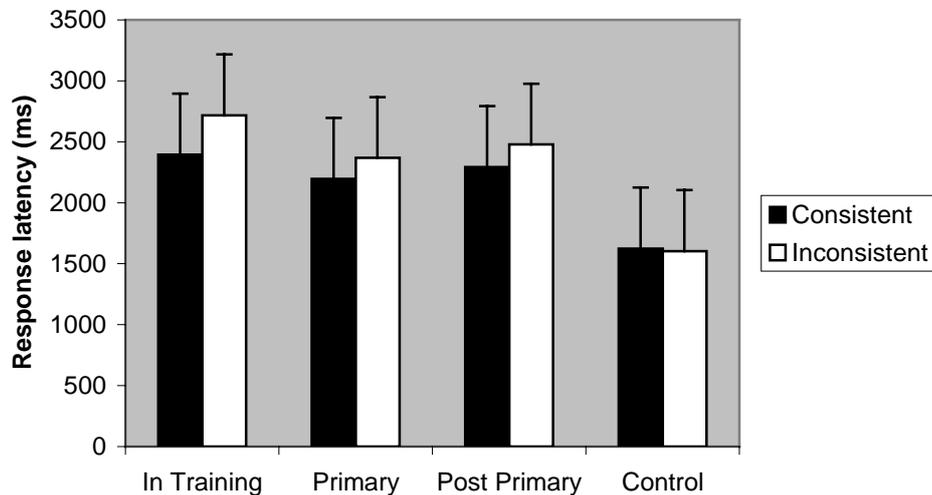


Figure 12: Group means for response latency for consistent and inconsistent responding.

No main effects were recorded for order for any group (all  $p$ 's  $>.34$ ). Blocks was non-significant for three of the groups (all  $p$ 's  $>.43$ ), but was significant for the Primary teachers [ $F(2, 18) = 9.639$ ,  $p = .0062$ ,  $\eta_p^2 = .027$ ]. Post hoc tests (Scheffe) indicated that the significant difference lay between Blocks 1 and 3, probably suggesting practice effects. No significant interaction effects were recorded (all  $p$ 's  $>.19$ ).

#### *Preliminary Trial Analysis*

Trial-type analyses were conducted on the latency data to explore the putative relationship between trial-type and IRAP effect. The overall mean response latencies

for each participant on consistent and inconsistent trials for each of trial-type are presented in Table 18.

TABLE 18.  
Mean Response Latencies (+SE) for Consistent and Inconsistent Responding by Trial-Type

Group	Teacher/Positive		Teacher/Negative		EBD/Positive		EBD/Negative	
	Con	Incon	Con	Incon	Con	Incon	Con	Incon
<i>Control</i>	1503	1639	1698	1686	1761	1677	1727	1744
<i>In-Training</i>	2057	2477	2358	2601	2606	3357	2544	2918
<i>Primary</i>	1870	2085	2075	2247	2419	2497	2280	2634
<i>Post-Primary</i>	1938	2203	2211	2503	2545	2578	2594	2660

The table indicates that across the four groups, participants overall responded faster to TEACHER/POSITIVE in comparison to the other three trial types, although the differences across trial-types were small. Four separate 2x2x2x2 mixed repeated measures ANOVA with sample (TEACHER/EBD), target-type (POSITIVE/NEGATIVE), order and IRAP condition as repeated measures variables were conducted (one for each group). The results of these analyses were identical to those recorded above, with a significant main effect for IRAP condition for the three groups of teachers: In-Training [F (1, 18) = 8.10, p = .0107,  $\eta_p^2 = .249$ ]; Primary [F (1,18) = 11.08, p = .0037,  $\eta_p^2 = .066$ ]; and Post-Primary [F (1,18) = 8.30, p = .0099,  $\eta_p^2 = .054$ ], but not for the Control Group (p > .64). Taken together, these outcomes suggest that the preliminary IRAP effect that had been recorded previously for the teachers was retained during the trial-type analyses.

Sample proved significant for all groups: In-Training [F (1, 18) = 11.35,  $p = .0034$ ,  $\eta_p^2 = .291$ ]; Primary [F (1, 18) = 49.930,  $p < .0001$ ,  $\eta_p^2 = .224$ ]; Post Primary [F (1, 18) = 94.92,  $p < .0001$ ,  $\eta_p^2 = .194$ ]; and Control [F (1, 18) = 16.06,  $\eta_p^2 = .060$ ]. These outcomes indicated that all groups responded differently to TEACHER relative to EBD PUPIL. Interestingly, the largest effect size was for the In-Training teachers (.249); then Primary teachers (.066); then the Control Group (.060); and finally Post-Primary teachers (.054).

Target-type was significant only for the Post-Primary teachers [F (1, 18) = 13.43,  $p = .0018$ ,  $\eta_p^2 = .065$ ], but not for the other three groups (all other  $p$ 's  $> .22$ ). Hence, only these teachers responded differently to POSITIVE versus NEGATIVE. Order was not significant for any group (all  $p$ 's  $> .10$ ), nor were any interaction effects (all  $p$ 's  $> .12$ ).

#### *The D-IRAP Algorithm and Between-Group Comparisons*

Consistent with the analyses conducted on the D scores in Experiment 3, the  $D_{\text{IRAP-(trial-type)}}$  scores were calculated for each participant in Experiment 4 and both types of  $D_{\text{IRAP}}$ -algorithm analyses were employed. These scores are presented in Table 19.

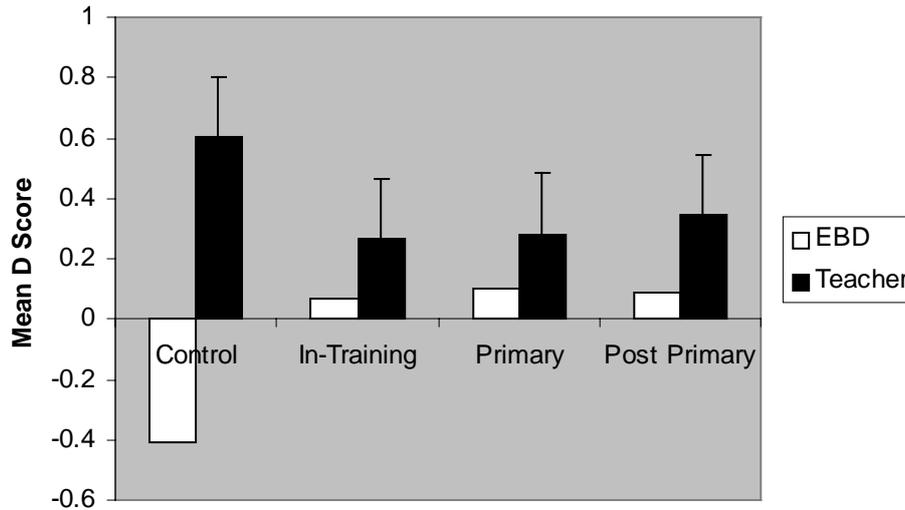
TABLE 19  
 $D_{\text{IRAP-(trial-type)}}$  scores (+SE) for the Four Trial-types

<b>Group</b>	<b>Teacher/Positive</b>	<b>Teacher/Negative</b>	<b>EBD/Positive</b>	<b>EBD/Negative</b>
<i>Control</i>	.267 (.267)	-.022 (.578)	-.194 (.719)	.058 (.585)
<i>In-Training</i>	.342 (.555)	.186 (.587)	.001 (.694)	.137 (.614)
<i>Primary</i>	.294 (.648)	.269 (.769)	.085 (.653)	.119 (.732)
<i>Post-Primary</i>	.357 (.619)	.338 (.772)	.124 (.592)	.053 (.678)

The  $D_{\text{IRAP-(trial-type)}}$  scores recorded with the three groups of teachers indicated that responding was always in the predicted direction (teacher-positive/EBD pupil-negative). Specifically, all three groups of teachers responded strongly to TEACHER/POSITIVE as SIMILAR but not OPPOSITE and to TEACHER/NEGATIVE as OPPOSITE but not SIMILAR. Their attitudes to EBD PUPIL were less strong, although in the predicted direction (i.e. EBD/POSITIVE/OPPOSITE more readily than SIMILAR and EBD/NEGATIVE/SIMILAR more readily than OPPOSITE).

In contrast, some of the trial-type performances of the Control Group differed considerably to the teachers. Similar to the teachers, they showed strong positivity towards TEACHER (i.e. they more readily selected TEACHER/POSITIVE/SIMILAR than OPPOSITE), but did not disconfirm that teachers were negative TEACHER/NEGATIVE (i.e. they more readily selected TEACHER/NEGATIVE/SIMILAR than OPPOSITE), although the size of the D score was very small. Furthermore, the Control Group showed positivity towards EBD PUPIL (i.e. they more readily selected EBD/POSITIVE/SIMILAR than OPPOSITE), although they displayed some agreement with EBD PUPIL/NEGATIVE). Again, the D score was very small.

For illustrative purposes, the trial-types were grouped by sample, in order to highlight potential differences in responding to TEACHER versus EBD PUPIL-- see Figure 13. The teachers and the Control Group showed clear positivity towards TEACHER, but only the latter failed to show clear negativity towards EBD PUPIL.




---

*Figure 18: Group mean D scores for trial-types.*

---

The  $D_{\text{IRAP-(trial-type)}}$  scores were entered into a 2x3x4 mixed repeated measures ANOVA, with group as the between participant variable and blocks and trial-type as the within participant variables. The analysis revealed an approaching significant main effect for group [ $F(3, 76) = 2.61, p = .0571, \eta_p^2 = .103^1$ ] and a significant main effect for trial-type [ $F(3, 76) = 9.50, p < .0001, \eta_p^2 = .255$ ]. However, there was no significant main effect for blocks ( $p = .98$ ) and no interaction effects (all  $p$ 's  $> .26$ ).

Post hoc tests (Scheffe) indicated that the significant trial-type differences lay between TEACHER/POSITIVE and EBD PUPIL/POSITIVE ( $p < .0001$ ); TEACHER/POSITIVE and EBD PUPIL/NEGATIVE ( $p = .0049$ ); and TEACHER/NEGATIVE and EBD/NEGATIVE ( $p = .0258$ ).

*Within-Group Comparisons*

Four separate 1x2x4 mixed repeated measures ANOVA with blocks and trial-type as the within participant variables were conducted (one for each group), in order to examine the trial-type effects for each group. No significant main effects for blocks were recorded for any group (all p's > .14). Trial-type was found to be non-significant for both Primary and Post-Primary teachers (all p's >.0506), but was significant for both the Control Group [F (3, 19) = 6.49, p =.0007  $\eta_p^2$  =.416) and the In-Training teachers ( [F (3, 19) = 2.91 , p = .0417  $\eta_p^2$  = .296). Post hoc tests (Scheffe) indicated that the significant trial-type differences for both groups lay between TEACHER/POSITIVE and EBD PUPIL/POSITIVE (Control Group: p =.0009 and In-Training teachers: p = .0451).

Four separate one-group t-tests (one per group) were conducted to determine whether each of the D-IRAP scores for the four individual trial-types differed significantly from 0 -- see Table 20.

TABLE 20  
T-test Results for D<sub>IRAP-(trial-type)</sub> Scores for All Groups

	<b>TEACHER /POSITIVE</b>	<b>TEACHER /NEGATIVE</b>	<b>EBD PUPIL /POSITIVE</b>	<b>EBD PUPIL /NEGATIVE</b>
<i>Control</i>	p <. 0023	p=.7744	p< .0411	P>. 4479
<i>In-Training</i>	p<. 0001	p<.0169	p>. 9913	p>. 0897
<i>Primary</i>	p<. 0008	p<. 0088	p >. 2111	p>. 2111
<i>Post-Primary</i>	p<.0001	P<. 0012	p>.1102	p>.5467

These analyses revealed a significant result for all groups for TEACHER/POSITIVE. TEACHER/NEGATIVE was also significant for the three groups of teachers, but not for the Control group. EBD/POSITIVE was non-significant for the three groups of teachers, but was significant for the Control group.

EBD PUPIL/NEGATIVE was non-significant for all groups, although the In-Training group approached significance.

*Explicit Measure*

Prior to the exposure to the IRAP, the three groups of teachers were exposed to the ORM Scale as an explicit measure, which comprised of three basic components. The results for each group on each of the three sub-scales are presented in Table 21.

*Years Experience.* A demographic questionnaire was also presented to Primary and Post-Primary teachers in relation to the number of years they had been engaged in mainstream teaching. The total mean score for each group was calculated and the results indicated that Primary teachers had 10 years and 3 months mean experience – considerably less than the mean of 16 years and 5 months for the Post-Primary teachers.

TABLE 21  
The Results of the ORM Scale.

<b>ORM Scale</b>			
<b>Attitudes to Inclusion</b>			
	In-Training	Primary	Post-Primary
<b>Strongly Disagree/Disagree</b>	-	-	-
<b>Undecided</b>	10%	20%	5%
<b>Strongly Agree/Agree</b>	90%	80%	95%

**Emotional Reaction Scale 1**  
*Reactions to the inclusion of EBD child in participant's classroom*

	In-Training	Primary	Post-Primary
<b>Extremely Negative/Negative</b>	-	10%	15%
<b>Neutral</b>	30%	45%	25%
<b>Positive/Extremely Positive</b>	70%	45%	60%

**Emotional Reaction Scale 2**  
*Exclusion of child from other school to be present in participant's classroom*

	In-Training	Primary	Post-Primary
<b>Extremely Negative/Negative</b>	25%	40%	35%
<b>Neutral</b>	25%	30%	40%
<b>Positive/ Extremely Positive</b>	50%	30%	25%

**Willingness to Adapt Teaching Practice**

	In-Training	Primary	Post-Primary
<b>Strongly Disagree/ Disagree</b>	5%	10%	10%
<b>Undecided</b>	45%	45%	65%
<b>Agree/Strongly Agree</b>	50%	45%	25%

(1) *Attitudes to Inclusion.* On the Likert Scale assessing teachers' attitudes to the inclusion of children with EBD, all three groups were predominantly positive (see top row of Table 21), with 90% of the overall sample in favour. No-one explicitly indicated that they disagreed with inclusion, although approximately 10% (mostly Primary teachers) were undecided.

(2) *Emotional Reaction Scale 1: Inclusion of EBD Child in Participant's Classroom.* In spite of their positive attitudes to the principles of inclusion, all three groups of teachers showed more negative reactions to the possibility of having EBD

children directly placed within their classrooms (second row of Table 21).

Specifically, only around 55% were in favour of inclusion in this context, 35% were neutral (mostly Primary teachers), and 10% were explicitly negative (included mostly both Primary and Post-Primary teachers).

*(2a) Emotional Reaction Scale 2: Inclusion of Child Previously Excluded.*

Because teachers had been reticent when asked about the direct inclusion of an EBD child in their own classrooms, it was not surprising when they reported even greater negativity at the prospect of including a child with EBD who had previously been excluded from another educational facility (third row of Table 21). Indeed, only around 35% were positive about such a prospect, 25% were neutral, and as many as 40% were overtly negative (mostly Primary and Post-Primary teachers).

Interestingly, when comparing responding to sub-scales 1 and 2, the Post-Primary teachers were more negative about an excluded child (scale 2) than about an EBD child who had not been excluded (scale 1). In contrast, the inclusion of such a child rendered the Primary teachers more negative than the inclusion of a non-excluded child.

*(2c) What extra things would you need to make your responses to the above three questions more positive?* Teachers from both groups were unanimous in their requests for support within the classroom situation, including designated SEN and traditional classroom assistants, and resource teaching. Four other areas of support were identified as essential to the mainstream inclusion of this population. (1) Standard class sizes were deemed an impediment to the inclusion of more than one or two children with EBD in any one classroom. (2) At a professional level, a multidisciplinary approach was perceived as crucial for keeping teachers informed of a child's progress in other areas (i.e. behavioural therapy, medication,

etc.), with particular emphasis placed on the need for relevant information *prior to* the child's introduction in a particular classroom. (3) At a more personal level, teachers stressed the need for support from closer colleagues, particularly school principals and parents. (4) In-service training was fundamental to meeting the children's academic needs because their learning difficulties impact upon their ability to cope in a mainstream classroom and on their behaviour therein.

One other key issue was raised by practically all of the Post-Primary teachers regarding SEN inclusion at this level. These teachers expressed serious concerns about their physical safety, particularly with teenagers with EBD. In order to deal with this issue effectively, they identified a number of areas of critical support. (5) The physical presence of relevant others (especially SEN assistants) was central to managing disruptive behaviour. (6) A whole school policy that governed disciplinary procedures for all students was essential and should be agreed by both parents and children *before* commencement at that school. (7) Behavioural support units were deemed to be beneficial to both teachers and students (e.g. in the provision of time-out facilities).

(3) *Willingness to Adapt Teaching Practices.* Only 40% of teachers indicated that they would be willing to adjust their teaching practices to facilitate inclusion of an EBD child, 50% were neutral and 10% were unwilling (bottom row of Table 22). The teachers who were neutral or negative were well dispersed across the three groups, hence it was perhaps most surprising that the In-Training teachers were no less willing than experienced teachers to adapt their teaching practices in this regard.

*Explicit Likert Analyses*

All participants were presented with a Likert Scale that required them to provide explicit ratings of the IRAP words in order to determine the relationship between their explicit and implicit attitudes in this regard. The mean Likert ratings for each of the positive and negative words for the four groups are presented in Table 22. The ratings were generally as expected, with positive ratings for positive words and negative ratings for negative words. In fact, the overall mean rating for Positive words was 3.343 and -1.928 for Negative words (not shown). Unexpectedly, the Control Group provided negative ratings for CALM.

TABLE 22  
Mean Ratings for Positive and Negative IRAP Words.

<b>Group</b>	<b>Positive Words</b>					
	ACCOMMODATING	POSITIVE	SUITABLE	CALM	PLEASANT	CO-OPERATIVE
<i>Control</i>	4.11	4.69	3.37	-1.632	2.79	3.95
<i>In-Training</i>	2.00	4.40	2.10	3.50	3.50	3.85
<i>Primary Teachers</i>	2.50	4.95	2.40	4.25	4.15	4.65
<i>Post-Primary Teachers</i>	2.30	4.70	1.85	4.75	3.10	4.25
<b>Negative Words</b>						
	NON-COMPLIANT	NEGATIVE	ANGRY	DIFFICULT	UNPLEASANT	INAPPROPRIATE
<i>Control</i>	-2.80	-3.65	-4.40	-2.10	-4.05	-2.75
<i>In-Training</i>	-1.05	-2.15	-2.05	-.300	-1.56	-1.76
<i>Primary Teachers</i>	-2.68	-3.05	-2.53	-.842	-2.47	-1.63
<i>Post-Primary Teachers</i>	-2.95	-2.25	-1.80	-.850	-2.25	.450

Four separate one-group t-tests (one per group) were conducted to determine whether there were significant differences in the ratings for positive versus negative words as provided by each group. As expected, the results revealed a significant difference for all groups (all p's <.0001).

## DISCUSSION

The IRAP effects observed in the current study indicated that the three groups of teachers showed strong implicit bias towards teachers. Although their attitudes to EBD PUPIL were weaker, they were predictably negative. A somewhat different IRAP profile was observed with the Control Group. Although they also indicated that teachers were positive and EBD pupils were negative, they did not disconfirm that teachers were negative and confirmed that EBD pupils were positive

The outcomes on the explicit ORM Scale indicated that although teachers were generally in favour of the inclusion of EBD children into mainstream education, they were more reticent when asked about the placement of this population within their own classrooms. The Primary teachers were even more reticent when this inclusion involved a child who had been previously excluded from another educational facility. Interestingly, all four groups (including the In-Training Teachers) expressed considerable unwillingness to adapt their teaching practices to foster such inclusion.

Taken together, therefore, the teachers showed considerably negative attitudes to children with EBD on both the implicit and explicit measures. Although the explicit measure specifically targeted inclusion, this offered a useful index of teachers' attitudes to this sample and how willing they would be to accommodate them within their direct educational environments. Although the Control Group did not complete the explicit measure, the IRAP showed a somewhat differentiated profile relative to teachers that was characterised by less negativity towards EBD pupils. The ability of the implicit measure to differentiate in this way among the two types of respondent was particularly positive.

