DEVELOPMENTAL DIFFERENCES IN THE ABILITY TO DECODE FACIAL EXPRESSION OF PAIN

by

Kathleen Susan Deyo

B.Sc., McMaster University, 1980

M.Sc. University of Waterloo, 1984

THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

PSYCHOLOGY

THE UNIVERSITY OF NORTHERN BRITISH COLUMBIA

JUNE 1999

0-612-62471-4

Approval

Name: Kathleen Susan Deyo

Degree: Master of Science

Thesis Title: DEVELOPMENTAL DIFFERENCES IN THE ABILITY TO DECODE FACIAL EXPRESSION OF PAIN

Examining Committee:

Chair: Donald Munton, Professor
International Studies

Supervisor: Kenneth Prkachin, Professor
Psychology

Member: Sherry Beaumont, Associate Professor
Psychology

Member: Judith Lapadat, Assistant Professor
Education

Lara Robinson PhD
External Examiner: Lara Robinson
Prince George Regional Hospital

Date Approved: June 14 / 99

ABSTRACT

The primary purpose of this thesis was to investigate age-related differences in the ability to recognize facial expression of pain. A secondary purpose was to examine the relationship between decoding performance and selected personality variables. A final exploratory purpose was to determine which facial cues were most predictive of participants' pain ratings. Previous research has indicated that the ability to decode facial expressions of emotion improves with increasing age throughout childhood. Little research has investigated variables that influence the recognition of pain expressions, and none have examined age differences. Thirty-three young adults and 102 children, in three age groups (six, nine and 12 years-of-age), viewed a videotape containing 90 two-second excerpts depicting clinical pain. Based on previous facial measurements, the excerpts fell into three intensity ranges - no, low and moderate to high pain. The participants' pain ratings were converted to sensitivity indices using signal detection theory. There was a linear increase in sensitivity with increasing age for children; however, the oldest children did not differ from the adults. As an additional way to evaluate observers' judgments, participants' ratings were correlated with patients' self-reported pain and facial actions. These results also indicated that participants' decoding performance improved with increasing age. Measures of empathy and self-perception were not systematically related to any dependent measures. The importance of different facial actions to pain ratings was examined using hierarchical regression analyses. These analyses indicated that fewer facial actions were predictive of pain ratings for the younger children in comparison to the other participants. Overall, these results imply that sensitivity to pain expression is largely developed by late childhood. The implications of this study were considered in terms of development, decoding methodology and the pain communication model.

Table Of Contents

List of Tables

Appendix D Tables

List of Figures

Acknowledgements

I must express my appreciation to the many people who have contributed to the completion to this thesis. First, I would like to thank my supervisor, Ken Prkachin, for his unique ability to provide guidance and allow autonomy throughout this endeavour. I appreciate his respect for my multiple roles (and associated responsibilities) and my compulsion to take more courses; which undeniably delayed progress. Secondly, I would like to extend my gratitude to my other two committee members, Sherry Beaumont and Judith Lapadat, for their support, insightful comments, constructive feedback and attention to detail. I am thankful for all the time and effort everyone on the committee devoted.

The videotape for this project was constructed by Reiko Graham. Her editing expertise is gratefully acknowledged. Recruiting participants was facilitated with the assistance of the staff at Highglen Elementary School and the Science Alliance and Y.M.C.A. Day Camps. I thank all of the children and families who volunteered to participate and allowed me into their homes to collect data.

Finally, I need to express my sincerest appreciation to my friends and family for their support and encouragement. My friends have been willing to listen and share their ideas. I thank my daughter Elizabeth for the inspiration for this thesis. According to the words on the yellowed comic strip on our fridge - "no accomplishment is without sacrifice" (Lynn Johnston). It would not have been possible without the love and patience of my family.