One issue that emerged from Experiment 4 concerned the possibility that what the IRAP was tapping into was not in fact the difference between the teachers' implicit attitudes to teachers relative to EBD pupils, but was simply the difference in attitudes to teachers versus all pupils. In other words, it remained possible that teachers would have equally negative implicit attitudes to normally-developing pupils. Hence, the IRAP outcomes reported here could not be interpreted as having specific relevance to EBD children. This issue was the focus of Experiment 5.

## Chapter 6

# Examining the Utility of the IRAP in the Study of Teachers' Implicit Attitudes to Normally-Developing Children

## **Chapter 6: Experiment 5**

### **Examining the Utility of the IRAP in the Study of Teachers' Implicit Attitudes to Normally-Developing Children**

#### INTRODUCTION

The IRAP effects observed in Experiment 4 indicated that teachers from three different levels of the education sector demonstrated, as expected, consistently positive bias towards teachers that contrasted strongly with reticence and negativity towards pupils with EBD. These views were not entirely consistent with non-teachers whose implicit attitudes to both teachers and EBD pupils were mixed, but did indicate greater negativity towards teachers and more positivity towards EBD pupils than was recorded with the teachers.

There was considerable overlap between the teachers' implicit attitudes to pupils with EBD and their explicit opinions regarding the inclusion of this group in mainstream education, measured by the ORM Scale. Although the teachers expressed general favour for this policy, they were reticent about the placement of these children within their own classrooms and were unwilling to adapt their teaching practices to facilitate this.

One issue that emerged from Experiment 4 concerned the possibility that what the IRAP was tapping into was not in fact the difference between the teachers' implicit attitudes to teachers relative to EBD pupils, but was simply the difference in attitudes to teachers versus all pupils. In other words, it remained possible that teachers would have equally negative implicit attitudes to normally-developing pupils. Hence, the IRAP

outcomes reported previously could not be interpreted as having specific relevance to EBD children. This issue was the focus of Experiment 5.

## METHOD

### *Participants*

Forty adults participated voluntarily in Experiment 5. These individuals were recruited by the Researcher through contacts with Primary and Post-Primary schools in the Dublin and Kildare regions. The participants were divided into two groups that were differentiated in terms of their levels and type of educational training and experience with children. The *Primary Teachers* ( $n = 20$ ) were all currently employed in Primary schools; had a formal qualification in teaching; and had at least two years teaching experience in mainstream education. The *Post-Primary Teachers* ( $n = 20$ ) were currently employed in Post-Primary schools; had a formal qualification in teaching; and had at least two years teaching experience in mainstream education.

### *Setting*

For both groups, all aspects of Experiment 5 were conducted in a quiet room in each of the respective schools. In all cases, the experimental setting remained free from noise and distraction throughout all aspects of the study. All participation was conducted on an individual basis. The Researcher remained present in the room at all times.

### *Apparatus and Materials*

The current study involved the same basic sets of materials/apparatus as Experiment 4. The stimuli employed here for the IRAP were identical in format to Experiment 4, except that the sample stimuli in this case contained the words TEACHER

and PUPIL. The target stimuli (six positive and six negative words) and the relational terms (SIMILAR and OPPOSITE) were the same. The stimulus arrangements for Experiment 5 are presented in Table 23.

TABLE 23  
The Stimulus Arrangements Employed in Experiment 5.

<b>Sample 1</b> TEACHER	<b>Sample 2</b> PUPIL
<b>Response Option 1</b> SIMILAR	<b>Response Option 2</b> OPPOSITE
<b>Targets deemed consistent with Sample 1</b>	<b>Targets deemed consistent with Sample 2</b>
CO-OPERATIVE SUITABLE POSITIVE CALM ACCOMODATING PLEASANT	DIFFICULT NON-COMPLIANT ANGRY NEGATIVE UNPLEASANT INAPPRORIATE

The screen shots employed here as visual aids were identical in format to the previous experiments, but were simply modified to incorporate the new sample stimuli.

#### *Procedure*

All aspects of the procedure employed in Experiment 5 were identical to Experiment 4, except that the two groups here completed the ORM *after* the IRAP and the instructions presented to participants were modified in accordance with the experimental stimuli. All participants completed the study in a single experimental session that lasted between 20 and 30mins.

*Trial-types.* The four basic trial-types presented here were referred to as TEACHER/POSITIVE; PUPIL/NEGATIVE; PUPIL/POSITIVE; and TEACHER/NEGATIVE -- see Figure 14. Each block of IRAP trials comprised of twelve

exposures to each of the four trial-types, which contained two exposures to each of the twelve target words (e.g. ANGRY appeared twice with TEACHER). Correct responses on consistent trials involved positivity towards TEACHER and negativity towards PUPIL, whereas correct responding on inconsistent trials was reversed (i.e. TEACHER-NEGATIVE and PUPIL-POSITIVE).

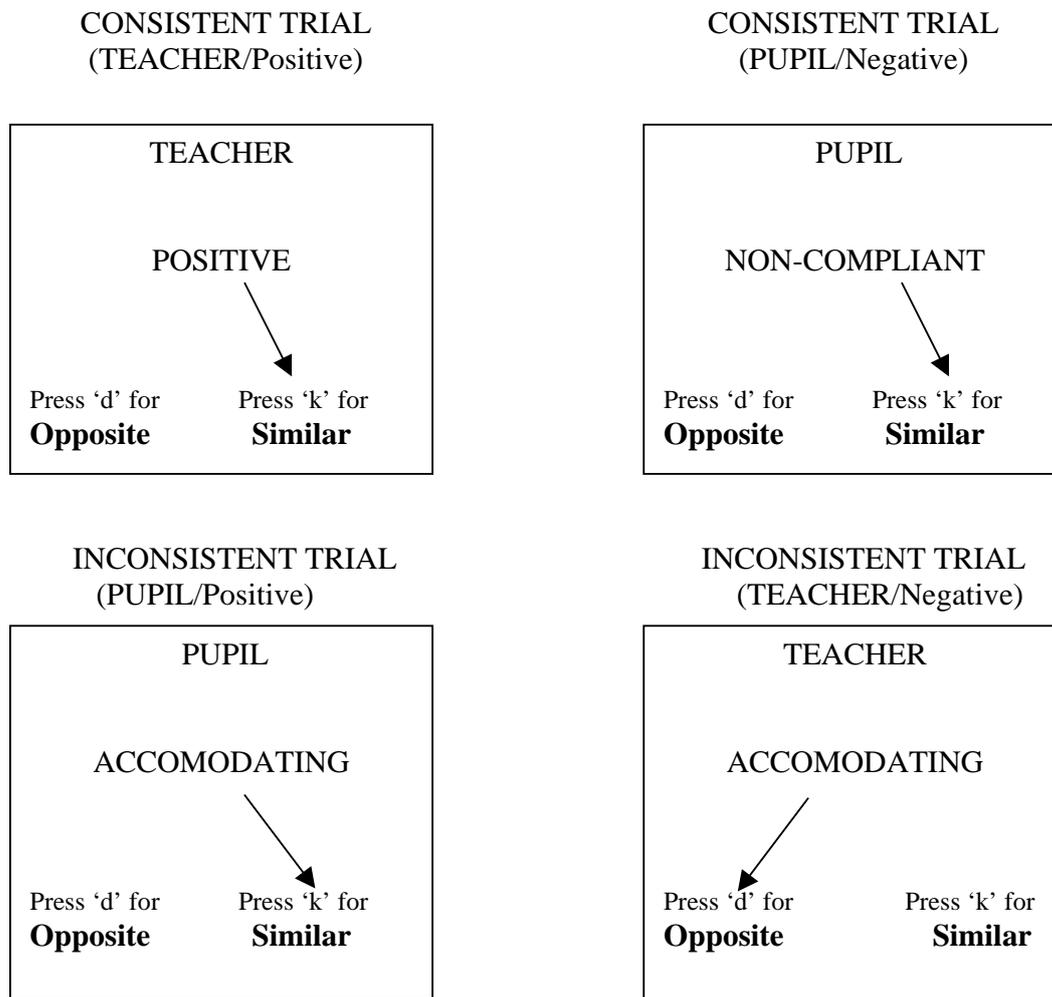


Figure 14: Schematic representation of the four basic trial-types in Experiment 5.

## RESULTS

### *Response Latency Data*

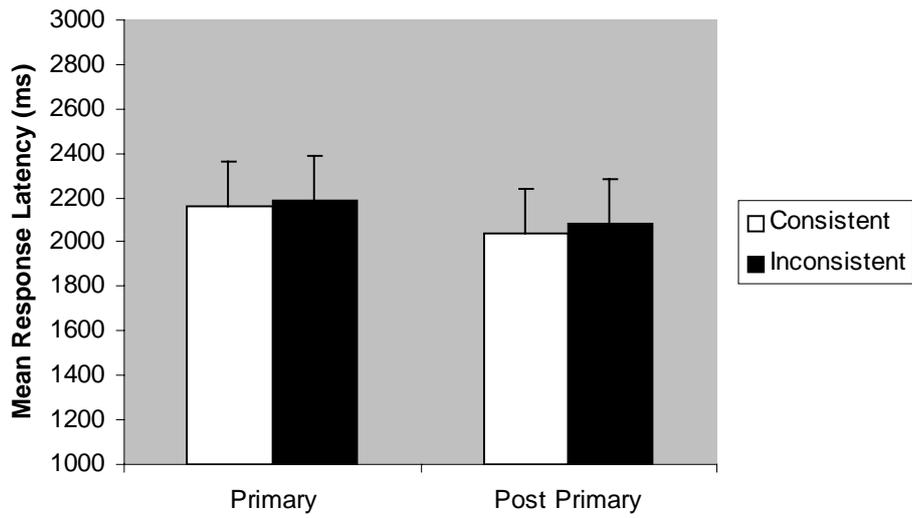
Mean response latencies were calculated (three consistent and three inconsistent) (after adjustments) for each participant on each of the six test blocks and were then collated by group --- see Table 24. On most test blocks, both groups of teachers demonstrated IRAP effects, although the differences between consistent and inconsistent responding were frequently small. Only on Blocks 5 and 6 the Primary Teachers responded more quickly on inconsistent trials over consistent (i.e. reverse IRAP effect).

TABLE 24  
The Mean Response Latencies for Both Groups on  
Consistent and Inconsistent Responding.

<b>Group</b>	<b>Blocks 1 and 2</b>		<b>Blocks 3 and 4</b>		<b>Blocks 5 and 6</b>	
	Con	Incon	Con	Incon	Con	Incon
<i>Primary</i>	2203	2268	2184	2229	2093	2068
<i>Post-Primary</i>	2026	2070	2035	2119	2047	2052

A 2x3x2 mixed repeated measures ANOVA was conducted on the latency data for each group, with IRAP condition and blocks as within participant variables and order as the between participant variable. Interestingly, neither group showed main effects for IRAP condition (Primary:  $p = .10$  and Post-Primary  $p = .11$ ) or order (Primary:  $p = .20$  and Post-Primary:  $p = .91$ ). For illustrative purposes, the IRAP effects are depicted in Figure 15. Although blocks were non-significant for the Post-Primary teachers ( $p = .73$ ), it was unexpectedly significant for the Primary teachers [ $F(2, 18) = 10.178, p = .0003$ ,

$\eta_p^2 = .153$ ]. In the latter case, Post-hoc tests (Scheffe) indicated that the significant difference lay between Blocks 1 and 2 and between Blocks 3 and 4, probably suggesting practice effects. No significant interaction effects were recorded (Primary,  $p = .79$ ; and Post-Primary  $p = .65$ ).



---

Figure 15: Group mean response latencies for consistent and inconsistent responding.

---

### *Preliminary Trial Analyses*

Trial-type analyses (TEACHER/POSITIVE; TEACHER/NEGATIVE; PUPIL/POSITIVE; and PUPIL/NEGATIVE) were conducted on the latency data to explore the putative relationship between trial-type and the IRAP effects noted above. The mean response latencies for both groups on consistent and inconsistent trials for each of the four trial-types are presented in Table 25. Both groups of teachers responded

fastest overall to TEACHER/POSITIVE relative to the other three trial-types, across which similar response patterns were observed.

TABLE 25  
The Mean Response Latencies for Consistent and Inconsistent Responding  
By Trial-type for Each Group.

Group	TEACHER/POSITIVE		TEACHER/NEGATIVE		PUPIL/POSITIVE		PUPIL/NEGATIVE	
	Con	Incon	Con	Incon	Con	Incon	Con	Incon
<i>Primary</i>	2076	2530	2303	2467	2580	2460	2533	2644
<i>Post-Primary</i>	2089	2131	2324	2446	2519	3110	2337	2441

A 2x2x2x2 mixed repeated measures ANOVA was conducted with sample (TEACHER/PUPIL), target-type (POSITIVE/NEGATIVE), order and IRAP condition as repeated measures variables for each group. The results revealed a significant effect for IRAP condition for the Primary Teachers [ $F(1, 18) = 10.566, p = .0044, \eta_p^2 = .071$ ], but not for Post-Primary ( $p = .37$ ). Similarly, sample was significant for Primary, [ $F(1, 18) = 10.548, p = .0045, \eta_p^2 = .135$ ], but not for Post-Primary ( $p = .74$ ). Target-type was non-significant for both groups (Primary:  $p = .10$  and Post-Primary:  $p = .12$ ) and there were no significant interaction effects (Primary:  $p = .62$  and Post-Primary:  $p = .52$ ).

In order to examine the sample effect obtained for the Primary teachers, two separate one-way repeated measures ANOVA were conducted -- one for PUPIL, the other for TEACHER. The analyses revealed a significant main effect for TEACHER ( $p < .0001$ ), but not for PUPIL ( $p = .21$ ).

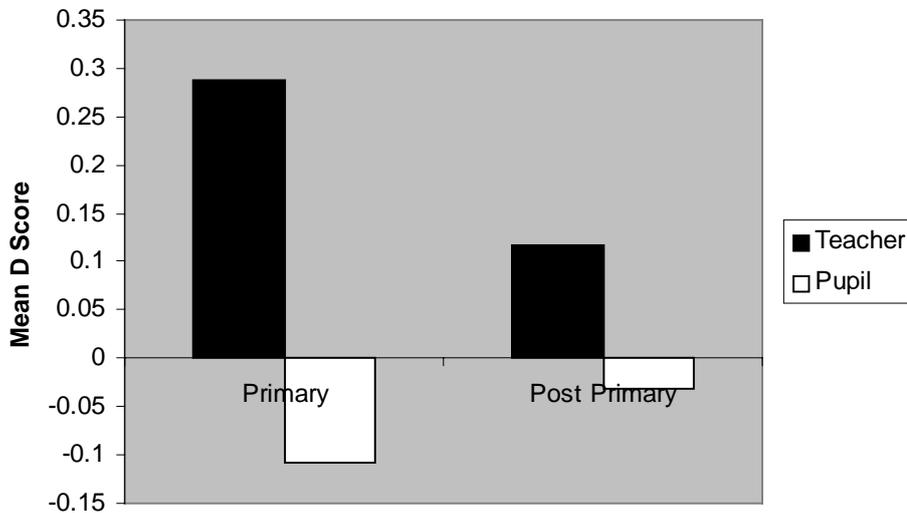
*The D-IRAP Algorithm and Within-Group Comparisons*

Consistent with the analyses conducted previously on the D scores, both types of  $D_{IRAP}$ -algorithm analyses were employed. The mean  $D_{IRAP-(trial-type)}$  scores recorded for each group are presented in Table 26. The patterns of responding indicated by the  $D_{IRAP-(trial-type)}$  scores were very similar for both groups. Both responded as predicted on TEACHER/POSITIVE and TEACHER/NEGATIVE, although the Primary Teachers produced a particularly strong D score on the former relative to the Post-Primary Teachers. The groups were again consistent with each other and with predictions in their responses to PUPIL/NEGATIVE, although the D scores in both cases were very small. Unexpectedly, however, both groups were inconsistent with predictions on PUPIL/POSITIVE, indicating that they more readily responded with PUPIL/POSITIVE/SIMILAR than OPPOSITE.

TABLE 26  
The  $D_{IRAP-(trial-type)}$  scores (+SE) for the Four Trial-types by Group.

Group	TEACHER/POSITIVE	TEACHER/NEGATIVE	PUPIL/POSITIVE	PUPIL/NEGATIVE
<i>Primary</i>	.421 (.644)	.155 (.654)	-.215 (.667)	.001 (.743)
<i>Post-Primary</i>	.147 (.725)	.147 (.641)	-.150 (.633)	.071 (.095)

For illustrative purposes, the trial-types were grouped by sample, in order to highlight differences in responding to TEACHER versus PUPIL -- see Figure 16. The figure illustrates, as expected, that both groups of teachers, particularly the Primary Teachers, showed strong positivity towards TEACHER. The figure also indicates that their attitudes to PUPIL were inconsistent with predictions, and thus were not as negative as was expected.




---

Figure 16: Group responses on  $D_{IRAP-(trial-type)}$  scores by sample.

---

The  $D_{IRAP-(trial-type)}$  scores were entered into a  $2 \times 3 \times 4$  mixed repeated measures ANOVA, with group as the between participant variable and blocks and trial-type as the within participant variables. The analysis revealed no significant main effect for group ( $p = .49$ ) or Blocks ( $p = .39$ ), although a significant main effect was recorded for trial-type. No interaction effects were recorded (all  $p$ 's  $> .68$ ). Post hoc tests (Scheffe) indicated that the significant differences for trial-type lay between: TEACHER/POSITIVE and PUPIL/POSITIVE ( $p < .0001$ ); and between TEACHER/NEGATIVE and PUPIL/NEGATIVE ( $p = .0142$ ).

### *Within-Group Comparisons*

Two separate 1x3x4 mixed repeated measures ANOVA's, one for each group, were conducted with blocks and trial-type as within participant variables. The analysis revealed no significant main effect for blocks for both groups (all  $p$ 's  $>.23$ ). Trial-type was non-significant for the Post-Primary Teachers ( $p = .15$ ), but significant for Primary Teachers ( $p < .0001$ ). Post hoc tests (Scheffe) indicated that the significant differences for trial-type for this group lay between: TEACHER/POSITIVE and PUPIL/POSITIVE ( $p < .0001$ ); TEACHER/NEGATIVE and PUPIL/NEGATIVE ( $p = .0109$ ) and for TEACHER/NEGATIVE and PUPIL/POSITIVE ( $p = .0318$ ).

Four separate one-group t-tests (one per group) were conducted to establish whether each of the D-IRAP scores for the four trial-types differed significantly from 0 -- see Table 27. In line with the trial-type analyses above, TEACHER/POSITIVE was significant for the Primary Teachers but not for Post-Primary. TEACHER/NEGATIVE approached significance for both groups. PUPIL/POSITIVE was significant for the Primary Teachers and approached significance for Post-Primary (responding here was in the non-predicted direction). PUPIL/NEGATIVE was not significant for either group.

TABLE 27  
The T-test Results By Trial-type for Both Groups.

<b>Group</b>	<b>TEACHER/POSITIVE</b>	<b>TEACHER/NEGATIVE</b>	<b>PUPIL/POSITIVE</b>	<b>PUPIL/NEGATIVE</b>
<i>Primary</i>	$P < .0001$	$p > .0712$	$p < .0155$	$p = .9925$
<i>Post-Primary</i>	$p > .1218$	$p > .0802$	$p < .0721$	$p > .4608$

Taken together, these findings indicated that there were some differences between the Primary and Post-Primary Teachers' attitudes to teachers and pupils. The Primary

teachers showed particularly strong positivity towards teachers that was not shared to the same extent by the Post-Primary teachers (i.e. they confirmed that teachers were similar to positive and opposite to negative). In contrast, the teachers' implicit attitudes to pupils were mixed when they confirmed that pupils were positive but also confirmed that pupils were negative, although the latter D score was very small.

#### *Explicit Measure*

Immediately after the exposure to the IRAP, both groups of participants were exposed to the ORM Scale. The results of each of the three ORM sub-scales for both groups are presented in Table 28.

*Years Experience.* The two groups of teachers had extensive teaching experience, with a mean of 20 years recorded for the Primary Teachers and 16 years/6 months for the Post-Primary teachers.

TABLE 28  
The Results of the ORM Scale.

<b>ORM Scale</b>		
<b>Attitudes to Inclusion</b>		
	Primary	Post-Primary
<b>Strongly Disagree/Disagree</b>	-	-
<b>Undecided</b>	10%	20%
<b>Agree/Strongly Agree</b>	90%	80%

<b>Emotional Reaction Scale 1</b>		
<i>Measured reactions to the inclusion of EBD child in participant's classroom</i>		
	Primary	Post-Primary
<b>Extremely Negative/Negative</b>	30%	10%
<b>Neutral</b>	55%	65%
<b>Positive/Extremely Positive</b>	15%	25%

<b>Emotional Reaction Scale 2</b>		
<i>Exclusion of child from other school to be present in participant's own classroom</i>		
	Primary	Post-Primary
<b>Extremely Negative/Negative</b>	65%	55%
<b>Neutral</b>	30%	40%
<b>Positive/Extremely Positive</b>	5%	5%

<b>Willingness to Adapt Teaching Practice</b>		
	Primary	Post-Primary
<b>Strongly Disagree/Disagree</b>	10%	-
<b>Undecided</b>	60%	70%
<b>Agree/Strongly Agree</b>	30%	30%

(1) *Attitudes to Inclusion.* Over 80% of both groups of teachers were positive about the general concept of inclusion (top row of Table 29). No-one explicitly indicated that they disagreed, although approximately 15% of the total sample (mostly Post-Primary Teachers) was undecided.

(2) *Emotional Reaction Scale: Inclusion of EBD Child in Participant's Classroom.* In spite of their positive attitudes to the principle of inclusion, both groups of teachers reacted more negatively to the possibility of having EBD children directly

placed within their classrooms (second row of Table 29). Specifically, only around 20% of the total sample, were in favour of inclusion in this context, 60% were neutral (mostly Post-Primary) and 20% were negative.

*(2a) Emotional Reaction Scale 2: Inclusion of Child Previously Excluded.* Even greater negativity was expressed when both groups were asked about the prospect of including a child with EBD who had previously been excluded from another educational facility (third row of Table 29). Indeed, only around 5% of the sample was positive, 35% were neutral, and 60% were negative.

*(2c) What extra things would you need to make your responses to the above three questions more positive?* As in the previous study, teachers from both groups were unanimous in their requests for support within the classroom situation, including designated SEN and traditional classroom assistants, and resource teaching. The areas identified in Chapter 5 were also highlighted in the current study, with the areas of support being highlighted as (1) reduction of class sizes; (2) the establishment of a multidisciplinary approach; (3) The necessity of developing a whole-school approach in dealing with school discipline problems in relation to children with EBD; (4) The opportunity to meet with the students in question (5) The establishment of an effective counselling service for both students and their families, while In-service training was also implicated as being fundamental to meeting the children's needs.

*(3) Willingness to Adapt Teaching Practices.* Across the two groups, 30% of teachers indicated that they would be willing to adjust their teaching practices to facilitate this type of inclusion, 65% were neutral and only 5% were unwilling (bottom row of Table 29).

### *Explicit Likert Analyses*

The mean Likert ratings for each word were calculated for each group and these are presented in Table 29. The ratings were generally as expected, with positive ratings for positive words and negative ratings for negative words. In fact, the overall mean rating for Positive words 3.52 and -1.85 for Negative words. Two separate one-group t-tests (one per group) indicated that the ratings for positive versus negative words differed significantly for each group ( $p < .0001$ ).

TABLE 29  
The Group Mean Likert Ratings for Positive and Negative Words.

<b>Positive Words</b>						
<b>Group</b>	ACCOMMODATING	POSITIVE	SUITABLE	CALM	PLEASANT	CO-OPERATIVE
<i>Primary</i>	3.11	4.57	2.32	4.53	3.48	4.21
<i>Post-Primary</i>	2.22	4.67	1.72	3.61	3.83	3.83
<b>Negative Words</b>						
	NON-COMPLIANT	NEGATIVE	ANGRY	DIFFICULT	UNPLEASANT	INAPPROPRIATE
<i>Primary</i>	-2.65	-2.60	-1.65	-1.75	-2.40	-1.55
<i>Post-Primary</i>	-1.75	-1.95	-1.65	-1.20	-2.90	-1.15

## DISCUSSION

The IRAP effects observed in the current study indicated that Primary and Post-Primary teachers showed implicit bias towards teachers, although this effect was much stronger for the former relative to the latter. The IRAP also indicated mixed attitudes towards pupils, in which they confirmed that they were positive but also agreed that they were negative. Both groups of teachers were similar in this regard.

The outcomes on the ORM Scale indicated that although these teachers were generally in favour of the inclusion of EBD children into mainstream education, they were more reticent when asked about the placement of this population within their own classrooms. Both groups were between willing and neutral in terms of the adaptation of their teaching practices.

The primary aim of Experiment 5 was to determine whether similar IRAP outcomes would be recorded with PUPIL and EBD PUPIL (Experiment 4), when each was compared with TEACHERS. In Experiment 4, all three groups of teachers showed strong implicit bias towards TEACHERS and weaker negative attitudes to EBD PUPIL. Interestingly, the Control Group of non-teachers had more mixed reactions to both types of stimulus sample. Although the teachers who participated in Experiment 5 showed equally strong positive bias towards TEACHER as recorded previously, they had mixed reactions to PUPIL. The difference, therefore, between the two studies indicated that teachers had more negative reactions to EBD PUPIL than to PUPIL. As a result, it is reasonable to conclude that the EBD outcomes recorded in Experiment 4 did not solely result from teachers' generic attitudes to pupils, but were specific (at least to some extent) to EBD.

## Chapter 7

### Systematic Comparisons among the IRAP Studies of Childhood and Adulthood

## **Chapter 7**

### **Systematic Comparisons Among The IRAP Studies Of Childhood and Adulthood**

The research conducted thus far comprised of a series of five experiments that might loosely be divided into two sections – studies with children (Experiments 1-3) and studies with adults/teachers (Experiments 4 and 5). Although the studies within each of the two divisions were conducted separately, it is possible to get a broader view of the work by drawing comparisons among the relevant experiments. For example, it would be useful to compare all three groups of children sampled in order to compare the profiles in terms of implicit and explicit self-esteem of normally developing children with those diagnosed with Dyslexia and with ADHD (i.e. by comparing Experiments 2 and 3). Similarly, one would be able to compare the teachers' implicit attitudes to normally-developing children (in terms of their reactions to the word PUPIL) with children with EBD by comparing Experiments 4 and 5. These two issues constitute the basic aims of the current chapter, which is divided accordingly into two parts. In Part 1, Experiments 2 and 3 are compared in an analysis of the childhood studies of gender and self-esteem. In Part 2, Experiments 4 and 5 are compared in an analysis of teachers' attitudes to normally-developing versus EBD children. Each of these parts is presented separately below.

## PART 1

### Childhood Comparisons from Experiments 2 and 3

Although Experiments 2 and 3 each comprised numerous types of analyses conducted on both the implicit and explicit outcomes, there are several critical dimensions along which comparisons can be drawn between the two if one wishes to identify potential differences between the three target groups of children – normally-developing, Dyslexic and ADHD. These include: the preliminary trial-type data; the  $D_{\text{IRAP-(trial-type)}}$  scores; the significance of the  $D_{\text{IRAP-(trial-type)}}$  scores relative to zero; overall D scores; and the Piers Harris 2 explicit measure. Each of these is described separately below.

#### *Preliminary Trial-Type Data*

Although the Normally-Developing children had participated as three separate samples in Experiments 1-3, the data from all three were collated as a single normally-developing sample for the purposes of comparison with the two other groups of children. The mean response latencies for consistent and inconsistent responding by trial-type for the three groups of children are presented in Table 30. Overall, the Normally-Developing participants produced the fastest response latencies, the ADHD children produced the slowest and the Dyslexic children were in between the two. These overall differences may, of course be due simply to the groups' differential levels of experience with computers, although there was no additional anecdotal evidence to support this. On the contrary, the differences, especially in the ordering of the three groups, may result from subtle differences in

their implicit self-esteem. Indeed, this latter supposition is supported by further analyses of the data.

TABLE 30  
Mean Response Latencies for Consistent and Inconsistent Responding By Trial-Type and Group.

Group	OWN NAME /POSITIVE		OWN NAME /NEGATIVE		OPPOSITE GENDER NAME/POSITIVE		OPPOSITE GENDER NAME/NEGATIVE	
	Con	Incon	Con	Incon	Con	Incon	Con	Incon
Normally-Developing	1849	2219	2063	2156	2303	2226	2351	2397
Dyslexic	2342	2851	2306	2692	2473	2548	2686	2626
ADHD	2650	2614	2510	3018	2652	2675	3058	3078

*The  $D_{IRAP-(trial-type)}$  Scores*

The  $D_{IRAP-(trial-type)}$  scores for each group of children are presented in Table 31. The normally-developing children responded in the predicted direction on all four trial-types. In contrast, while the Dyslexic and ADHD children also responded in the predicted direction on three of the trial-types, they both responded in the non-predicted direction for OTHER/NEGATIVE. This latter finding indicated that these children did not demonstrate negativity towards others that was shown by the normally-developing children. Although all three groups of children responded as predicted on the critical OWN NAME-POSITIVE trial-type, showing positivity towards the self, the size of the D score recorded with the ADHD children was almost six times smaller than that recorded for the other two groups. This was the first key evidence that the ADHD children had lower implicit self-esteem than the normally-developing and dyslexic children.

TABLE 31  
The  $D_{IRAP-(Trial-Type)}$  Scores (+SE) For The Four Trial-Types by Group.

Group	OWN NAME /POSITIVE	OWNNAME /NEGATIVE	OPPOSITE GENDER NAME/POSITIVE	OPPOSITE GENDER NAME/NEGATIVE
<i>Dyslexic</i>	.442 (.582)	.245 (.626)	.072 (.670)	-.270 (1.691)
<i>ADHD</i>	.072 (.625)	.383 (.526)	.006 (.715)	-.084 (.530)
<i>Normally-Developing</i>	.403(.716)	.263(.653)	.003(.607)	.092(.565)

*The Significance of The  $D_{IRAP-(Trial-Type)}$  Scores Relative to Zero*

The results of the t-tests conducted on the  $D_{IRAP-(trial-type)}$  scores for each group are presented in Table 32. Although all three groups of children responded in the predicted direction on the critical OWN NAME-POSITIVE trial-type, only the D scores of the Normally-Developing and Dyslexic children differed significantly from zero. This suggested that the ADHD children, although in agreement with the predictions, did not have particularly strong attitudes in this regard. Nonetheless, the latter children had strong opinions that they were not negative, with their D scores significantly different from zero on this trial-type (this was also the case for the other two groups). Although the preliminary trial-type data had previously indicated that both the Dyslexic and ADHD children responded in the non-predicted direction for OTHER/NEGATIVE, suggesting some reticence towards viewing others as negative (while the Normally-Developing children did not demonstrate this), in both cases, the D scores on this trial-type did not differ significantly from zero. Interestingly, the sub-set of Normally-Developing children who had participated in Experiment 2 had previously shown significant and predicted D scores for OTHER-NEGATIVE, indicating that they had strong negativity towards others, but this effect was reduced to non-significance when the three groups of Normally-Developing children were combined. In summary therefore, one must conclude that none of the groups held particularly strong negativity towards others.

TABLE 32  
The T-tests Results for Trial-type for Each Group.

Group	OWN NAME /POSITIVE	OWN NAME /NEGATIVE	OPPOSITE GENDER NAME/POSITIVE	OPPOSITE GENDER NAME/NEGATIVE
<i>Normally-Developing</i>	p<.0134	p<.0377	p= .7826	p>.1977
<i>Dyslexic</i>	p<.0003	p<.0405	p>.3890	p>.2675
<i>ADHD</i>	p>.5350	p<.0004	p.=9614	p>.3920

*Overall D Scores*

Figure 17 presents the differences in responding between the three groups of children in terms of the overall D scores. The Normally-Developing children produced the strongest D-Score (.222), followed by the Dyslexic group (.122), and finally the ADHD group (.094), thus indicating the characteristic pattern of responding on the IRAP amongst the three groups.

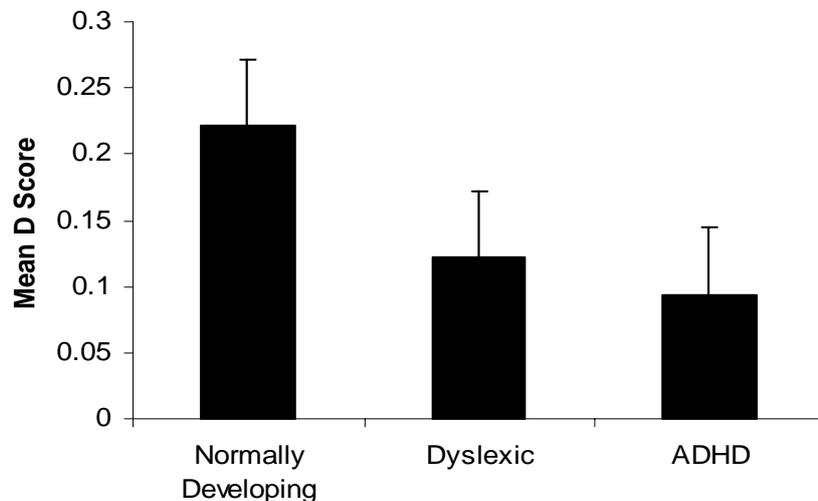


Figure 17: Group mean overall D scores.

*Explicit Measure*

The mean TOT scores for each group on the Piers Harris 2 indicated that all three groups of children fell within the average (45T-55T) range of self-esteem:

Normally-Developing (55.2); ADHD (48) and Dyslexic (48). Because the TOT score is informed by the more specific domain scores, further analysis were conducted on each score and the means and norms for each score by group are presented in Table 33.

Table 33  
Piers-Harris Mean Domain Scores and Norms for All Groups

Domain Scales	Normally-Developing		Dyslexic		ADHD	
	Mean	Range	Mean	Range	Mean	Range
<i>Behavioural Adjustment</i>	55	(A)	53	(A)	38	(L)
<i>Intellectual/School Status</i>	54	(A)	43	(LA)	43	(LA)
<i>Physical Appearance</i>	50	(A)	44	(LA)	41	(LA)
<i>Freedom from Anxiety</i>	52	(A)	48	(A)	41	(LA)
<i>Popularity</i>	55	(A)	49	(A)	39	(L)

The means for the Normally-Developing children fell within the average range (A) on all five domain scales, indicating that these children had high explicit positivity with regards to themselves. Lower than average (L.A) scores began to emerge with the Dyslexic group on the Intellectual/School Status and Physical Appearance domains, suggesting that this group were less happy with aspects of their physical appearance and perceived intellectual abilities. The latter, of course, is consistent with the specific intellectual difficulties faced by these children in educational contexts. It was perhaps not surprising then that even lower scores on the domains were presented by the ADHD group. In fact, these children produced low-average (LA) to Low (L) scores on all five domains, with the lowest scores on Behavioural Adjustment and Popularity. Again, this latter outcome is consistent with the children's poor perceptions of themselves particularly in educational and social contexts.

In summary, the results on the IRAP indicates the differences between the Normally-Developing group and those groups of children with SEN. Specifically, the ADHD group had the smallest D-Score on OWN-NAME POSITIVE relative to the other two groups. A similar profile emerged in relation to the overall D score for the three groups with the ADHD children presenting with lower implicit self-esteem than the Normally Developing and Dyslexic children, while the Dyslexic group had lower implicit self-esteem than their Normally-Developing peers. On the explicit measure, the Dyslexic and ADHD groups had less positive self-regard on the sub domain scales in relation to the areas of difficulties they experience. Taken together, the outcomes observed on both the implicit and explicit measures appear to discriminate between groups of children on the self-esteem continuum.

## **PART 2**

### **Teachers' Comparisons from Experiments 4 and 5**

Although Experiments 4 and 5 each comprised of numerous types of analyses conducted on the teachers' implicit and explicit performances, there are again several critical dimensions along which comparisons can be drawn between the two studies if one wishes to identify potential differences in the teachers' implicit attitudes to normally-developing versus EBD pupils. These comparisons include: the  $D_{\text{IRAP-(trial-type)}}$  scores; the significance of the  $D_{\text{IRAP-(trial-type)}}$  scores relative to zero; overall D scores for the samples PUPIL versus EBD PUPIL. Each of these is described separately below.

*The  $D_{IRAP-(trial-type)}$  Scores*

In order to highlight comparisons between the two target samples PUPIL and EBD PUPIL, the data from the two groups of teachers exposed to both samples (i.e. Primary and Post-Primary teachers) were abstracted from Experiments 4 and 5 and compared (the Control group and the In-Training teachers had been exposed only to EBD PUPIL and thus are excluded here). The two groups of teachers'  $D_{IRAP-(trial-type)}$  scores for each of the four trial-types on each of the two experiments (EBD PUPIL: Experiment 4 and PUPIL: Experiment 5) are presented in Table 34.

TABLE 34  
 $D_{IRAP-(trial-type)}$  scores (+SE) for the Two Groups of Teachers on the Four  
 Trial-types Grouped by Sample

PRIMARY TEACHERS				
	TEACHER /POSITIVE	TEACHER /NEGATIVE	PUPIL /POSITIVE	PUPIL /NEGATIVE
EBD PUPIL (Experiment 4)	.294 (.648)	.269 (.769)	.085 (.653)	.119 (.732)
PUPIL (Experiment 5)	.421 (.644)	.155(.654)	-.215 (.667)	.001 (.743)
POST-PRIMARY TEACHERS				
	TEACHER /POSITIVE	TEACHER /NEGATIVE	PUPIL /POSITIVE	PUPIL /NEGATIVE
EBD PUPIL (Experiment 4)	.357 (.619)	.338 (.772)	.124(.592)	.053 (.678)
PUPIL (Experiment 5)	.147 (.725)	.147 (.641)	-.150 (.633)	.071 (.095)

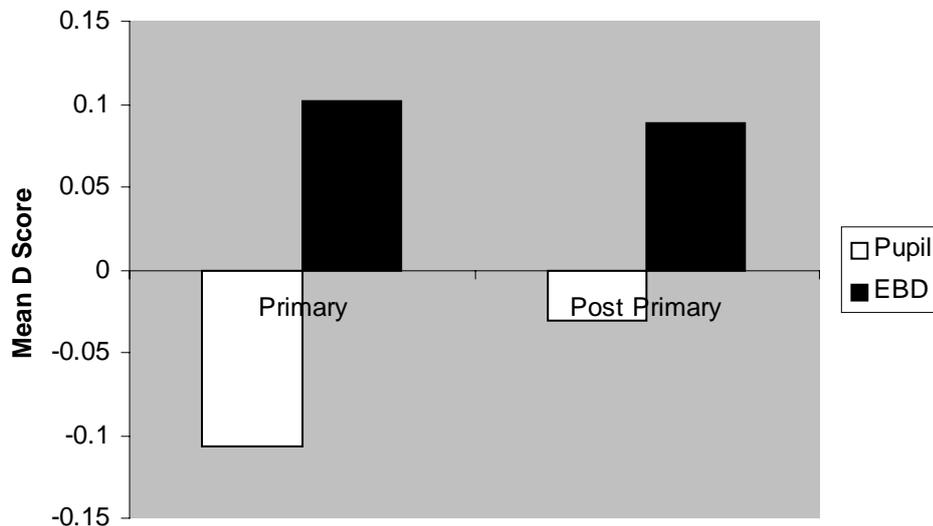
In the context of both EBD PUPIL and PUPIL samples, the Primary teachers showed strong positivity towards teachers. However, their attitudes to TEACHER-POSITIVE were stronger when PUPIL was the sample than when it was EBD. This outcome might imply that Primary teachers perceive themselves to be a more cohesive group when dealing with Normally-Developing pupils than when faced with

the challenges of dealing with EBD pupils. When presented with PUPIL, Primary teachers unexpectedly disagreed that PUPIL was negative in terms of the PUPIL-POSITIVE trial-type. However, when presented with EBD PUPIL, performances on both trial-types (i.e EBD PUPIL POSITIVE: EBD PUPIL NEGATIVE) indicated consistent negativity towards this sample (although the D scores in both cases were small). Hence, across the two experiments, the IRAP appeared to detect subtle differences in the Primary teachers' levels of positivity towards themselves as well as more negative attitudes to EBD PUPIL relative to PUPIL.