INTRODUCTION

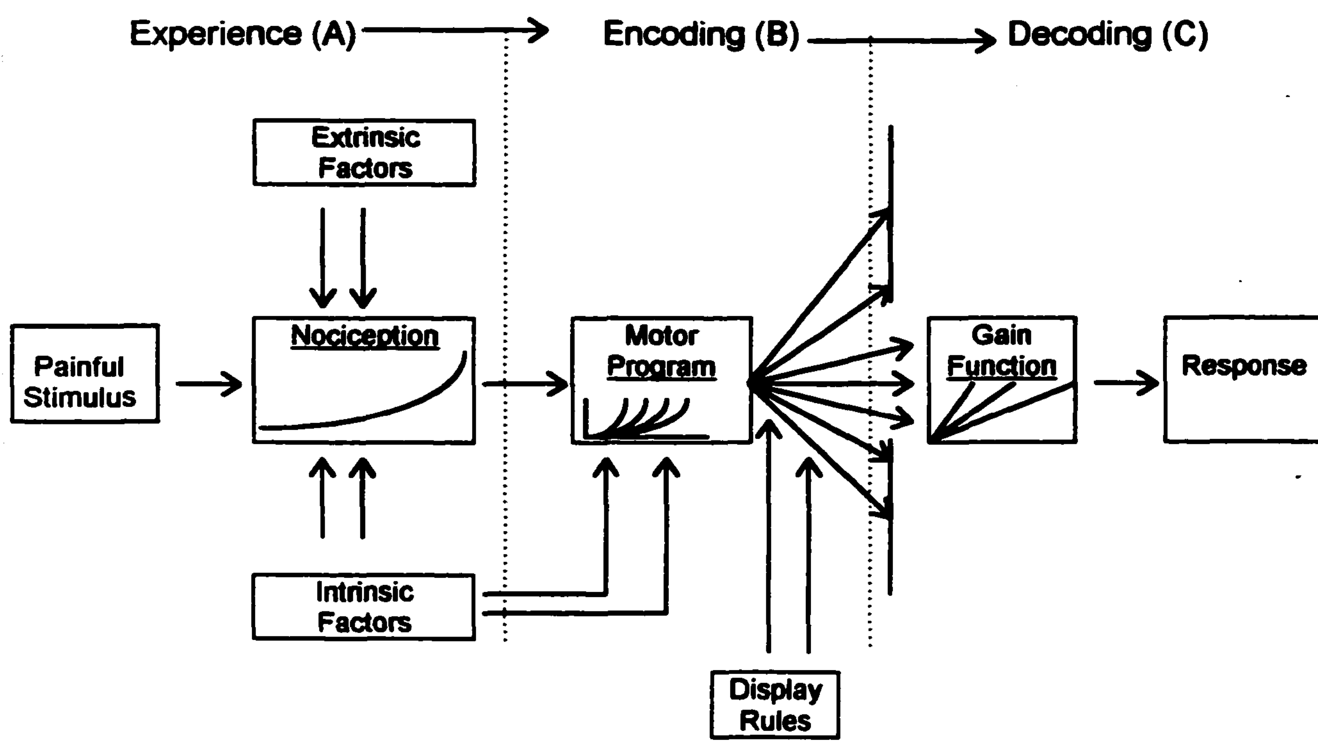
The primary purpose of this thesis was to investigate developmental differences in the perception of facial expression of pain using a signal detection theory (SDT) approach. The SDT rationale and methodology will be described in detail in subsequent chapters. The introductory section highlights research findings about the development of the ability to decode facial expressions of emotion in general and the influence of social variables on decoding performance focusing on previous methods employed. This section will be followed by a description of the prototypic facial expression of pain and a brief discussion of the relevant literature. Then, the results of studies that have examined characteristics of adult observers that influence their judgments of pain expressions is presented. Hypotheses about the development of the ability to decode facial expression of pain were based on the facial expression of emotions literature. The method section describes, in detail, the procedures used to examine sensitivity to pain expressions. The results initially focus on the signal detection analyses and will then address a number of ancillary analyses. The implications of the study are discussed in terms of development, methodology and the pain expression model.