Unlike the Primary teachers, the Post-Primary teachers showed stronger positivity towards teachers when EBD PUPIL was the sample relative to the PUPIL sample on both teacher trial-types. This finding implies that this group of teachers perceived themselves and other teachers as more cohesive when dealing with EBD students than Normally-Developing students. Interestingly however, the performances of these teachers resembled those of the Primary teachers on the pupil trial-types. Specifically, when presented with PUPIL, these teachers also disagreed that PUPIL was negative on the PUPIL-POSITIVE trial-type, but showed mixed reactions to PUPIL-NEGATIVE. However, when presented with EBD PUPIL, performances on both trial-types were as predicted, showing negativity towards this sample (although the D scores in both cases were small). Hence, across the two experiments, the IRAP again detected subtle differences in the Post-Primary teachers' levels of positivity towards themselves as well as more negative attitudes to EBD PUPIL relative to PUPIL.

In order to determine whether the differences in the teachers' implicit attitudes to PUPIL versus EBD PUPIL were statistically significant, the  $D_{\text{IRAP-(trial-type)}}$  scores from both studies were entered into two separate one-way ANOVA's – one for

each group of teachers, with the D scores grouped by sample (i.e. PUPIL versus EBD PUPIL). For the Primary teachers, there was a highly significant difference between the two samples [ $F(1, 19) = 4.473, p = .0053, \eta_p^2 = .385$ ], but not for the Post-Primary teachers ( $p = .15$ ). Figure 18. illustrates the differences between the two groups in this regard and highlights the fact that both groups of teachers, particularly the Primary teachers) showed positivity to PUPIL (i.e. the reverse direction indicates a lack of predicted negativity), but negativity to EBD PUPIL (as predicted).




---

Figure 18: Primary and Post-Primary Teachers' D scores for the Samples PUPIL versus EBD PUPIL.

---

*The Significance of The  $D_{IRAP-(Trial-Type)}$  Scores Relative to Zero*

The results of the one-group t-tests conducted on the  $D_{IRAP-(trial-type)}$  scores for each group of teachers are presented in Table 35. While all four groups of teachers responded in the predicted direction on the TEACHER-POSITIVE trial-type, only the D scores of the Primary teachers for both PUPIL and EBD PUPIL, and the D scores

for Post-Primary teachers on EBD PUPIL differed significantly from zero. No significant results were obtained for the Post-Primary teachers presented with PUPIL

TABLE 35  
The T-test Results for Primary and Post-Primary Teachers on the  $D_{IRAP-(Trial-Type)}$  Scores

PRIMARY TEACHERS				
	TEACHER /POSITIVE	TEACHER /NEGATIVE	PUPIL /POSITIVE	PUPIL/ NEGATIVE
EBD PUPIL	p<.0008	p<. 0088	p>.3157	p>.2111
PUPIL	p<.0001	p>.0712	p<.0155	p = .9925
POST-PRIMARY TEACHERS				
	TEACHER /POSITIVE	TEACHER /NEGATIVE	PUPIL /POSITIVE	PUPIL/ NEGATIVE
EBD PUPIL	p<.0001	p<.0012	p>.1102	p>.5467
PUPIL	p>.1218	p>.0802	p<.0721	p>.4608

The results of the t-tests conducted on the  $D_{IRAP-(Trial-Type)}$  scores confirmed the teachers' attitudes to teachers and the two types of pupils as described above. In practically all cases, both groups of teachers had strong and positive attitudes towards teachers that were significant (only the Post-Primary teachers' attitudes to TEACHER-POSITIVE in the context of PUPIL were non-significant), as predicted. In terms of attitudes to pupils, both groups of teachers had shown strong negativity only to EBD PUPIL on PUPIL-POSITIVE, but this was not significant. However, their strong positive and non-predicted attitudes to PUPIL were either significant (Primary) or approaching significance (Post-Primary). The t-tests, therefore, provided further confirmation of the teachers' implicit attitudes to PUPIL as positive and EBD PUPIL as negative.

In summary, both groups of teachers had strong implicit positive attitudes to teachers on both EBD and PUPIL samples, however the Primary teachers attitudes

were stronger when the sample PUPIL was presented. Negativity towards EBD pupils was observed for both groups of teachers but a significant difference between PUPIL and EBD PUPIL was only observed for the Primary teachers.

## Chapter 8

### General Discussion

## Chapter 8

### General Discussion

The purposes of the current thesis were threefold. The first aim was primarily methodological and attempted to determine the utility of the IRAP as a measure of children's implicit attitudes to self. Within this aim, the utility of the procedure with children was examined with samples of individuals who varied along the SEN continuum. The second aim of the thesis was to determine whether the children who did successfully proceed through the IRAP would show a profile of implicit attitude to the self and others that was consistent with experimental and developmental predictions and whether this profile was influenced by their relative position on the SEN continuum. The third aim of the thesis concerned what might be referred to as the other side of attitudes to SEN in terms of teachers' rather than the children's own views. Within this aim, we attempted to compare the implicit attitudes of different groups of teachers towards normally-developing and EBD pupils. The findings that emerged from these aims are reviewed below, followed by theoretical discussions of the issues raised therein.

#### *Summary of Findings from Part 1: Chapters 2, 3, 4 and 7.*

*Chapter 2* was the first study that attempted to employ the IRAP as a measure of children's implicit attitudes. The research that comprised Experiment 1 targeted the children's attitudes to gender, with the child's own gender label (BOY or GIRL, as appropriate) juxtaposed against the label for the opposite gender (BOY or GIRL, as appropriate) in combination with Positive and Negative evaluative terms. The study predicted that the children (males and females) would show an implicit bias towards

their own gender and against the other gender. Indeed, both the young males and females more readily paired their own gender label with Positive than Negative terms and more readily paired the opposite gender label with Negative than Positive terms. There were no differences between the two genders in this regard.

*Chapter 3* attempted to extend the experimental aims of the previous study by targeting the more amorphous concept of self-esteem, although the IRAP samples were again directed at gender. Specifically, the IRAP in Experiment 2 employed each child's own name as a means of tapping into the children's implicit attitudes to self, rather than their gender identities. OWN NAME was juxtaposed with a standard name that characterised the opposite gender (i.e. LOLA as a girls' name and PETER as a boy's name). Again the children were required to relate these to Positive and Negative evaluative terms. Because of the new focus on self-esteem, Experiment 2 also involved a comparison between Normally-Developing children and a small sample who presented with Dyslexia. This latter feature of the study was designed to determine the utility of the IRAP with children who did not present with 'normal' levels of academic attainment and to examine whether this group would show a different pattern of implicit self-esteem relative to the normally-developing children. The results of the study indicated that both groups of children proceeded easily through the IRAP and showed strong and positive implicit attitudes to self. There were no differences between the two groups of children in the latter regard (i.e. the Dyslexic children did *not* show evidence of lower implicit self-esteem relative to those who were Normally-Developing). Interestingly, however, small (but non-significant) differences did emerge between the groups in their implicit attitudes to others. That is, the Normally-Developing children were consistently negative towards others (i.e. they agreed that they were negative and disagreed that they were positive),

whereas the Dyslexic children had mixed views in this regard (i.e. they did not agree that others were either positive or negative). Although the two groups of children showed similarly strong self-esteem profiles on the explicit Piers-Harris 2 (the TOT scores in both cases were within the average range), differences emerged on further analyses of specific sub-domain scales. Specifically, the Normally-Developing children scored within the average range on both Intellectual/School Status and Physical Appearance, whereas the Dyslexic children scored low-average on both.

Experiment 3 reported in *Chapter 4* attempted to develop the research program with children further with the inclusion of a sample of children with EBD, namely those with a formal diagnosis of ADHD. Hence, the study was concerned with the utility of the IRAP with a sample of children with more established difficulties with regard to academic attainment than those presented by the former sample with Dyslexia, and who might subsequently present with different implicit attitudes to self. The results recorded with the Normally-Developing children were consistent with the two previous studies and indicated strong positive self-regard (i.e. they agreed that they were positive and disagreed that they were negative). Although the ADHD children also showed positive self-regard, this outcome overall was more mixed. Specifically, the D score for the OWN NAME-POSITIVE trial-type was non-significant and considerably smaller than that recorded with the Normally-Developing children. Interestingly, however, this group of children had strong views that they were not negative, with significant D scores on OWN NAME-NEGATIVE in the predicted direction (i.e. they readily disagreed that they were negative). Again, similar to the two previous studies, the Normally-Developing children had strong negativity towards others (i.e. they agreed that they were negative and disagreed that they were positive). However, the children with ADHD demonstrated mixed opinions

of others that were very similar to the children in Experiment 2 who presented with Dyslexia (i.e. they did not agree that others were either positive or negative).

Although the two groups of children showed similarly strong self-esteem profiles on the Piers-Harris 2 (the TOT scores in both cases were again within the average range), differences emerged here also on specific sub-domain scales. Specifically, the children with ADHD scored within the low range on both Behavioural Adjustment and Popularity, and on the low-average range on Intellectual/School Status, Physical Appearance and Freedom from Anxiety. In contrast, the Normally-Developing children were average across all domain scales.

*Chapter 7* included comparisons of the findings across Experiments 1, 2 and 3, highlighting in particular the similarities across the three groups of normally-developing children and their differences between those with Dyslexia and those with ADHD in terms of implicit self-esteem and attitudes to others. Many of the issues touched on in that chapter are discussed further in the sections below.

#### *Theoretical Issues from Part 1: Chapters 2, 3 4 and 7.*

Three key theoretical issues emerged from the running and results of the three experiments conducted with children, namely the utility of the IRAP with this population; the relationship between the current findings and those reported elsewhere; putative explanations for the differences among the groups of children on the implicit and explicit measures; and correlations between the implicit and explicit measures. Each of these is discussed separately below.

*The Utility of the IRAP as a Measure of Children's Implicit Attitudes.* A total of twenty-eight normally-developing children participated across Experiments 1, 2 and 3, along with ten children diagnosed with Dyslexia and ten with ADHD. All

forty-eight children proceeded rapidly and easily through all aspects of the procedure and invariably reported that their experiences of the program were challenging and positive. Indeed, there were no recorded differences in the relative abilities of the three groups to proceed with the program.

The basic program to which the children were exposed was identical to that employed in the original research by Barnes-Holmes et al. (2006), with the exception of adjustments to the instructions, including illustrated screen shots. Although it is difficult to determine precisely the impact of this modification to the outcomes, the smooth running of the programs, the children's rapid progression and the consistent outcomes all suggest that the children clearly understand what they were being asked to do and that the modified instructions contributed to this. However, the fact that the length of the response latencies and the number of practice blocks required by the children were very similar to those recorded with adults suggests that the modifications did not alter the performance outcomes.

Indeed it was perhaps surprising that the two groups of children who presented with Dyslexia and ADHD required little or no substantive additional assistance in the completion of the program (i.e. the latter group required some additional encouragement for staying on task and the program incorporated additional aspects of visual feedback). However, although both conditions are associated with poorer academic attainment, they are not necessarily associated with overall impairments in IQ (Hechtman, et al., 2004, Snowling, 2000). In any case, the simple visual format and trial repetitions of the IRAP clearly render the procedure child-friendly and most likely usable even by children who would present with some degree of IQ impairment.

*Relationship with Existing Evidence.* The findings from Experiments 1, 2 and 3 represent the first IRAP evidence with children and thus make a considerable contribution to the existing IRAP literature. This contribution is expanded further by the focus of the studies on gender and self-esteem in particular, neither of which has been examined previously with the procedure.

There are only several published IAT studies that have been conducted with children and these have primarily targeted implicit attitudes to race and gender. Although both studies of race demonstrated predictable pro-White/anti-Black bias that were consistent with data obtained with adults, both involved a substantively modified procedure (the Ch-IAT) primarily designed to be more child-friendly (Baron & Banaji, 2006; Rutland et al., 2005). While many of the reported changes to the original procedure appear to be instructional in nature, there were also changes to the response options that render it difficult to determine precisely what impact these may have had on the outcomes. In any case, it is interesting to note that no such changes were necessary to the IRAP in comparisons of the data obtained with adult and child samples.

The only published IAT study of children's implicit attitudes to gender involved the use of the original procedure with eleven-year-olds (Skowronski & Lawrence, 2001). Unexpectedly however, the children showed no preference for their own gender. As noted previously, these outcomes are inconsistent with a wealth of research evidence that suggests that gender preferences, particularly in terms of one's associations with one's own gender, are well-established by age ten.

In an attempt to account for their unexpected findings, Skowronski and Lawrence (2001) raised the possibility that they had failed to tap into the children's gender identities because they had not used samples that identified the self or other

directly. In other words, it is possible that the children did not interpret the genders targeted in terms of themselves. This seems to be a likely explanation for their findings when one considers the outcomes obtained in the IRAP here in which the children's direct associations with their own gender were targeted. As a result, the findings reported by Skowronski and Lawrence may be more methodological than developmental and highlight, in any case, the importance of carefully selecting appropriate sample stimuli on implicit measures, particularly with children.

*Why the Children Differed on the IRAP.* Although we had not expected the samples of children to differ in their gender associations, some differences were expected with regard to self-esteem, based on the assumption that weaker levels of academic attainment would be associated with lower levels of self-regard. Hence, we identified three groups of children who appeared to differ along this continuum. In their implicit attitudes to self, the Normally-Developing children produced the same profile across the three studies, and these were consistent with those recorded for the children with Dyslexia. That is, all presented with strong positive self-regard on both of the target IRAP trial-types. The children with ADHD, however, presented a more mixed view of themselves in which they only marginally agreed that they were positive, but strongly and significantly disagreed that they were negative. Interestingly, both the Dyslexic and ADHD children differed from the normal-developers in presenting mixed views about others, compared to stronger negativity to others displayed by the latter.

It was perhaps ambitious (or even pessimistic) to expect the children with Dyslexia to demonstrate a different profile of implicit self-esteem from the Normally-Developing children, especially in light of the fact that the two groups scored within the average range of the Piers-Harris 2. This similarity may, of course, result in part

from the fact that Dyslexia is a form of *specific* learning difficult that does not impair IQ output overall and perhaps more importantly is generally not associated with overtly challenging behaviour in educational or social contexts. Indeed, Dyslexia is difficult to detect and thus it is likely that other children will not know that a peer presents with specific educational challenges associated with the condition. Furthermore, it is common practice for children with Dyslexia to be provided with additional school- and community-based supports, including one-to-one assistance from learning resource teachers. Consequently, while formal self-esteem enhancement may not be targeted directly, this assistance may inadvertently serve to support the children's self-esteem. As a result, they may be less likely to have negative regard for themselves and for others (from whom they receive help and support). Of course, it is also the case that strategic supports enable the Dyslexic child to identify him or herself as different, and thus negativity towards others (who are also different) would seem unlikely.

The profile associated with ADHD, however, differs markedly from the characteristics of normal development and Dyslexia. These differences appear to come from three key areas that increase the likelihood that self-esteem will be impaired in this population, in spite of the fact that they do not present with overall impairments in IQ.

First, a range of underlying deficits in executive function have been associated with ADHD, including; working memory, response inhibition, temporal processing, inattention, and poor organisational and study skills (Barkely, 1997; Castellanos & Tannock, 2002; Rapport, Chung, Shore, & Denney et al., 2000; Shelton Barkley, & Crowsswait, 2000). Perhaps relatedly, these individuals appear to have immature appraisal systems in which they incorrectly determine the locus of control of events

around them (Selikowitz, 2004). Although inappropriate appraisal is a natural developmental tendency, it is usually absent by late childhood in normal development, but appears to be retained longer in ADHD, particularly where the child operates regularly in a broad context of confused autonomies or competences. Indeed, there is some empirical evidence to suggest that children with ADHD attribute their failure to internal causes (such as effort and ability) but attribute their success to external unstable causes, such as luck or chance (Selikowitz, 2004). As a result, they may be denied a sense of pride during success and are thereafter susceptible to learned helplessness and lack of motivation to continue (Cooley & Ayres, 1988).

Second, many have argued that individuals with ADHD experience problems in all areas of their lives (HADD, 2005). It is not surprising, therefore, that this group of individuals are reported to be highly sensitive to the perceptions of others, including peers, family and teachers (Guevremont & Dumas, 1994; Roes, 1998; Wheeler & Carlson, 1994). The view that these perceptions are in fact consistent with reality (rather than inaccurate perceptions) is supported by additional evidence that levels of self-efficacy reported by parents of this population are lower than those of normally-developing children (Breen & Barkely, 1988; Mash & Johnson, 1983) and that stress is greater (Anastopolous, Guevremont, Shelton, & du Paul, 1992).

Third, researchers have also argued that the self-esteem of children with learning difficulties are influenced by the type of assistive services they receive (Gurney, 1988; Humphrey, 2002) and the length of time spent in difficulty but without assistance (Scanlon & McGilloway, 2006). Without the necessary academic and emotional support, learning can quickly become aversive and result in loss of motivation and the exacerbation of challenging behaviours -- a learning curve that

seems to characterise the ADHD profile (Wilkinson, 2003). Consequently, results from follow-up studies of children with ADHD have indicated significantly higher rates of: grade retention, placement in special education classrooms, and school drop-out relative to their peers (Weiss & Hechtman, 1993). In light of these three sources of difficulty, it is not surprising that the ADHD children here showed lower than normal implicit self-regard.

*Why the Children Differed on the Piers-Harris 2.* The mean TOT score calculated for each group from Experiments 3 and 4 indicated that all three sets of children fell within the average range of self-esteem. However, differential performances on the domain scales separated the groups. Specifically, as expected, the Normally-Developing children were at least average across the board and all children were individually consistent with this profile. In contrast, the Dyslexic children were invariably low-average on Intellectual/School Status and Physical Appearance. The perceived weakness in the former domain is hardly surprising given that these children by age 10 (i.e. the average age of the current Dyslexic participants) would already be acutely aware of their learning difficulty. For example, Humphrey (2002) reported that Dyslexic children are acutely aware of their deficits in reading and spelling relative to peers. Such perceptions of course likely result in the children feeling some sense of separateness from others who were 'normal' in the context of discrepancies between aspirations and accomplishments (Harter, 1999). Furthermore, if self-esteem is to be viewed as a multi-dimensional concept it is likely that perceived discrepancies on one dimension of the self-concept may impact on other domains (Burton, 2004). In the case of the current research, the separateness from the norm appears to manifest itself in terms of Physical Appearance. Put simply, although these children are not physically different from non-Dyslexic children they perceive

that they are by virtue of the fact that they differ academically. However, the level of difference is not so overt to warrant differences in other domain scales, such as perceptions of popularity.

Although the children here with ADHD also differed from the normal-developers on specific Piers-Harris domain scales, the profile differed also from those who presented with Dyslexia. Indeed, the ADHD group produced low-average to low scores across the board, but were invariably low on Behavioural Adjustment and on Popularity. Of course, the breadth of weaknesses is entirely consistent with the condition. Perceived difficulties on Behavioural Adjustment also indicate that the children in this group see that their own behaviour is problematic for themselves and others, hence popularity concerns would also be expected. Again, these perceptions are entirely consistent with reality, with deficient and troubled peer relationships a characteristic of the condition (Barkley, 1998; Hinshaw & Melnick, 1995; Erhardt & Hinshaw, 1994) and well-established social maladjustment associated with inept behavioural control or regulation.

Although all three groups of children produced overall scores on the explicit Piers-Harris that indicated average self-esteem, the sub-scales were perhaps more informative. Indeed, the latter were surprisingly consistent with common characteristics differentially associated with Dyslexia and ADHD. These outcomes clearly highlight the greater utility of the sub-scales and suggest less benefit from the overall TOT score.

*Correlation of Implicit and Explicit Measures.* One of the primary aims of Experiment 2 was to assess potential correlations between the implicit and explicit measures. Not surprisingly a correlational analysis conducted between the overall D scores and the TOT scores proved non-significant. Furthermore, a Pearson's  $r$

calculated between Behavioural Adjustment for the ADHD group and their overall D IRAP scores was also non-significant ( $r = .179$   $p = .5917$ ). In contrast, a correlation between the D score and the Popularity scores approached significance ( $r = -.648$ .  $p = .0517$ ).

Because the multi-dimensional and complex nature of self-esteem make it difficult to capture the same aspect of self-esteem across two different types of measure, it seems likely that while both were measuring some aspect of self-esteem, they were not measuring the same one. Indeed, it is difficult to know how this might best be done given the breadth of explicit measures and the narrowness of implicit measures, such as the IRAP. At best, the fact that both explicit and implicit outcomes recorded here were consistent with experimental and diagnostic predictions and that some correlations were recorded with such small sample sizes suggests that potential correlations between the two measures in the future should not be ruled out.