The ability to recognize facial expressions of pain has important clinical and social implications (Prkachin & Craig, 1995). For the clinician, facial expressions are often important in the diagnosis of a medical problem and evaluation of treatment (Manne, Jacobsen, & Redd, 1992). The importance of clinicians' abilities to recognize facial displays of pain is especially true for infants, those with handicaps that preclude accurate verbal reporting and patients for whom distorting a verbal report is a concern (Bieri, Reeve, Champion, Addicoat, & Ziegler, 1990; Craig, Hyde, & Patrick, 1991). In the social environment, expressing pain or distress may serve as a warning of threat or may solicit helping behaviour on the part of the observer (Prkachin & Craig, 1995). Research is beginning to demonstrate that, at least among adults, there are individual differences in the way that people interpret others' facial expressions of pain (e.g., Prkachin & Craig, 1995; von Baeyer, Johnson & McMillan, 1984). Nevertheless, to date, no study has addressed differences between adults and children of various ages in their detection and interpretation of pain expressions.

Prkachin and Craig (1995) have proposed a model of the nonverbal communication of pain that will form the basic theoretical framework for the present study. Their model integrates a general model of nonverbal communication (Rosenthal, 1982) with Ekman's (1977) neurocultural model of emotion and recent data on pain expression. In general, nonverbal communication depends on two aspects of performance that influence the effectiveness of a communication (Zaidel & Mehrabian, 1969). Encoding refers to the process that occurs when a person converts his or her emotion into a facial expression. Good encoders are able to emit clearly discriminable cues to their emotions in their facial expressions. In contrast, decoding is the process of discriminating different facial cues. A competent decoder is able to discriminate different feelings from a variety of cues.

According to the pain model (Prkachin & Craig, 1995), the three processes that may occur during an episode of pain are experiential (A), encoding, (B) and decoding (C) (Figure 1). The experience begins when the

Figure 1. Pain Communication Model



person is exposed to a stimulus that exceeds the pain threshold. The person then may encode this experience in a facial expression that is transmitted to the social environment. This expression is then decoded by observers. Decoding refers to both the detection and interpretation of the facial expression of pain. At each stage of the model, there are a number of variables that may affect the pain display or the interpretation of it.

Although there are a large number of variables that affect the initial experience of pain, these will not be addressed here. The present discussion will be limited to the encoding and decoding of pain expressions. Furthermore, encoding will be discussed insofar as it is necessary to demonstrate the existence of a "universal" or prototypical pain expression. Then, variables that are related to the accurate judgment of pain expressions will be addressed.

Facial expression of emotion: General considerations

First of all, it should be made clear that pain is not being interpreted as an emotion. Nevertheless, referring to the emotion expression research is appropriate because, like pain, emotional expressions have an interpersonal function; to convey feelings (Levenson, 1994; cited in Levenson, 1996). According to Schweder, (1994) pain is a nonemotional feeling that is "associated with perceptible facial icons" (p. 39).

There is ample evidence to show that nonverbal cues (e.g., gestures, facial expressions) are critical to effective interpersonal communication. They transmit information to the social world that may emphasize, complement or contradict verbal messages (Zimbardo & Leippe, 1991; Bugental, Kaswan, & Love, 1970; Riggio, 1992). In fact, Philippot, Feldman, and McGee (1992) suggest that facial expression may be more effective than verbal cues in revealing an individual's internal state. Presumably, this suggestion is because nonverbal behaviour is assumed to be spontaneous and under less voluntary control than verbal behaviour (Egan, 1986; Levenson, 1994).

At this point, it is also important to address the concept of universality of facial expressions of emotion and how universality relates to the development of encoding ability. If facial expressions of emotion are "universal," then encoding of them should occur naturally given the appropriate stimuli and maturity of the neural and muscular control mechanisms. Therefore, a discussion of "universal" and "encoding ability" will be provided below.

There is considerable debate in the literature (see Russell, 1995, for a review) about the degree of universality of facial expressions of emotion. Nevertheless, there is agreement that there is at least "minimal universality" (Russell, 1995). According to this position, across cultures, humans produce the same facial expressions in similar situations (actual or imagined), and they use the same emotion labels (or equivalent translations for speakers of different languages) when asked to decode facial expressions in photographs. Ekman (1973; cited in Russell 1995) claims that the universal facial expressions are "happiness, sadness, anger, fear, disgust, and surprise" (p. 380).

Most researchers also agree that cultural or social display rules will influence what different people will express (e.g., Russell, 1995). According to this position, humans may be predisposed to reveal their emotions in recognizable facial expressions, but there is significant voluntary control over the actual expression that is displayed. Display rules are unwritten guidelines that govern how a person is expected to reveal one's emotion. For example, in our culture, display rules dictate that males and females should control their expression of anger or sadness in different ways. The use of display rules demonstrates that expression of emotion is, to some degree, under voluntary control. Rinn (1991) indicates that spontaneous or involuntary facial expressions and voluntary expressions are under different types of neurological control; which would suggest that encoding ability could be different for involuntary and voluntary expressions.

The majority of the research concerning age-related changes in encoding, to be described below, involves the posing of expressions of emotion rather than the spontaneous expression that accompanies an emotion. The ability to pose nonemotional expressions (i.e., facial expressions of emotion that are not associated with the experience of emotion) leads to questions concerning the role of either verbal ability or knowledge of emotion vocabulary in the performance of the encoding tasks. Is a good encoder able to control facial musculature better than a poor encoder? Or, does a good encoder have a better understanding of emotion words and expectations about how to demonstrate that awareness?

There is one additional point to be aware of when considering the facial expression research. It appears that encoding and decoding accuracy are often confounded. For example, encoding accuracy is often operationally defined as a function of the degree of consensus among the observers (i.e., decoding agreement). As will be shown below, in a typical encoding study, children will be asked to pose the emotion that is appropriate for a given situation. Then, their accuracy at encoding the expression is assessed by the judgments of adult (e.g., university students) observers. Therefore, the decoders' judgments determine encoding accuracy.

Encoding and decoding of universal facial expressions of emotion

Given the significance of facial expressions of emotion to interpersonal communication, it is necessary to determine how the ability to encode and decode them develops. Developmental changes in the decoding of universal expressions of emotion will be emphasized as they are relevant to the hypotheses and design of this study. In addition, the methods typically used to assess encoding and decoding performance will be described because they differ from the measures employed in this study. Specifically, most facial expression research involves the discrimination among different expressions, whereas in this study, the participants' task was to discriminate between different intensities of the same expression (i.e., a pain expression). A SDT approach was employed in this study for two reasons. First of all, in general, the SDT method allows for the calculation of independent sensitivity and bias indices (See, Warm, Dember & Howe, 1997). According to Ellermeier (1997), the willingness to report pain (i.e., bias) affects pain ratings and therefore should be distinguished from the sensory or sensitivity factor. Secondly, the SDT method has been used previously by Prkachin (1992b; Prkachin & Craig, 1985) to examine observers' sensitivity to expressions of pain resulting from exposure to different intensities of painful stimulation.

With respect to facial expression of emotions, both encoding (i.e., often of posed expressions) and decoding improve during childhood until adult performance is achieved. Adults, in general, perform at nearly perfect levels[1] and sometimes, but not always, there are sex differences in performance. For example, Zuckerman, Lipets, Koivumaki & Rosenthal (1975) found that women were slightly better encoders and significantly better decoders than men for the six universal facial expressions of emotion.