#### *Summary of Findings from Part 2: Chapters 5, 6 and 7.*

*Chapter 5* was the first attempt to employ the IRAP as a measure of teachers' implicit attitudes to students. Because a primary focus of the research program concerned implicit attitudes to SEN, the first experiment with teachers examined their attitudes to pupils who were categorised in this way. Hence, the IRAP employed here juxtaposed TEACHER with EBD PUPIL in the context of Positive and Negative terms. The study predicted that teachers would show an implicit positive bias for TEACHER over EBD PUPIL. Hence, they should more readily associate TEACHERS with Positive terms than Negative and more readily associate EBD PUPILS with Negative terms than Positive. Along similar lines to the research in Part 1, Experiment 4 involved comparisons among several groups. Specifically, three

groups of teachers with different levels of teaching experience were compared with each other, namely; Post-Primary Teachers in Training with limited teaching experience; experienced Primary Teachers; and experienced Post-Primary Teachers. The aim of this aspect of the research was to determine whether the nature of teachers' implicit attitudes towards these pupils was influenced by the educational sector on which they operated or by their level of teaching experience. The teachers were also compared with a group of non-teachers (i.e. undergraduates) in order to determine whether these two types of groups shared similar attitudes towards these children. The study predicted that the non-teachers would show less implicit bias against EBD PUPILS.

The results of Experiment 4 were consistent with experimental predictions when all three groups of teachers showed strong implicit bias towards TEACHERS and negative (but weaker) attitudes towards EBD PUPILS. Primary teachers in particular had stronger implicit attitudes in comparison to the Post-Primary teachers. As predicted, these differed from the implicit attitudes of the Control Group, who were more mixed with regard to both samples. Specifically, they indicated that teachers were positive and EBD pupils were negative, but they did not disconfirm that teachers were negative and did confirm that EBD pupils were positive.

The teachers in Experiment 4 were also presented with an explicit measure (i.e. the ORM Scale) that concerned their attitudes to EBD pupils in terms of their willingness to have this sample integrated into mainstream classrooms. In line with their implicit bias on the IRAP, teachers indicated that although they were generally in favour of the principles of inclusion, they were more reticent about the placement of these pupils within their own classrooms. Indeed, all four groups of teachers

expressed considerable unwillingness to adapt their teaching practices to foster mainstream inclusion of this population.

Although the outcomes from Experiment 4 were largely consistent with experimental predictions, there remained the possibility that these simply reflected the teachers' implicit attitudes to all pupils, rather than EBD pupils in particular. Hence, Experiment 5 reported in Chapter 6 attempted to determine whether similar IRAP outcomes would be recorded when TEACHER was juxtaposed with PUPIL. Again, the study predicted that teachers would show an implicit positive bias for TEACHER over PUPIL, although it was anticipated that the bias against PUPIL would not be of the same magnitude as had been recorded with EBD PUPIL. Given some minor differences that emerged between the Primary and Post-Primary teachers in Experiment 4, Experiment 5 again employed two sets of teachers, one from Primary and the other from Post-Primary to determine whether they differed in their implicit attitudes to PUPILS.

The IRAP outcomes were consistent with Experiment 4 when both Primary and Post-Primary teachers showed implicit bias towards TEACHERS. Unexpectedly, however, this effect was stronger for the former relative to the latter. Both groups had mixed attitudes towards pupils in which they confirmed that they were positive but also agreed that they were negative. The results on the ORM Scale were almost identical to those recorded previously.

*Chapter 7* involved systematic comparisons of the findings from Experiments 4 and 5 in order to highlight differences in the teachers' implicit attitudes to PUPILS versus EBD PUPILS. Only Primary and Post-Primary Teachers were exposed to both samples on the IRAP. In the context of samples, Primary teachers showed strong

implicit positivity towards teachers, but they differed in their attitudes to pupils across samples. That is, they showed greater negativity towards EBD PUPIL than PUPIL. Post-Primary teachers were equally strong in their preference for TEACHERS and particularly so when EBD PUPIL was the sample. Again, their negativity towards EBD PUPIL was greater than towards PUPIL. Interestingly, when the two groups of Primary teachers from Experiments 4 and 5 were compared, a significant difference was recorded in their attitudes to PUPIL and EBD PUPIL. This outcome did not emerge for the Post-Primary teachers.

#### *Theoretical Issues from Part 2: Chapters 5, 6 and 7*

Three theoretical issues emerged from the two experiments conducted with teachers, namely: the relationship between the current findings and those reported elsewhere; why the teachers differed on the implicit and explicit measures; and directions for future research. Each of these is discussed separately below.

*Relationship with Existing Evidence.* Experiments 4 and 5 were the first studies to employ the IRAP as a measure of teachers' implicit attitudes. The findings from this work supported the only existing IRAP study that has been conducted with professionals, reported by Barnes-Holmes et al. (2006) who examined the attitudes of professionals working directly with children with autism. Although this original study found strong and predicted IRAP effects in terms of bias towards professionals and against the children, it failed to find significant differences across the groups of professionals who presented with different levels of experience with this population. Nonetheless, there were small group differences on the implicit and explicit measures. The results from Experiments 4 and 5 contribute to this existing evidence with similar bias towards professionals and against the children, particularly when these present

with SEN, rather than being normally-developing. The systematic comparisons of teachers' attitudes to EBD PUPILS versus PUPILS here and the differences that were recorded therein, suggest that the outcomes recorded with autism professionals in the original study resulted from attitudes to that population, rather for example to all children, although such a direct comparison was not originally conducted.

*Why Teachers Differed on the IRAP.* The IRAP outcomes recorded here differentiated the teachers from non-teachers, as well as highlighting differences among teachers working at different levels of education. Put simply, the teachers collectively showed greater negativity towards EBD pupils relative to non-teachers and teachers from the Primary sector were the most negative in this regard.

There are a number of reasons why Primary teachers showed implicit bias against EBD pupils. In the Irish context, Primary level education has witnessed huge change in the last decade, with a rapid progression towards inclusive practices, dictated largely by the Education for Persons with Disabilities Bill (EPSEN 2004: INTO, 2006). However, there is broad consensus that this process has occurred too quickly and without adequate preparation (Task Force on Behaviour, 2006; Travers, 2006). While teachers generally admit that educational provision for children with SEN in Irish Primary schools has improved considerably in that time, they still report a lack of appropriate training and provision for challenging behaviour in particular (Irish National Teachers Organisation, 2007).

A broader issue that has likely exacerbated Primary teachers' negativity towards EBD pupils and perhaps pupils generally concerns the increasing and rapid diversity of the Irish population, with strong demand for Primary school places for children for whom English is not their first language. Of course, this sub-set of children is not without SEN individuals, whose educational and social needs are

difficult to meet without adequate translation and mediation. Under EPSEN, for example, teachers have a statutory obligation to generate and implement Individual Education Plans (IEPs) for all children. However, Primary teachers have little or no training in SEN and such an IEP would be almost impossible for a Primary teacher to construct for a child with SEN who spoke little English (Gash, 2006). Indeed, the INTO recently estimated that an IEP for a child with SEN would require as many as 36 hours work.

The facts that children at Post-Primary level are older and that teachers there expressed serious concerns about their physical safety in the context of pupils at this level with EBD would render it surprising that Primary teachers showed stronger negativity to EBD pupils than Post-Primary teachers. However, a number of issues may have influenced this difference. First, Primary teachers spend all day with the same children, relative to Post-Primary teachers who spend a maximum of one and a half hours with each class of students. Second, by Post-primary level students with EBD have most likely already been identified and thus there is less of an onus on teachers at this level to detect difficulties of this nature. Third, teachers at Post-Primary teach a limited number of subjects and thus are concerned about the impact of a child's difficulties on those subjects only, rather than across the board. Indeed, numerous researchers have highlighted the greater challenge of Primary teaching in this regard (Olson & Cooper, 2003). Fourth, the emphasis at Post-Primary level on state-based examinations (e.g. the Leaving Certificate) has ensured that even pupils with SEN get resources to enable them to leave school with adequate qualifications. Fifth, in recent open debates about inclusion, Post-Primary teachers have been particularly vocal about the difficulties therein and their plight has been highlighted by high-profile initiatives, such as the recent Task Force on Behaviour and Discipline

in Schools (2006). Taken together, these outcomes suggest that the plight of Primary teachers in detecting and managing children with SEN has gone relatively unnoticed, compared to their Post-Primary counterparts and thus greater negativity towards this population by the former over the latter is not entirely surprising.

Post-Primary teachers, however, remained negative towards EBD Pupils, so there remain clear difficulties in educating SEN individuals at this level also. For example, under section 29 of the Education Act (1998), pupils have the right to appeal against school expulsions and a recent audit of these appeals indicated that 23.5% of appeals between 2001 and 2006 concerned SEN students. Indeed, ADHD was the most common type of disability recorded (Lodge & Feeney, in press).

*Teachers' Explicit Attitudes.* Similar to the first studies with children, Part 2 recorded considerable overlap between the implicit and explicit outcomes. The differences between the teachers were relatively minor on the explicit measure, with more agreement across the groups than disagreement. On the ORM Scale, teachers appeared to offer honest answers in which they indicated that they had general agreement with the principles of SEN inclusion in mainstream, but did not favour the direct impact of this within their on classrooms. In light of the issues discussed above this is hardly surprising and indeed all teachers focused primarily on lack of adequate support as a key concern. One respondent's answers effectively highlight teachers' feelings on this issue:

"Nothing would persuade me to be positive about these moves. I have 34 children in class, 10 remedial and one with little English. I do not want anymore"

*Future Directions for Research.* Taken together, the current findings highlight two issues that may further hamper the process of mainstream inclusion for children with SEN and which suggest the need for additional research. (1) The negative

attitudes of teachers likely exacerbate existing feelings of inadequacy and thus impact negatively on direct interactions with the children. (2) Relatedly, the lack of specific training and support for teachers in this regard reduces potential opportunities for positive interactions and increases possibilities for negative interactions. Amongst the two most difficult issues to resolve in this context are as follows. Although the teachers sampled here appeared to be very candid in their responses to the explicit measure, it remains culturally unwise for teachers to openly express negative attitudes to children, even where these clearly exist and are justified within a difficult and tense educational context. Even if the level of generic training and support for dealing with SEN pupils is increased, it will be difficult to determine whether more specific difficulties remain and to know how these might be assessed. For example, perhaps teachers could be presented with repeated IRAPs in an ABA design involving specific training with SEN and open discussions about their negative implicit attitudes. Furthermore, future research will need to direct attention towards the relationship between implicit attitudes and overt behaviour, because changing one may not necessarily change the other. There may also be self-esteem issues that exacerbate teachers' lack of self-efficacy in their teaching that could be targeted by the IRAP as a guide towards intervention.

### *Conclusions*

The research reported in the current thesis represents perhaps an unusual hybrid between applied social research and experimental psychology. While its primary aim was to explore the general utility of an experimental tool (the IRAP), the findings revealed more than was anticipated about attitudes surrounding SEN from both the children themselves and the teachers who teach them. In this context, therefore, the IRAP permitted a fuller picture of the SEN environment than was

previously available in research that relied heavily on explicit methodologies. Furthermore, the IRAP appeared to offer a useful assessment of implicit attitudes to gender, self-esteem and others. Indeed, the ability of the procedure to detect subtle differences between the groups was unexpected, particularly in light of the fact that many of these differences made sound developmental and social sense. These outcomes suggest that the IRAP was tapping into important issues that comprise one's attitudes to oneself and others. Hence, perhaps the most important outcome of the current research concerns the benefits of integrating different types of question and different methodologies in a discipline too long fragmented. The current work suggests that such integration is not only useful for the solving of research questions, but might actually impact directly on the lives of those who need our attention most.

## References

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4<sup>th</sup> edition). Washington D.C. Author.

Anastopoulous, A.D., Guevremont, D.C., Shelton, T.L., & DuPaul, G.(1992). Parenting stress among families of children with attention deficit hyperactivity disorder. *Journal of Abnormal Child Psychology*, 20, 503-20.

Antonak, R., & Larrivee, B. (1995). Psychometric analysis and revision of the opinions relative to mainstreaming scale. *Exceptional Children*, 62, 139-149.

Asendorpf, J.B., Banse, R., & Mucke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behaviour. *Journal of Personality and Social Psychology*, 83, 380-393.

Avarmidis, E., Bayliss, P., & Burden, R. (2000). Student teachers' attitudes towards the inclusion of children with special educational needs in the ordinary school. *Teaching and Teacher Education*, 16, 277-293.

Barkely, R.A., Anastopoulos, A.D., Guevremont, D.C., & Fletcher, K.E. (1991). Adolescents with ADHD: patterns of behavioural adjustment, academic functioning and treatment utilization. *Journal American Academic Child Adolescent Psychiatry* 30, 752-761.

Barnes-Holmes, D., Barnes-Holmes, Y., Hayden, E., Milne, R., Power, P., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*,

Banaji, M. R., Blair, I.V., and Glasser, J. (1997). In R.S. Wyer (Ed), *Advances in social cognition* (Vol 10, pp. 63 –74). Mahwah, NJ: Lawrence Erlbaum.

Banise, R. (1999). Automatic evaluation of self and significant others; affective priming in close relationships. *Journal of Social and Personal Relationships*, 16, 272-81.

Banise, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes toward homosexuality: Reliability, validity and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie*, 48, 145-160.

Barkley, R.A. (1997). Behavioural inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, 121, 65-94.

Barkley, R.A. (1998). *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment*. New York. Guilford Press.

Baron, A.S., & Banaji, M.R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science*, 17, 53 –58.

Bay, M., & Bryan, T.H. (1991). Teachers' reports of their thinking about at-risk learners and others. *Exceptionality*, 2, 127-139.

Baumeister, R.F. (1993). *Self-esteem: The puzzle of low self-regard*. New York: Plenum Press

Beck, A.T. (1976). *Cognitive therapy and the emotional disorders*. Penguin. U.S.A.

Bekle, M. (2004). Knowledge and attitudes about attention-deficit hyperactivity disorder (ADHD): A comparison between practicing teachers and undergraduate education students. *Journal of Attention Disorders*, 7, 151-160.

Biederman, J., Faraone, S.V., & Chen, W. (1993). Social adjustment inventory for children and adolescents: Concurrent validity in ADHD children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 32, 1059-1064.

Biederman, J., Mick, E., Faraone, S.V., Braten, E., & Doyle, A. (2002). Influence of gender on ADHD in children referred to a psychiatric clinic. *American Journal of Psychiatry*, 159, 36-42.

Blair, I.V., Ma, J.E., & Lenton, A.P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology*, 81, 828-841.

Blascovich, J., & Tomaka, J. (1991). Measures of self-esteem. In J. Robinson, P. Shaver, & L. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 161-194). Academic Press. New York.

Bosson, J.K., Brown, R.P., Zeigler-Hill, V., & Swann, W.B. (2003). Self-enhancement tendencies among people with high explicit self-esteem: The moderating role of implicit self-esteem. *Self and Identity*, 2, 169-187.

Bosson, J.K., Swann, W., & Pennebaker, J.W. (2000). Stalking the perfect measure of implicit self-esteem: the blind men and the elephant revisited? *Journal of Personality and Social Psychology*, 79, 631-643.

Breen, M.J., & Barkley, R.A. (1988). Child psychopathology and parenting stress in girls and boys having attention deficit disorder with hyperactivity. *Journal of Paediatric Psychology*, 13, 265-80.

Buhrmester, D. (1993). *Adolescent friendship and the socialization of gender differences in social interaction styles*. Paper presented at the biennial meeting of the Society for Research in Child Development, New Orleans.

- Burton, S. (2004). Self-esteem groups for secondary pupils with dyslexia. *Educational Psychology in Practice, 20*, 56-72.
- Bussing, R., & Zima, B. (2000). Self-esteem in special education children with ADHD: Relationship to disorder characteristics and medication use. *Journal of American Academy of Child and Adolescent Psychiatry, 39*, 1260-1269.
- Butterworth, G. (1990). Self-perception in infancy, In D. Cicchetti & M. Beeghly (Eds), *The Self in Transition*. Chicago. University of Chicago Press.
- Castellanos, F.X., Lee, P.P., Sharp, W., Jefferies, N.O., Greenstein, D, K., Clasen, L.S., et al., (2002). Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Medical Association, 288*, 1740-1748.
- Castellanos, F.X., & Tannock, R. (2002). Neuroscience of attention-deficit hyperactivity disorder. The search for endophenotypes. *Nature Review Neuroscience, 3*, 617-628.
- Coleman, J. (2001). Self-esteem: towards a better understanding. *Young Minds Magazine, 51*, 18-19.
- Cooley, E.J., & Ayres, R.R. (1988). Self-concept and success-failure attributions of non handicapped students and students with learning disabilities. *Journal of Clinical Psychology, 41*, 554-561.
- Cooper, H, (1979). Pygmalion grows up: A model for teacher expectation communication and performance influence. *Review of Educational Research, 49*, 389-410.
- Cooper, H, . & Good, T. (1983). *Pygmalion grows up: Studies in the expectation communication process*. Longman. New York

Coppersmith, S. (1967). *The antecedents of self-esteem* .. W.H. Freeman. San Francisco

Craik, F.I. M., & Lockhar, R.S. (1972) Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour*, *11*, 671-684.

Cullinan, V., Barnes, D., & Smeets, P.M. (1998). A precursor to the relational evaluation procedures: Analyzing stimulus equivalence. *The Psychological Record*, *48*, 121-145.

Cullinan, V., Barnes-Holmes, D., & Smeets, P.M. (2000). A precursor to the relational evaluation procedure: Analyzing stimulus equivalence II. *The Psychological Record*, *50*, 467-492.

Cunningham, W.W., Johnson, M.K., Raye, C.L., Gatenby, J.C., Gore, J.C., & Banaji, M.R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science*. *15*, 806-813.

Dasgupta, N., &Greenwald, A.G. (2001). On the malleability of automatic attitudes. Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*, 800-814.

Department of Education and Science. (1998). *The Education Act*. Government Publications. Dublin.

Department of Education and Science. (2001). *Report of the Task Force on Dyslexia*. Government Publications. Dublin.

Department of Education and Science. (2001). *The report of the task force on autism: Educational provision and support for persons with autistic spectrum disorders*. Government Publications. Dublin.

Department of Education and Science. (2004). *Education for persons with disabilities bill*. Government Publications. Dublin.

Department of Education and Science. (2006). *School Matters: The report of the task force on student behaviour in second level schools*. Government Publications. Dublin.

Department of Education and Science. (In Press). *Section 29 Appeals: An analysis of the first 1,000 cases*. Government Publications. Dublin.

De Raedt, R., Schacht, R., Franck, E., & De Houwer, J. (in press). Self-esteem and depression revisited: Implicit positive self-esteem in depressed patients? *Behaviour Research and Therapy*.

DeHart, T. & Pelham, W.B. (in press). Fluctuations in state implicit self-esteem in response to daylily negative events. *Journal of Experimental Social Psychology*,

De Houwer, J. (2002). The Implicit Association Test as a tool for studying dysfunctional associations in psychopathology: Strengths and limitations. *Journal of Behaviour Therapy and Experimental Psychiatry*, 33, 115 –133

De Houwer, J. & Eelen. (1998). An affective variant of the Simon paradigm. *Psychology Press*, 12, 45-61.

De Houwer, J., Geldof, T., & De Bruycker, E. (2005). The implicit association test as a general measure of similarity. *Canadian Journal of Experimental Psychology*, 59, 228-239.

de Jong, P., Pasman, W., Kindt, M., & van den Hout, M.A. (2001). A reaction time paradigm to assess (implicit) compliant-specific dysfunctional beliefs. *Behaviour Research and Therapy*, 39, 101-113.

de Jong, P. (2002) Implicit self-esteem and social anxiety: Differential self-positivity effects in high and low anxious individuals. *Behaviour Research and Therapy*, 40, 501-508.

Dovidio, J. F., & Fazio, R.H. (1992). New technologies for the direct and indirect assessment of attitudes. In J. Tanur (ED.) *Questions about questions: Meaning, memory, expression and social interaction in surveys* (pp. 204-237). New York; Sage.

Dovidio, J.K. Kawakami, K. Johnson, C. Johnson, B. & Howard A. (1997). On the nature of prejudice: automatic and controlled processes. *Journal of Experimental Social Psychology*, 33, 510-40.