One of the earliest studies examined children's production and discrimination of eight facial expressions (Odom & Lemond, 1972). The materials for this study were a total of 32 black and white photographs with each of the eight expressions (i.e., fear, anger, distress, shame, disgust, surprise, joy and interest) represented four times. Children in kindergarten and grade five participated in one of two discrimination and one of two production tasks. In the matching-discrimination task, the child was to select a photograph, from a series of four provided at one time, to match the way the person in the standard photo felt. For the situation-discrimination task the child was told that the people in the four photographs being shown felt a different way. The experimenter read a situation, and the child was to choose which photograph corresponded to it. For

[1] Encoding accuracy is the percentage of facial expressions that are correctly identified by observers, whereas decoding accuracy is the correct identification of a universal facial expression or proportion of agreement among observers.

the production tasks, the children were told to make faces as well as they could with their faces in a frame so that their picture could be taken. In the imitation-production task, the children were shown two photographs of the same expression and were instructed to make a face like those shown. In the situation-production task, the experimenter read situations and the children were asked to make a face like they would feel if they were in the specific situation. Results demonstrated that grade five children were more accurate on both encoding and decoding tasks than kindergarten children. There were no significant effects of gender on either task. Discrimination was more accurate than production indicating that children were better at recognizing than posing facial expressions. Furthermore, because the older children made errors, the authors suggested that maximum sensitivity had not been attained by the 10-year-old children.

Similarly, Profyt and Whissell (1991) read stories to children aged 4 - 6 years and asked them to pose how the child in each of the stories would feel. The children's poses were videotaped. One week later, the tapes were shown to the child, another child and to adults. Overall, encoding accuracy (i.e., recognizability of the expressions produced) increased with increasing age. Performance for girls and boys was the same with the exception that girls were better at encoding the fear expression.

With the exception of one study to be described later (Kolb, Wilson, & Taylor, 1992), most research suggests that differences in encoding and decoding abilities are a function of social influences or social skill. For example, Tucker and Riggio (1988) examined the relationship between social skills in encoding of posed and spontaneous facial expressions by adults. The spontaneous expressions were videotaped unobtrusively while subjects viewed slides that were chosen to arouse three emotions (disgust, happiness and sadness). The participants were informed of the videotaping. They were then given cards with an emotion and a neutral message to recite and were instructed to pose the given emotion. The Social Skills Inventory, which measures four aspects of social skill, was administered following the videotaping. Groups of three judges rated the segments for which emotion a participant was expressing. Encoding accuracy was defined as the percentage of judges who correctly identified which emotion the participant was expressing. Adults with higher social skills scores were superior encoders of posed expressions in comparison to those with lower scores. Spontaneous expressions were less related to social skills than were the posed expressions.

In contrast to the previously discussed study which included adults, Feldman, White, and Lobato (1982) conducted two studies that investigated the relationship between social skill and encoding and decoding abilities of children. In the first study, boys and girls ranging from five to 12 years-of-age were administered a role-taking task as a measure of social skill. Then, each child's nonverbal encoding skill was evaluated. They were instructed to sample two drinks (sweetened and unsweetened) and try to convince (i.e., fool) an interviewer that they either liked or disliked both drinks. The children were being videotaped during their interviews. Untrained observers rated the children in segments as truthful or deceptive. A deception ability score, the measure of encoding ability, was the percentage of observers who identified as truthful a child who was being deceptive. Role-taking skill was significantly correlated with encoding ability and this was more pronounced for girls than for boys. Unlike other studies, encoding ability did not increase with age. However, role-taking did improve with age.

In their second study, Feldman et al. (1982) examined encoding and decoding ability of normal and institutionalized, "emotionally disturbed" adolescents. Seventeen specific social competencies were rated by teachers or counsellors as the indicator of social skill. Participants viewed three two-minute videotapes that were intended to elicit positive, negative and neutral emotional responses. The adolescents were videotaped while they viewed the tapes. Encoding was defined as undergraduate judges' ratings of the participants' facial expressions on a 5-point unpleasant-pleasant scale. During the decoding phase of the study, the participants rated the facial expressions of ten undergraduates viewing the same tapes. Results indicated that normal adolescents were better encoders and decoders than "emotionally disturbed" adolescents. Furthermore, social skill was correlated with encoding but not with decoding.