Dunham, Y., Baron, A., & Banaji. M. (2006). From American city to Japanese village; A cross-cultural investigation of implicit race attitudes. *Child Development*.71, 1223-1240.

Eagly, A. H., & Chaiken, S. (1998). Attitude structure and function. In D.T. Gilbert.S.T. Fiske & G Linksay (Eds). *The handbook of social psychology 4<sup>th</sup> edition, . 1*, pp.269-322. New York: McGraw Hall.

Eagly, A. H., Mladinic, A., & Otto, S. (1991). Are women evaluated more favourably than men? An analysis of attitudes, beliefs and emotions. *Psychology of Women Quarterly*, 15, 203-216.

Egloff, B., & Schmukle, S.C. (2002). Predictive validity on an Implicit Association Test for measuring anxiety. *Journal of Personality and Social Psychology*, 83,1141-1455.

Elbaum, B., & Vaughn, S. (2001). School-based interventions to enhance the self-concept of students with learning disabilities: a meta-analysis, *Elementary School Journal*,3, 303-329.

Erhardt, D., & Hinshaw, S.P. (1994). Initial sociometric impressions of attention deficit hyperactivity disorder and comparison boys: Predictions from social

behaviours and from nonbehavioural variables. *Journal of Consulting and Clinical Psychology*, 62, 833-842.

Erikson, E.H. (1968). *Childhood and Society*. New York. W.W. Norton.

Farnham, S.D. (1999). *From implicit self-esteem to in-group favouritism*.

Unpublished doctoral dissertation. University of Washington. Seattle. WA.

Fagot, B.I., & Leinbach, M.D. (1993). Gender development in young children. From discrimination to labelling. *Developmental Review*, 13, 205-224.

Fazio, R.H. Jackson., J.R. Dunton., B.C. & Williams. C.J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, 69, 1013-27.

Fazio, R. H., & Olson, M.A. (2003). Implicit measures in social cognition research; Their meaning and use. *Annual Review of Psychology*, 54, 297-327.

Fiedler, K., & Bluemke, M. (2003). Faking the IAT; Aided and unaided response control on the Implicit Association Test. *Basic and Applied Social Psychology*, 27, 307-316.

Fredrick, L.D., Deitz, S.M., Bryceland, J.A., & Hummel, J.H. (2000). *Behaviour analysis education and effective schooling*. Context Press. Reno Nevada..

Frederickson, N., & Cline.T. (2002). *Special Educational Needs: Inclusion and Diversity*. Open University Press. Philadelphia.

Gash, H. (2006). Beginning primary teachers and children with mild learning difficulties. *Irish Educational Studies*, 25, 275-287.

Gemar, M.C., Segal., Z.V., Segratti, S., & Kennedy, S.J. (2001). Mood-induced changes on the Implicit Association Test in recovered depressed patients. *Journal of Abnormal Psychology*, 110, 282-289 .

Gilbert., D. T. & Hixon. J.T. (1991). The trouble with thinking: activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, 60, 509 –17.

Golombok, S., & Fivush, R. (1994). *Gender development*. Cambridge. UK: Cambridge University Press.

Greene, R. W., Biederman, J., Faraone, S.V., Ouelletter, C.A., Penn, C., & Griffin, S.M. (1996). Toward a new psychometric definition of social disability in children with attention deficit hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 35, 571-579.

Greene, R. W., Biederman, J., Farone, S. V., Sienna, M., & Garcia-Jettonn, J. (1997). Adolescent outcome of boys with attention-deficit/hyperactivity disorder and social disability: Results from a 4-year longitudinal follow-u[ study. *Journal of Consulting and Clinical Psychology*, 65, 758-767.

Greene, R. W., Biederman, J., Faraone, S.V., Wilens, T.E., Mick, E., & Blick, H.K. (1999). Further validation of social impairment as a predictor of substance use disorder: Findings from a sample of siblings of boys with and without ADHD. *Journal of Consulting and Clinical Psychology*, 28, 349-354.

Greenwald, A.G. & Banaji., M.R. (1995). Implicit social cognition: Attitudes, self-esteem and stereotypes. *Psychological Review*, 10, (1) 4 – 27.

Greenwald, A.G. (1990). What cognitive representations underlie social attitudes? *Bulletin of the Psychomaniac Society*, 28, 254-260.

Greenwald, A.G. McGhee. D. E. & Schwartz,. J.L.K. (1998). Measuring individual differences in implicit cognition; the implicit association test. *Journal Personality and Social Psychology*, 74, 1464-80.

Greenwald, A.G., & Farnham, S.D. (2000). Using the implicit association test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, 79, 1022-1038.

Greenwald, A.G., Nosek, B.A. (2001). Health of the implicit association test at age 3. *Zeitschrift fur Experimentelle Psychologie*, 4885-93.

Greenwald, A.G., Nosek, B.A., & Banaji, M.R. (2003). Understanding and using the Implicit Association Test: An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85, 197 –276.

Griffin, S., & Shevlin, M. (2007). *Responding to Special Educational Needs*. Gill & Macmillan. Dublin.

Gurney, P. (1988). *Self-Esteem in Children with Special Educational Needs*, London, Routledge.

HADD Family Support Group (2005). *Attention Deficit Hyperactivity Disorder. ADHD and Education: A Resource for Teachers*. Dublin.

Harter, S. (1990a). Issues in the assessment of the self-concept of children and adolescents. In A. LaGreca (Ed), *Through the eyes of a child*. Boston, Ally & Bacon.

Harter, S. (1999). *The construction of the self*. London. The Guilford Press.

Hasings, R. P., & Oakfor, S. (2003). Student teachers' attitudes towards the inclusion of children with Special Needs. *Educational Psychology*, 23, 87-94.

Hayes, S. C., & Barnes, D. (1997). Analyzing derived stimulus relations requires more than the concept of stimulus class. *Journal of the Experimental Analysis of Behavior*, 68, 225-233.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). Relational frame theory: A précis. In S. C. Hayes, D., Barnes-Holmes, & B. Roche (Eds.), *Relational*

frame theory: A post-Skinnerian account of human language and cognition (pp. 141-154). New York: Plenum.

Hayes, S. C., & Barnes-Holmes, D. (2004). Relational operants: Processes and implications: A response to Palmer's review of relational frame theory. *Journal of the Experimental Analysis of Behavior*, 82, 213-224.

Hechtman, M.D., Abikoff, H., Klein, R.G., Weiss, M.D., Resnitz, C., & Kouri, J., et al., (2004). Academic achievement and emotional status of children with ADHD treated with long-term methylphenidate and multimodal psychosocial treatment. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43, 812-819.

Heinonen, K., Räikkönen, K., & Keltikangas-Järvinen, L. (2005). Self-esteem in early and late adolescence predicts dispositional optimism-pessimism in adulthood: a 21-year longitudinal study. *Personality and Individual Differences*, 39, 511-521.

Helmstetter, E., Peck, C.C., and Giangreco, M.F. (1994). Outcomes of interactions with peers with moderate or severe abilities. *Journal of the Association for Persons with Severe Handicap*, 10, 263-276.

Hetts, J.J., Sakuma, M., & Pelham, B.W. (1999). Two roads to positive regard; implicit and explicit self-evaluation and culture. *Journal of Experimental Social Psychology*, 35, 512-59.

Hinshaw, S.P. (1992). Externalizing behaviour problems and academic underachievement in childhood and adolescence: Casual relationships and underlying mechanisms. *Psychological Bulletin*, 111, 127-155.

Hinshaw, S.P., & Melnick, S.M. (1995). Peer relations of boys with attention-deficit hyperactivity disorder with and without comorbid aggression. *Development and Psychopathology*, 7, 627-647.

Hodgens, J.B., Cole, J., & Boldizar, J. (2000). Peer-based differences among boys with ADHD. *Journal of Clinical Child Psychology, 29*, 443-452

Hofman, W., Gawronski, B., Gschwendner, T., Be, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*, 1369-1385.

Hoza, B., Pelham, E.E., Milich, R., Pillow, D., & McBride, K. (1993). The self-perceptions and attributions of attention deficit hyperactivity disorder and non-referred boys. *Journal of Abnormal Child Psychology, 3*, 271-286.

Hoza, B., Dobbs, J., Owens, J.S., & Pillow, D.R. (2002). Do boys with attention-deficit hyperactivity disorder have positive illusory self-concepts. *Journal of Abnormal Psychology, 2*, 268-278.

Hoza, B., Mrdug, S., Gerdes, A.C., Hinshaw, S.P., Bukowski, W.M., & Gold, J.A (2005). What aspects of peer relationships are impaired in children with ADHD? *Journal of Consulting and Clinical Psychology, 73*, 411-423.

Humphrey, N., & Mullins, P. (2002). Self-Concept and Self-Esteem in Developmental Dyslexia: *Journal of Research in Special Educational Needs, 2*, 1 – 13.

Humphrey, N. (2002) Teacher and pupil ratings of self-esteem in developmental dyslexia. *British Journal of Special Education 29*, 29–36.

Hunt, P., Alwell, M., Faroon, Davos, F., and Goetz, L. (1996). Creating socially supportive environments for fully included students who experience multiple disabilities. *Journal of the Association for Persons with Severe Handicaps, 21*, 53-71.

Irish National Teachers Organisation (2007). *Newcomer Children in the Primary Education System*. INTO.

Irwin, L.K., Tobin, T.J., Sprague, J.R., Sugain, G. & Vincent, C.G. (2004). Validity of office discipline referral measures as indices of school-wide behavioural status and effects of school-wide behavioural interventions. *Journal of Positive Behaviour Interventions*, 6, 131-147.

James, W. (1890). *Principles of Psychology Vol. 1*, Henry Hold. New York.

Jordan, C.H., Spencer, S.J., Zanna, M.P., Hosino-Browne, E., & Correll, J. (2003). Secure and defensive high self-esteem. *Journal of Personality and Social Psychology*, 85, 969-978.

Karpinski, A. (2004). Measuring self-esteem using the Implicit Association Test: The role of the other. *Society for Personality and Social Psychology, Inc*, 30, 22-34.

Kim, D.Y. (2003). Voluntary controllability of the implicit association test (IAT). *Social Psychology Quarterly*, 66, 83-96.

Koole, S.L. Dijksterhuis, A., & van Knippenberg.A. (2001). What's in a name: implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology*, 80, 669-85.

Laird, R.D., Jordan, K.Y., Dodge, K.A., Petit, G.S., & Bates, J.E. (2001). Peer rejection in childhood, involvement with antisocial peers in early adolescence, and the development of externalising behaviour problems. *Development & Psychopathology*, 13, 337-354.

Larrivee, B. (1982). Factors underlying regular classroom teachers' attitudes to mainstreaming. *Exceptional Child*, 33, 173-180.

Larsen, S.R., Steele, M.M. & Sailor, W. (in press). The relationship of school-wide positive behaviour support to academic achievement in an urban middle school. *Psychology Review*.

Latene, J.M., & Darley B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality & Social Psychology*, 4, 377-83.

Lawerence, D. (1996). *Enhancing Self-Esteem in the Classroom*, London, Paul Chapman.

Leader, G., Barnes, D., & Smeets, P.M. (1996). Establishing equivalence relations using a respondent-type training procedure. *The Psychological Record*, 46, 685-706.

Leader,G., & Barnes-Holmes, D. (2000). Matching-to-sample and respondent-type training as methods for producing equivalence relations: Isolating the critical variable. *The Psychological record*, 51, 429-444.

Lewin, L.M., Davis, B., & Hops, H. (1999). Childhood social predictors of adolescent antisocial behaviour: Gender differences in predictive accuracy and efficacy. *Journal of Abnormal Child Psychology*, 27,,272-292 .

Leslie, J.C., Tierney, K.J., Robinson, C.P., Keenan, M., Watt, A., & Barnes, D. (1993). Differences between clinically anxious and non-anxious subjects in a stimulus equivalence training task involving threat words. *The Psychological Record*, 43,153-161.

Lindeman, J., and. Verkasalo, M. (1995). Personality, Situation, and Positive-Negative Asymmetry in Socially Desirable Responding. *European Journal of Personality* 9, 125-134.

Lipkins, R.M., Hayes, S.C., & Hayes, L.J. (1993). Longitudinal study of the development of derived relations in an infant. *Journal of Experimental Child Psychology*, 56, 201-39.

Maccoby, E.E. (1998). *The two sexes: Growing up apart, coming together*. Cambridge, MA: Harvard University Press.

Maccoby, E.E., & Jacklin, C.N. (1987). Gender segregation in childhood. In H.W. Resse (Ed). *Advances in child development and behaviour*, 20, 239-287.

Maison, D., Greenwald, A.G., & Bruin, R. (2001). The implicit association tests as a measure of implicit consumer attitudes. *Polish Psychological Bulletin*, 32, 61-69.

Maison, D., Greenwald, A. G., & Bruin, R.H. (2004). Predictive validity of the implicit association test in studies of brands, consumer attitudes and behaviour. *Journal of Consumer Psychology*, 1, 405-415.

Mannuzza, S., Klein, R.G., Bessler, A., Malloy, P., & LaPadula, M. (1993). Adult outcome of hyperactive boys. Educational achievement, occupational rank, and psychiatric status. *Archives of General Psychiatry*, 42, 948-952.

Mash, E.J., & Johnson, C. (1983). Parental perceptions of child behaviour problems, parenting self-esteem and mothers' reported stress in younger and older hyperactive and normal children. *Journal of Consulting Clinic Psychology*, 51, 86-89.

McIntosh, K. (2005). *Use of DIBELS ORF trajectories to predict office discipline referrals*. Papers presented at DIBELS Summit 2005, Ratin, N.M.

Mellot, D.S., & Greenwald, A.G. (1999). *But I don't feel old! Implicit self-esteem, age identity and ageism in the elderly*. Unpublished manuscript, University of Washington, Seattle, WA.

- Mercugliano, M.D. (1999). What is attention-deficit/hyperactivity disorder? *Pediatric Clinics of North America*, *46*, 831-843.
- Mierke, J., & Klauer, K.C. (2003). Method-specific variance in the implicit association test. *Journal of Experimental Psychology: General*, *134*, 565-584.
- Milgram, S. (1963). Behavioural study of obedience. *Journal of Abnormal Psychology*, *67*, 371-8.
- Morgan, C. D., & Murray, H. A. (1935). A method for examining fantasies: The Thematic Apperception Test. *Archives of Neurology and Psychiatry*, *34*, 289-306.
- Nitnas, K.M., Neila, A., Nikolaidou, E., Papadimitriou, S., Papadopoulou, I., Fasoulas, A., et al., (2006). Inclusion and challenging behaviours: Greek general educator's perspectives. *The Behaviour Analyst Today*, *1*, 84-95.
- Nosek, B.A., & Banaji, M.R. (2001). The go/no-go association task. *Social Cognition*, *6*, 625-644.
- Nosek, B.A., Banaji, M.R., & Greenwald, A.G. (2002a). Harvesting implicit group attitudes and beliefs from a demonstration website. *Group Dynamics*, *6*, 101-115.
- Nosek, B.A., Banaji, M.R., & Greenwald, A.G. (2002b). Math = male, me = female, therefore math  $\neq$  me. *Journal of Personality and Social Psychology*, *83*, 44-59.
- Nosek, B.A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, *134*, 565-584.
- Nuttin, J.M. (1985). Narcissism beyond Gestalt and awareness: the name letter effect. *European Journal of Social Psychology*, *15*, 353-61.
- O'Brien, T. (2000). Increasing inclusion: Did anyone mention learning? *REACH Journal of Special Needs Education in Ireland*, *1*, 2-11.

- Ollendick, T.H., Weist, M.D., Borden, M.G., & Greene, R.W. (1992). Sociometric status and academic, behavioural and psychological adjustment: A five-year longitudinal study. *Journal of Consulting and Clinical Psychology, 60*, 80-87.
- Paulhaus, D.L. (1986). *Self-deception and impression management in test response*. In A. Angleitner & J.S. Wiggins (Ed). *Personality assessment questionnaire*, 142-165. New York. Springer.
- Pelham, B. W., & Swan, W. B., Jr. (1989). From self-conceptions to self-worth: On the sources and structure of global self-esteem. *Journal of Personality and Social Psychology, 57*, 672-680.
- Pennebaker, J.W. (1997). *Opening up; The healing power of expressing emotion*: Guilford Press. New York.
- Rathus, Spencer. A. (2006). *Childhood. Voyages in Development. Second edition*. Thomson. U.S.A.
- Rapport, M.D. Chung, K.M., Shore, G., Denney, C.B., & Issacs, P. (2000). Upgrading the science and technology of assessment and diagnosis laboratory and clinic-based assessment of children with ADHD. *Journal of Clinical Child Psychology, 29*, 555-568.
- Rogers, T.B., Kuiper, N .A. & Kirker, W.S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35*, 677-688.
- Rosenberg, Morris. (1989). *Society and the Adolescent Self-Image*, Revised edition. Middletown, CT: University Press. Wesleyan.
- Rudman., L.A. Greenwald., A.G. & McGhee.D.E. (2001b). Implicit self-concept and evaluative implicit gender stereotypes: self and ingroup share desirable traits. *Personality and Social Psychology Bulletin, 27*, 1164 –78.

Rudman, L.A., & Goodwin, S.A. (2004). Gender differences in automatic in-group bias: Why do women like women more than men like men? *Journal of Personality and Social Psychology*, 4, 494-509.

Rutland, A., Cameron, L., Milne, A., & McGeorge, P. (2005). Social norms and self-presentation: Children's implicit and explicit intergroup attitudes. *Child Development*, 76, 451-466.

Sachill, L., & Schwab-Stone, M. (2000). Epidemiology of ADHD in school-age children. *Child Adolescent Psychiatric Clinic. North America*, 3, 541-55.

Sackeim, H.A., & Gur, R.C. (1978). *Self-deception, self-esteem and depression: The adaptive value of lying to oneself*. In J Masling (Ed). *Empirical Studies of Psychoanalytic Theory*, 101-157. Hillsdale. N.Y Erlbaum.

Santrock, J.W. (2004). *Child Development, tenth edition*. McGraw – Hill. U.S.A.

Schacter, D.L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology; Learning, Memory, and Cognition*, 13, 501-518.

Semrud-Clikeman, M.S., Biederman, J., Sprich, S., Krifcher, B., Norman, D., & Farone, S. (1992). Comorbidity between ADHD and learning disability: A review and report in a clinically referred sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, 31, 439-448.

Scanlon, G. & McGilloway, S. (2006). Managing children with special needs in the Irish education system: A professional perspective. *Journal of the Irish Association of Teachers in Special Education (REACH)*, 19, 81-94.

Schmukle, S.C., & Egloff, B. (2004). Does the implicit association test for assessing anxiety measure trait and state variance? *European Journal of Personality*, 18, 483-494.

- Selikowitz. (2004). *ADHD: The facts*. Oxford University Press. Oxford.
- Semrud-Clikeman M., Biederman, J., Sprich-Buckminster, S., Lehman, B.K., Farone, S.V., & Norman, D. (1992). Comorbidity between ADHD and learning disability: a review and report in a clinically referred sample. *Journal of American Academic Child Adolescent Psychiatry*, 27, 78-87.
- Shelton, R.L., Barkley, R.A. & Crosswaite, C. (2000). Multimethod psychoeducational intervention for preschool children with disruptive behaviour: two-year post-treatment follow-up. *Journal of Abnormal Child Psychology*, 28, 253-266.
- Shimizu, M., & Pelham, B.W. (2004). The unconscious cost of good fortune: Implicit and explicit self-esteem, positive life events, and health. *Health Psychology*.
- Sidman. M. (1971). *Equivalence Relations and Behaviour: A research Story*. Cooperative Inc. Boston
- Simon. J.R. (1990). The effects of an irrelevant directional cue on human information processing. In Proctor, R.W. and Reeve, T.G. (Eds). *Stimulus-Response compatibility: An integrated perspective pp31-86*. Amsterdam. Holland
- Skowronski, J.J., & Lawrence, M.A. (2001). A comparative study of the implicit and explicit gender attitudes of children and college students. *Psychology of Women Quarterly*, 25, 155-165.
- Slomkowski, C., Klein, R.G., Manuzza, S. (1995). Is self-esteem an important outcome in hyperactive children? *Journal of Abnormal Child Psychology*, 23. 303-315.
- Snowling, M.J. (2000). *Dyslexia*. Blackwell. London.
- Spalding, L.R., & Hardin, C.D. (1999). Unconscious unease and self-handicapping: Behavioural consequences of individual differences in implicit and explicit self-esteem. *Psychological Science*, 10, 535-539.

Spence, J. T. (1974). The thematic apperception test and attitudes toward achievement in women: A new look at the motive to avoid success and a new method of measurement. *Journal of Consulting and Clinical Psychology, 42*, 427-437.

Sroufe, L.A. (1990a). An organizational perspective on the self. In D. Cicchetti & M. Beeghly (Eds). *The self in transition: Infancy to childhood*. 281-307. Chicago. University of Chicago Press.

Steele, R.S., & Morawski, J. G .G. (2002). Implicit cognition and the social unconscious. *Theory & Psychology, 12*, 37 -54.

Stewart, J., & Buggy, T. (1994). *Social status and self-esteem: Children with ADHD and their peers*. ERIC Document Reproduction Service No. ED 400 630).

Teachman, B.A., Gregg, A.P., & Woody, S.R. (2001). Implicit associations of fear relevant stimuli among individuals with snake and spider-fears. *Journal of Abnormal Psychology, 110*, 226-235.

Teachman, B.A., Gapinski, K.D., Brownwell, K.D., Rawlins, M., & Jeyaran, J. (2003). Demonstrations of implicit anti-fat bias: The impact of providing causal information and evoking empathy. *Health Psychology 22*, 68-78

Teachman, B.A., & Woody, S.R. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatment. *Journal of Abnormal Psychology, 112*, 100-109.

Teige, S., Schnabel, K., Banse, R., & Asendorpf, J. B. (2004). Assessment of multiple implicit self-concept dimensions using the Extrinsic Affective Simon Task (EAST). Unpublished manuscript.

Tafarodi, R.W., & Swann, W.B. Jr. (1995). Self-liking and self-competence as dimensions of global self-esteem: Initial validation of a measure. *Journal of Personality Assessment, 65*, 322-342.

Tobin, T., & Sugai, G. (1999). Predicting violence at school, chronic discipline problems, and high school outcomes from sixth graders' school records. *Journal of Emotional Disorders*, 7, 40-53.

Travers, J. (2006). Perceptions of learning-support teachers and resource teachers of each other's role in Irish primary schools. *Irish Educational Studies*, 25, 155-169.

Tretuing, J.J., & Hinshaw, S.P. (2001). Depression and self-esteem in boys with attention-deficit/hyperactivity disorder: Associations and comorbid aggression and explanatory attributional mechanisms. *Journal of Abnormal Child Psychology*, 1, 23-39.

Tulving, E., Schacter, D., & Stark, H.A. (1982). Priming events in word fragment completion are independent on recognition memory. *Learning Memory and Cognition* 8, 336-341.

Wade, B., & Moore, M. (1997). *Experiencing Special Education*. Open University Press. Buckingham-Philadelphia.

Warrington, E.K., & Weiskrantz, L. (1968). New method of testing long-term retention with special inference to amnesic patients. *Nature*, 217, 972-974.

Weighfield, A., & Karpathion, M. (1991). Whoa am I and what can I do? Children's self-concepts and motivation in achievement situations. *Educational Psychologist*, 26, 233-261.

Weinstein, R.S. (2002). *Reaching higher: The power of expectations in schooling*. Cambridge, MA: Harvard University Press.

Weiss, G., & Hechtman, L.T. (1993). *Hyperactive children grow up: ADHD in children, adolescents and adults*. Guilford Press. New York.

Wenar. C., & Kerig. P. (2005). *Developmental Psychopathology; From infancy through adolescence. Fifth Edition.* McGraw Hill. New York.

Wheeler, J. C., & Carlson, C.L. (1994). The social functioning of children with ADD with hyperactivity and ADD without hyperactivity: A comparison of their peer relations and social deficits. *Journal of Emotional and Behavioural Disorders*, 2, 2-12.

Wilkinson, W.K. (2003). *Straight Talk about AD/HD. A Guide to Attention Deficit/Hyperactivity Disorder for Irish parents and Professionals.* Collins Press. Ireland.

Wilson., T.D. Lindsey., D. & Schooler. T. Y. (2000). A model of dual attitudes. *Psychological Review*,. 10, 101 – 126

## **Table of Appendices**

<b>Appendix A</b>	The Piers Harris Self-Concept Scale for Children
<b>Appendix B</b>	Instructions for Children's Studies
<b>Appendix C</b>	Domain Scores from the Piers-Harris Self-Concept Scale for Experiment 2
<b>Appendix D</b>	Domain Scores from the Piers-Harris Self-Concept Scale for Experiment 3
<b>Appendix E</b>	The Opinions Relative to Mainstreaming Scale (ORM scale)
<b>Appendix F</b>	Likert scale representing the stimuli employed in the IRAP
<b>Appendix G</b>	Instructions for Teachers' Studies

## Appendix A

The Piers- Harris Concepts Scale for Children

‘The Way I Feel about Myself’ Piers-Harris 2

[copyright-protected]

## Appendix B

### First set of Instructions for Children's Studies

The first set of instructions to which participants were exposed were explicitly designed to ensure that they understood the meanings of all of the experimental words.

For example, the following instructions were presented to a male participant:

This game involves you having to match certain words together. So, if I show you the word BOY (participant is shown the two sample stimuli) because you are a boy, then you might have to match that word with one of these words (participant is shown the target stimuli), such as ACCEPTED. But I need to make sure that you know what all these words mean. So, if you had to match BOY with ACCEPTED, what do you think ACCEPTED means?

Participants were given ample time to respond and were asked to qualify each answer with an example. Consider the following transcript:

Child: People accept me because they like me.

Researcher: Yes, it's a bit like somebody in your class who might annoy you sometimes, because they have a hard time sitting down all the time. But that's ok because they are a nice person and you have fun with them. So you accept them even if they are a little silly sometimes.

The same procedure was then followed for each of the six target stimuli.

Once the correct meanings of the target stimuli had been determined, the participants were presented with specific instructions regarding the matching of the sample and target stimuli. For example, the following instructions were presented to a female participant:

Sometimes the computer will want you to match your name (eg. MARY) with this set of words (participant is shown the three positive target stimuli), and sometimes it will want you to match the word PETER to this set of words (participant is shown the three negative target stimuli). So, you might have to match MAR with ACCEPTED and PETER with FAULTY.

Then the computer will change its mind and it will want you to match your name e.g. MARY with this set of words (participant is shown the three negative target stimuli) and the word PETER with this set of words (participant is shown the three positive target stimuli). So, now you might have to match MARY with FAULTY and PETER with ACCEPTED.

In order to match the words in this game you have to press either the 'd' or 'k' button on the keyboard. Can you show me where they are? They are the only two keys you have to press. You do not have to press anything else. Ok?

Now, there are two other words involved in this game -- OPPOSITE and SIMILAR. So if I tell you that I want you to match your name (MARY) to ACCEPTED, POPULAR and PERFECT and you see this coming up on the screen (child is presented with schematic representation of a consistent trial with OWN NAME) which key will you press? Will you press 'd' or 'k'?

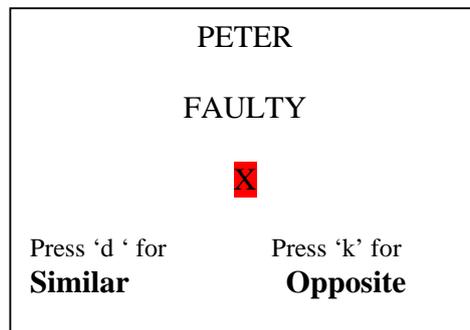
Ok, so what happens if PETER and ACCEPTED come up on the screen but I have told you that I want you to match your name to ACCEPTED what key would you press? Would you press 'd' or 'k'?

Now, like all games there is a little trick involved between the words OPPOSITE and SIMILAR. Sometimes the word OPPOSITE is on this side of the screen (child is shown schematic representation of OPPOSITE on the left-hand side of the screen) and sometimes it is over on the other side of the screen (child is shown schematic representation of OPPOSITE on the right hand side of the screen). It's also the same for the word SIMILAR. Sometimes it is on this side of the screen (pointing to left) and then it changes to the other side of the screen (pointing to right). So the trick is that you have to keep your eye on which side of the screen OPPOSITE and SIMILAR are on, because remember, you have to press 'd' or 'k' to match the words. So your job in this game is to keep your eye on the word at the top (i.e. OWN NAME or PETER), the word in the middle (e.g. ACCEPTED) and the places in which OPPOSITE and SIMILAR appear on the screen. Do you understand?

### *Feedback*

Although, the children had been instructed in all aspects of the IRAP procedure at this point, they had not completed any actual trials and thus had not yet been exposed to the feedback mechanism contained within the program. Although the IRAP is referred to as a test, all incorrect trials (but not correct trials) are consequated with written automatic feedback that indicates that a wrong response has been emitted. This feedback involves the presentation of a red X that appeared automatically in the middle of the screen (see Figure 4). The feedback remained on the screen until the correct response is emitted.

Figure 4



---

Figure 4: Visual aid containing a schematic representation of an incorrect response on an IRAP trial.

---

In order to instruct the children effectively on the feedback, a visual representation of feedback was presented along with the following instructions:

Ok. Like all games, you will sometimes match the words wrong and that's ok. This game is not about getting it right all of the time. Although it's important to get as many words right as you can, the trick is that you have to match the words as fast as you can while trying to get them right. Ok?

If you get one wrong, the computer will show you this (child is presented with the schematic representation of feedback) and the only way you can go forward in the game is to press the correct key. So the computer will actually tell you when you have made a mistake, ok?

*Practice Trials.* Although the format of all eight blocks was identical, the children were given specific instructions initially concerning whether a block was a practice block or whether it was a test block and the instructions explicitly depended upon whether the first block of practice trials was consistent or inconsistent. The instructions provided during the practice blocks were specifically targeted towards teaching the child: (1) how to press the appropriate keys and (2) how to match the sample and target stimuli appropriately. The instructions provided to a male child exposed to consistent trials first, for example, were as follows:

In order for you to learn how to play the game, we are going to do some practice. For this first practice I want you to match PETER to ACCEPTED, PERFECT and POPULAR, and your name (MARY) to FAULTY, BROKEN and USELESS. I will help you through the first few matching tasks. Ok?

Can you sit in close to the computer and make sure you are comfortable? Good, now show me where you are going to put your fingers. Remember for this game you only need to press the 'd' and 'k' keys on the keypad. When you are ready, just press the space bar (researcher indicated location of spacebar) and the game will start.

The first exposure within the first practice block was then presented. During the first few practice trials, the researcher prompted the child by indicating the correct response option and the correct key associated with it. Prompting for correct responding continued until each child was familiar with correct responding (this rarely exceeded five trials). If a child emitted an incorrect response during the initial practice trials they were instructed as follows:

That's ok. Remember that I told you it's ok to get some wrong. All you have to do now is to press the correct key in order to go forward with the game. So, which key will you press?

Throughout the first practice block, the researcher also intermittently provided verbal feedback on correct and incorrect responding.

At the end of each block of trials, the IRAP presented participants with automated feedback on the percentage of trials correct and the median response time (in ms.) achieved during that block. Because of the participants' ages, the researcher provided further clarification on what the automated scores meant. For example, if a child's median reaction time on the first block exceeded 3000ms. s/he was instructed as follows:

Ok .... (Child's name). That was very good. You seem to be getting the hang of this game. There is one other thing I need to tell you. In order to get to the next level of the game you need to try and keep your score under 3000ms. Ok? And I still want you to try to get as many right as you can. Ok?

Hence, irrespective of performance, all children then proceeded immediately with the second practice block.

Because the IRAP systematically presented blocks that alternated between consistent and inconsistent responding, the children required explicit instructions after the first practice block regarding the change in trials in the subsequent block. For instance, in the examples provided above the instructions concerned consistent responding and thus the children then required further instructions regarding the inconsistent trials presented in the subsequent practice block. Hence, the children were explicitly instructed that the second practice block of trials would involve changing their responses as follows:

Ok, now do you remember that I told you that the computer sometimes changes its mind? This time it wants you to match \peter to FAULTY, BROKEN and USELESS and your name (MARY) to ACCEPTED, PERFECT AND POPULAR. This is still only a practice just like the last time. But I want you to try and go as fast as you can, while still trying to get as many right as you can, Ok?

When you are comfortable and ready to go press the space bar, the first set of words will come up on the screen.

The instructional and presentation formats that accompanied the second block of trials were identical to the first, except that the children were instructed to emit alternative response patterns.

### *Test Trials*

After the presentation of the automated feedback following the second practice block of trials, the children were invited to take a short break prior to the first block of test trials. The first test block then commenced with the following verbal instructions:

Ok, it's obvious that you have caught onto this game really quickly. So I think you are ready to go on to the next stage. We are going to do this six more times, but as you can see you get through the stages really quickly, and we will count each task together when you have finished. Ok?

Now, this time I am not going to tell you which key to press. Ok. So, that means that the practice is now over and the computer is going to test you to see how much you remember. Ok? Before each matching task I will just check with you to see if you remember which words you are matching to PETER and which words you are matching your own name to. Is that ok?

Now, for the next ones can you tell me which words you have to match PETER to? (Researcher uses the word lists to prompt the child and waits for a response). Now can you tell me which words you have to match your own name to? (Again, researcher waits for the child's response).

So, are you comfortable and ready to go? Make sure you have your fingers on the right keys. Remember -- try and go as fast as you can but still trying to get them right. Ok, off you go and press the space bar to begin.

Although the children were provided with these specific instructions at the beginning of each block regarding the type of responding appropriate for that block (i.e. inconsistent or consistent – but these actual words were not employed), they were not prompted throughout the block as had been the case with the practice trials. However, the Researcher did remain seated beside each child throughout all trials and continued to provide positive verbal encouragement regarding appropriate on-task behaviour (e.g. “you are doing well and working really fast”), but not for accurate or inaccurate responding.

At the end of the sixth and final test block, the Researcher indicated to the children that the game was over and thanked them for their participation. Each child was then given a packet of gel pens as a reward for taking part in the study.

Appendix C

Table of Domain Scores from the Piers-Harris 2 for Experiment 2: Chapter 3

Dyslexic Group										
Response Bias	48	46	43	63	56	48	46	38	29	39
Self-Esteem( TOT)	55	47	53	33	44	44	46	67	37	48
Behavioural Adjustment	62	54	63	41	62	49	54	62	49	33
Intellectual/School Status	51	43	48	32	40	38	48	60	32	42
Physical Appearance	48	45	48	35	45	42	40	59	28	45
Freedom from Anxiety	51	51	51	35	46	48	43	65	51	35
Popularity	50	44	60	39	39	47	50	69	39	41
Happiness/Satisfaction	47	40	43	37	59	43	33	59	47	41

Normally Developing Group 1										
Response Bias	35	46	29	62	40	48	43	61	46	48
Self-Esteem( TOT)	41	60	58	53	55	60	58	50	61	53
Behavioural Adjustment	49	53	53	54	54	62	62	54	62	49
Intellectual/School Status	41	59	48	54	60	59	62	55	59	46
Physical Appearance	39	52	52	48	48	48	48	46	65	53
Freedom from Anxiety	41	58	65	46	58	48	54	43	54	54
Popularity	47	50	68	54	54	54	60	44	60	60
Happiness/Satisfaction	35	47	47	47	51	47	40	43	51	59

Appendix D

Table of Domain Scores from the Piers-Harris 2 for Experiment 3: Chapter 4

ADHD										
Response Bias	43	51	50	38	44	50	44	61	61	44
Self-Esteem( TOT)	45	59	70	59	43	34	44	46	37	45
Behavioural Adjustment	31	29	28	29	29	46	49	46	50	41
Intellectual/School Status	32	45	47	42	40	38	38	46	44	54
Physical Appearance	45	48	29	41	45	32	42	38	38	48
Freedom from Anxiety	31	28	29	28	65	43	59	54	37	39
Popularity	31	29	34	29	54	30	39	68	39	41
Happiness/Satisfaction	37	40	40	37	40	35	44	37	40	51

Normally Developing Group 2										
Response Bias	52	40	49	66	61	43	48	63	38	49
Self-Esteem( TOT)	55	61	51	61	50	58	60	38	66	56
Behavioural Adjustment	47	54	54	62	54	62	62	37	61	46
Intellectual/School Status	51	59	51	59	55	62	59	38	59	55
Physical Appearance	48	52	58	65	46	48	48	48	52	52
Freedom from Anxiety	55	63	46	54	43	54	48	37	58	54
Popularity	44	54	54	60	44	60	54	44	68	60
Happiness/Satisfaction	40	59	47	51	43	40	47	37	59	43

Appendix E

The Opinions Relative to Mainstreaming Scale Revised (ORM Scale)

**A Survey of Teachers' Attitudes Regarding Children with Emotional and Behavioural Difficulties (EBD) in Mainstream Education**

The Education for Persons with Special Educational Needs Bill 2003 (i.e. the EPSEN Act) seeks:

“to provide that people with special educational needs shall have the same right to avail of and benefit from appropriate education as do their peers”

The impact of the EPSEN Act for mainstream teachers is that it puts extensive responsibilities upon them for the effective management of children with special educational needs within their classrooms.

While the EPSEN Act collectively refers to all children with special educational needs, it subsumes specific categories of children such as those diagnosed with Attention Deficit Hyperactivity Disorder (ADHD), Conduct Disorder (CD), Oppositional Defiant Disorder (ODD), or those with Emotional and Behavioural Difficulties (EBD). The EPSEN Act, therefore, will require that children with EBD (whose diagnosis may traverse more than one of the previous conditions and who are often specifically noted for their difficult conduct) be placed with normally-developing and other children within mainstream classrooms. The current questionnaire is concerned only with the attitudes of mainstream teachers towards children with EBD.

The primary purpose of this questionnaire is to obtain information that will enhance our understanding of teacher attitudes towards children with emotional and behavioural difficulties (EBD) in mainstream education.

It is very important that you are completely aware that all of the information obtained from this measure is completely confidential and anonymous (i.e. at no time will you be asked to provide your name).

*We are very grateful indeed for your co-operation.*

**Official Use only:** ID No. \_\_\_\_\_ Date: \_\_\_\_\_ Gender: \_\_\_\_\_

*Please note that all the information you provide will be kept strictly confidential.*

**(1) Current Teaching Position (Please tick one)**

(a) Principal  (b) Vice-Principal/ In classroom  (c) Teacher

**(2) Teaching Experience: (Please specify number of years employed as a teacher)**

(a) Primary \_\_\_\_\_ (b) Post Primary \_\_\_\_\_

**(3) Teacher Training:**

**(Please tick appropriate training college and indicate year graduated)**

(a) Arts Degree:  Institution \_\_\_\_\_  
Graduated \_\_\_\_\_

(b) Mary Immaculate   
Graduated: \_\_\_\_\_

(c) St. Patrick's Drumcondra   
Graduated: \_\_\_\_\_

(d) Marino Institute   
Graduated \_\_\_\_\_

(e) Conversion Course from Arts to Primary teaching   
Graduated \_\_\_\_\_

(f) Overseas Training College  Institution \_\_\_\_\_  
Graduated \_\_\_\_\_

**(4) Additional Training in Special Education: (Please fill out appropriate section)**

(a) In-service provided by the Department of Education & Science

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

(b) Additional SEN professional development taken outside of departmental requirements

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Please circle the number under the column that best describes your agreement or disagreement with the following statements. There are no correct answers; the best answers are those that honestly reflect your feelings.(The term Emotional Behavioral Difficulties (EBD) in the following scale refers to children with significant challenging behaviors

**1 = Strongly Disagree    2 = Disagree    3 = Undecided    4 = Agree    5 = Strongly Agree**

		<u>SD</u>	<u>D</u>	U	A	SA
1.	The needs of students with EBD are best served through special, separate classes.	1	2	3	4	5
2.	The challenge of being in an ordinary classroom will promote the academic growth of the child with EBD	1	2	3	4	5
3.	Inclusion offers mixed group interaction which will foster understanding and acceptance of differences.	1	2	3	4	5
4.	Isolation in a special class has a negative effect on the social and emotional development of a student with EBD.	1	2	3	4	5
5.	The child with EBD will probably develop academic skills more rapidly in a special classroom than in a regular classroom	1	2	3	4	5
6.	The contact ordinary-class students have with included students with EBD may be harmful.	1	2	3	4	5
7.	Including the child with EBD will promote his/her social independence.	1	2	3	4	5
8.	The inclusion of students with special needs can be beneficial for regular students.	1	2	3	4	5
9.	Inclusion is likely to have a negative effect on the emotional development of the child with EBD	1	2	3	4	5
10.	The child with EBD will be socially isolated by other students.	1	2	3	4	5
11.	Students with EBD should be given every opportunity to function in the general-classroom setting, where possible.	1	2	3	4	5
12.	The presence of students with EBD will promote acceptance of differences on the part of other students.	1	2	3	4	5

### **Emotional Reaction Scale**

If a new student who was described as having severe behavioural problems (a child with ADHD, conduct disorder, oppositional defiant disorder, etc), was about to join your class tomorrow, you would feel..... (please circle the number which best describes your feelings)

<b>Uncomfortable</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Comfortable</b>
<b>Negative</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Positive</b>
<b>Unconfident</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Confident</b>
<b>Pessimistic</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Optimistic</b>
<b>Worried</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Self-assured</b>
<b>Disinterested</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Interested</b>
<b>Unhappy</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Happy</b>

If a new student who was described as having emotional and behavioural difficulties (a child with severe disruptive behaviour, a child with ADHD, a child who has been excluded from other schools because of his/her deviant behaviour etc), was about to join your class tomorrow, you would feel...

<b>Uncomfortable</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Comfortable</b>
<b>Negative</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Positive</b>
<b>Unconfident</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Confident</b>
<b>Pessimistic</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Optimistic</b>
<b>Worried</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Self-assured</b>
<b>Disinterested</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Interested</b>
<b>Unhappy</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>Happy</b>

What extra things would you need to make your responses to the above three questions more positive?

Please circle the number under the column that best describes your agreement or disagreement with the following statements. There are no correct answers; the best answers are those that honestly reflect your feelings.

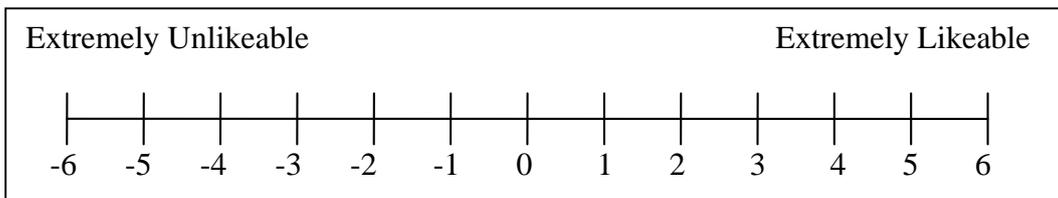
**SD = Strongly Disagree      D = Disagree      U = Undecided      A = Agree      SA = Strongly Agree**

		<b>SD</b>	<b>D</b>	<b>U</b>	<b>A</b>	<b>SA</b>
1.	I would be supportive towards the idea of including children with emotional and behavioural difficulties in my classroom.	1	2	3	4	5
2.	I would be willing to engage in in-service training on teaching children with severe emotional and behavioural difficulties.	1	2	3	4	5
3.	I would be willing to engage in developing the appropriate skills to teach children with emotional and behavioural difficulties in their classroom.	1	2	3	4	5
4.	I would be willing to engage in developing skills for managing the behaviour of children with emotional and behavioural difficulties.	1	2	3	4	5
5.	I would be willing to accept responsibility for teaching children with severe emotional and behavioural difficulties within a whole-school policy.	1	2	3	4	5
6.	I would be willing to continuously assess myself to inform my teaching practice.	1	2	3	4	5
7.	I would be willing to change my teaching processes to accommodate children with emotional and behavioural difficulties in my classroom.	1	2	3	4	5
8.	I would be willing to co-operate with the parents of the children with emotional and behavioural difficulties for the benefit of their children.	1	2	3	4	5

Thank you for completing this questionnaire

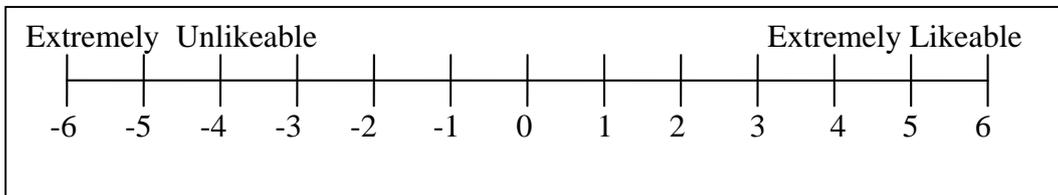
### Likert 1

## Accommodating



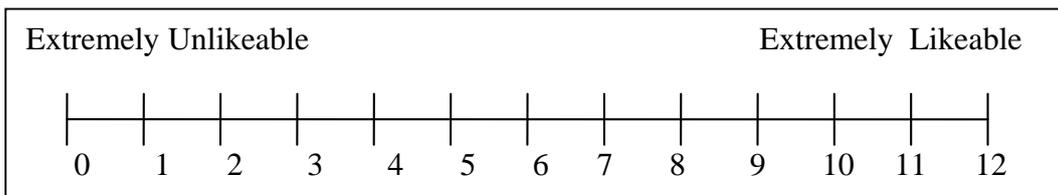
Please indicate the extent to which you would rate the word above as likeable or unlikeable by circling the appropriate number.

### Positive



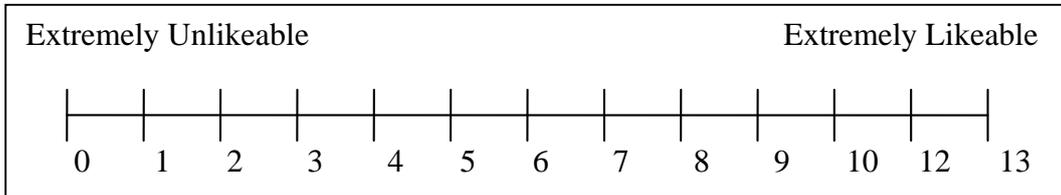
Please indicate the extent to which you would rate the word above as likeable or unlikeable by circling the appropriate number

### Non-Compliant



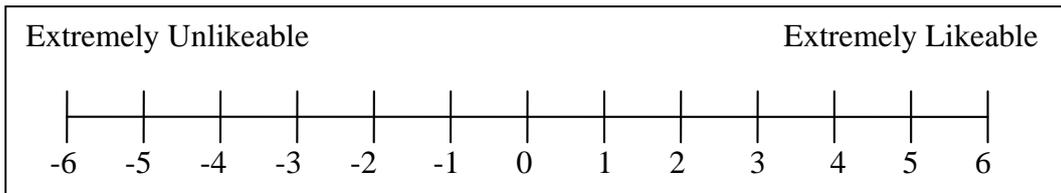
Please indicate the extent to which you would rate the word above as likeable or unlikeable by circling the appropriate number.

## Negative



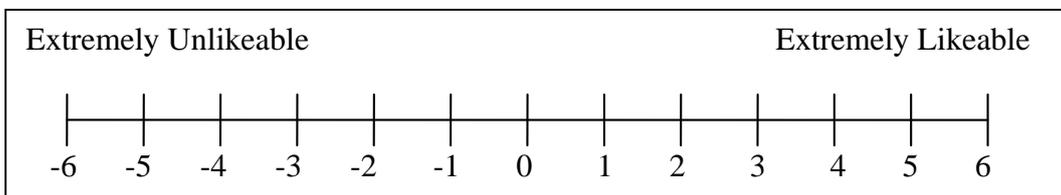
Please indicate the extent to which you would rate the word above as likeable or unlikeable by circling the appropriate number

## Suitable



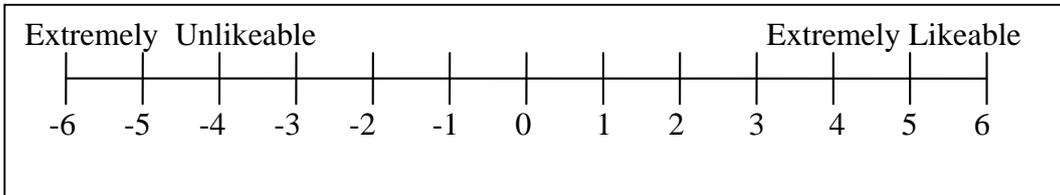
Please indicate the extent to which you would rate the word above as likeable or unlikeable by circling the appropriate number

## Calm



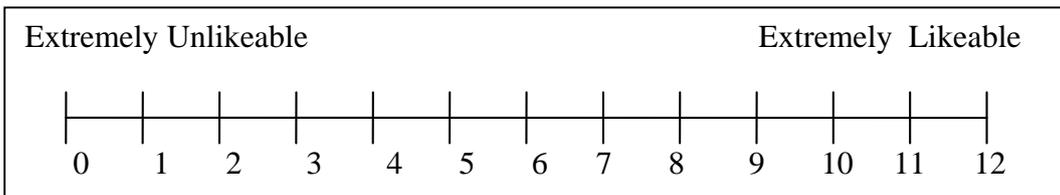
Please indicate the extent to which you would rate the word above as likeable or unlikeable by circling the appropriate number.

# Angry



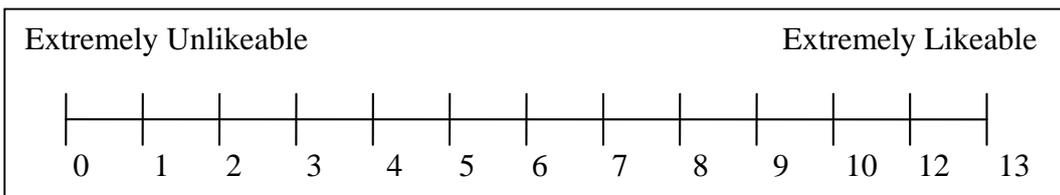
Please indicate the extent to which you would rate the word above as likeable or unlikeable by circling the appropriate number

# Pleasant



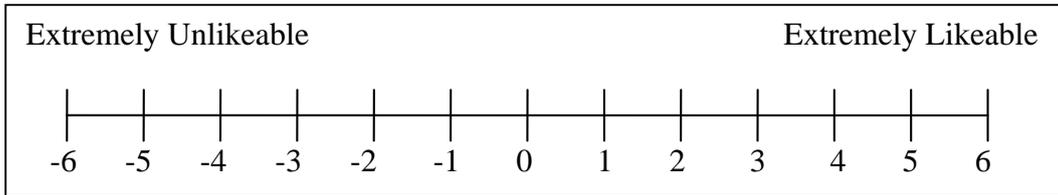
Please indicate the extent to which you would rate the word above as likeable or unlikeable by circling the appropriate number

# Difficult



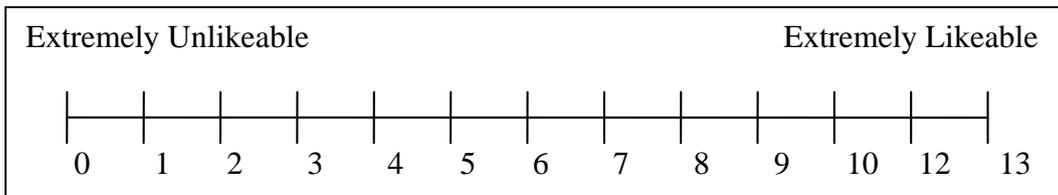
Please indicate the extent to which you would rate the word above as likeable or unlikeable by circling the appropriate number

## Unpleasant



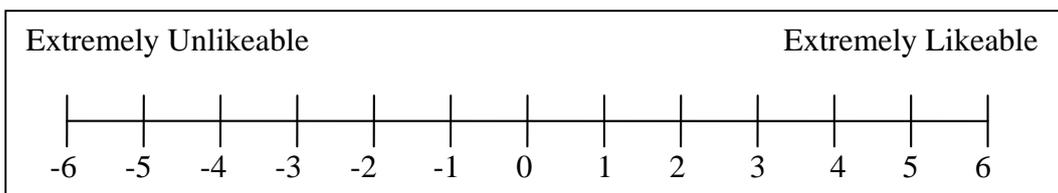
Please indicate the extent to which you would rate the word above as likeable or unlikeable by circling the appropriate number

## Cooperative



Please indicate the extent to which you would rate the word above as likeable or unlikeable by circling the appropriate number

## Inappropriate



Please indicate the extent to which you would rate the word above as likeable or unlikeable by circling the appropriate number

## Appendix G

### First set of Instructions for Teachers Study- Experiment 4: Chapter 5

The first set of instructions to which participants were exposed were explicitly designed to ensure that they understood the meanings of all of the experimental words.

For example, the following instructions were presented to a participant:

This is a computer-based task that involves you having to match certain words together to two target words. The two target words comprise of TEACHER and EBD PUPIL The target word EBD Pupil represents pupils with emotional behavioural disturbances in mainstream classes. The target word TEACHER represents the teacher in the mainstream class Ok?

Participants were then presented with the two word lists that represent each of the twelve target stimuli and asked if they understood what each word represented. Once the correct meanings of the target stimuli had been determined, the participants were presented with specific instructions regarding the matching of the sample and target stimuli. For example, the following instructions were presented to a participant

Sometimes the computer will want you to match your name TEACHER with this set of words (participant is shown the list of positive target stimuli), and sometimes it will want you to match the word EBD PUPIL to this set of words (participant is shown the list of negative target stimuli). So, you might have to match TEACHER with ACCOMODATING and EBD PUPIL with NON-COMPLIANT.

Then the computer will change and it will want you to match TEACHER with this set of words (participant is shown the negative target stimuli) and the word EBD PUPIL with this set of words (participant is shown the positive target stimuli). So, now you might have to match TEACHER with NON-COMPLIANT and EBD PUPIL with ACCEPTED.

In order to make a response to the stimuli on the screen you have to press either the 'd' or 'k' button on the keyboard. Do you know where they are? They are the only two keys you have to press. You do not have to press anything else. Ok?

Now, there are two other response options that enable you to respond -- OPPOSITE and SIMILAR. So for example, if I tell you that I want you to match TEACHER to

ACCEPTED and you see this coming up on the screen (participant is presented with schematic representation of a consistent trial with TEACHER) which key will you press? Will you press 'd' or 'k'?

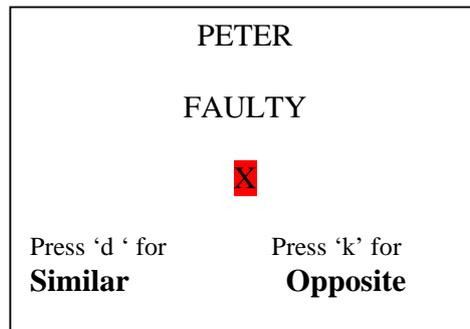
Ok, so what happens if EBD PUPIL and ACCEPTED come up on the screen but I have told you that I want you to match TEACHER to ACCEPTED what key would you press? Would you press 'd' or 'k'?

Finally, sometimes the word OPPOSITE is on this side of the screen (participant is shown schematic representation of OPPOSITE on the left-hand side of the screen) and sometimes it is over on the other side of the screen (participant is shown schematic representation of OPPOSITE on the right hand side of the screen). It's also the same for the word SIMILAR. Sometimes it is on this side of the screen (pointing to left) and then it changes to the other side of the screen (pointing to right). So the trick is that you have to keep your eye on which side of the screen OPPOSITE and SIMILAR are on, because remember, you have to press 'd' or 'k' to match the words. So you need to keep your eye on the word at the top (i.e. TEACHER or EBD PUPIL), the word in the middle (e.g. ACCEPTED) and the places in which OPPOSITE and SIMILAR appear on the screen. Do you understand?

### *Feedback*

Although, the participants had been instructed in all aspects of the IRAP procedure at this point, they had not completed any actual trials and thus had not yet been exposed to the feedback mechanism contained within the program. Although the IRAP is referred to as a test, all incorrect trials (but not correct trials) are consequated with written automatic feedback that indicates that a wrong response has been emitted. This feedback involves the presentation of a red X that appeared automatically in the middle of the screen (see Figure 4). The feedback remained on the screen until the correct response is emitted.

Figure 4



---

Figure 4: Visual aid containing a schematic representation of an incorrect response on an IRAP trial.

---

In order to instruct the participants effectively on the feedback, a visual representation of feedback was presented along with the following instructions:

Ok. Like all word-matching tasks you will probably get some wrong. This task is not about getting it right all of the time. Although it's important to get as many words right as you can, the trick is that you have to match the words as fast as you can while trying to get them right. Ok?

If you get one wrong, the computer will show you this (participant is presented with the schematic representation of feedback) and the only way you can continue the task is to press the correct key. So the computer will actually tell you when you have made a mistake, ok?

*Practice Trials.* Although the format of all eight blocks was identical, the participants were given specific instructions initially concerning whether a block was a practice block or whether it was a test block and the instructions explicitly depended upon whether the first block of practice trials was consistent or inconsistent. The instructions provided during the practice blocks were specifically targeted towards

familiarising the participants with the procedure. For example the following instructions were provided to participants exposed to consistent trials first.

Ok this first part of the procedure is a practice. For this first practice I want you to match TEACHER to ACCEPTED, COOPERATIVE ETC and EBD PUPIL to NON COMPLIANT, ANGRY ETC .

If you sit in close to the computer and make sure you are comfortable and familiarise yourself with where you will place your fingers. You need to keep your fingers on these keys at all times Ok? And remember these are the only two keys you have to press. . When you are ready, just press the space bar (researcher indicated location of spacebar) and the task will begin.

The first exposure within the first practice block was then presented. During the first few practice trials, the researcher sat beside the participant and prompted them if required by indicating the correct response option and the correct key associated with it. Prompting for correct responding continued until each participant was familiar with correct responding (this rarely exceeded three trials). If a participant emitted an incorrect response during the initial practice trials they were instructed as follows:

That's ok. Remember that I told you it's ok to get some wrong. All you have to do now is to press the correct key in order to go forward with the game. So, which key will you press?

Throughout the first practice block, the researcher also intermittently provided verbal feedback on correct and incorrect responding.

At the end of each block of trials, the IRAP presented participants with automated feedback on the percentage of trials correct and the median response time (in ms.) achieved during that block. The researcher provided further clarification on what the automated scores meant. For example, if a participants median reaction time on the first block exceeded 3000ms. s/he was instructed as follows:

Ok ..... (Participants name). That was very good. You seem to be getting the hang of the task. There is one other thing I need to tell you. I need you to try and keep your score under 3000ms. Ok? And I still want you to try to get as many right as you can. Ok?

Hence, irrespective of performance, all participants then proceeded immediately with the second practice block.

Because the IRAP systematically presented blocks that alternated between consistent and inconsistent responding, the participants were reminded about the computer switching the task. For instance, in the examples provided above the instructions concerned consistent responding and thus the participants then required further instructions regarding the inconsistent trials presented in the subsequent practice block. Hence, the participants were explicitly instructed that the second practice block of trials would involve changing their responses as follows:

Ok, now do you remember that I told you that the computer sometimes changes its mind? This time it wants you to match TEACHER to NON COMPLIANT, ANGRY ETC and EBD PUPIL to ACCEPTED, COOPERATIVE ETC. This is still only a practice just like the last time. But I want you to try and go as fast as you can, while still trying to get as many right as you can, Ok?

When you are comfortable and ready to go press the space bar, the first set of words will come up on the screen.

The instructional and presentation formats that accompanied the second block of trials were identical to the first, except that the participants were instructed to emit alternative response patterns.

### *Test Trials*

After the presentation of the automated feedback following the second practice block of trials, the participants were invited to take a short break prior to the first block of test trials. The first test block then commenced with the following verbal instructions:

Ok, now that you are familiar with task we will move on to the next stage. We are going to do this six more times, but as you can see you get through the stages really quickly, Now, this time I am not going to tell you which key to press. Ok. So, that means that the practice is now over and the computer is going to test you to see how much you remember. Ok? I will not tell you which tasks you are on you have to remember. But I will remain in the room with you. Is that ok?

Ok do you know which task you are on? (Researcher prompts if required only at the beginning of the first test trial) So, if are you comfortable and ready to go, Make sure you have your fingers on the right keys. Remember -- try and go as fast as you can but still trying to get them right. Ok, off you go and press the space bar to begin.

The participants were not provided with these specific instructions at the beginning of each block regarding the type of responding appropriate for that block (i.e. inconsistent or consistent – but these actual words were not employed), and they were not prompted throughout the block, as had been the case with the practice trials. At the end of the sixth and final test block, the Researcher indicated to the participant that the task was over and thanked them for their participation.