In Feldman's earlier research (i.e., Feldman, et al., 1982), social competencies were rated by teachers or counsellors. In his later work, he and his colleagues have used the social competence scale of the Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1982). For example, Custrini and Feldman (1989) investigated the relationship between social competence, as measured by the CBCL, and encoding and decoding performance of 9 - 12 year olds. Naturally occurring expressions, elicited during viewing of videos, were used for the encoding task. Encoding accuracy was determined by undergraduate judges. The children's decoding task was to view videotapes of university students' reactions to videos and to indicate which of five emotional expressions was being displayed. The videotaped segments had been previously rated by undergraduate judges. A child's response was considered to be correct if it agreed with the students' judgments. Decoding was more accurate than encoding for all five emotion categories assessed. An important finding was the significant interaction between social competence and gender. Boys, in both high and low social competence groups, performed with the same degree of accuracy. However, there was a difference in the performance of girls. High social competence girls performed better than boys, whereas low social competence girls were worse than the boys. It should be noted, however, that there were only three girls in the low social competence group.

According to Feldman, Philippot and Custrini (1991), there is no definition of social competence that is acceptable to most researchers or theoreticians. Many terms are considered synonymous, and many aspects of social behaviour have been studied. Feldman and his colleagues (e.g., Feldman et al., 1991) assume that various social skills underlie competence (e.g., the ability to manage impressions, communication behaviours such as patterns of eye contact or voice intonation). Therefore, they consider social competence as a multidimensional domain. From their perspective, "both decoding and encoding skills can be viewed as manifestations of social competence" (p. 331).

The remaining studies to discuss will focus on decoding. Examining infants' recognition of facial expressions requires methods quite different from those employed with older children and adults. For example, looking times are commonly used to assess either recognition or memory. It has been found that four-month-old infants' looking times are longer for happy expressions than for angry or neutral expressions (Labarbera, Izard, Vietze & Parisi, 1976; cited in Gosselin, 1995). Furthermore, by 5 months-of-age, infants are able to discriminate happy and sad facial expressions as measured by a habituation/dishabituation method of assessing infant memory (Walker-Andrews & Lennon, 1991). In the habituation phase of this procedure, infants were presented slides of one facial expression repeatedly until looking time decreased to a criterion amount of initial looking time. The infants were then presented with a new slide that either matched or mismatched the expression on the previous slides. Looking times were longer for mismatched expressions suggesting that the infants recognized that the mismatched expression was different. Nelson and de Haan (1996) investigated whether 7-month-old infants' brain activity was influenced by the type of emotion expressed in a face.

Electroencephalogram (EEG) recordings were taken while infants were presented slides of either happy or fearful faces or fearful or angry expressions. They demonstrated that the infants were able to discriminate happy from fearful expressions but not angry from fearful expressions. Therefore, the ability to discriminate different facial expressions is present early in life. Gosselin (1995) suggests that this discrimination of facial expressions of emotion is the foundation of social competence that appears early in development.

It appears that decoding ability continues to develop throughout childhood. For example, Missaghi-Lakshman and Whissell (1991) had children in grades 2, 4, and 7 draw faces for the six universal expressions. Two weeks later, the faces were decoded by adults and the child who drew the faces. Decoding accuracy improved with increasing grade. There were no differences between boys and girls. Similarly, Tremblay, Kirouac, and Dore (1987) observed that children's decoding ability improved with age. Their methodology was different in that children were presented with photographs of the universal facial expressions displayed by adults and children of both sexes. They observed that decoding performance was not affected by either the age of the model nor the sex of the observer.

Other researchers have demonstrated that decoding ability is related to level of intelligence. Xeromeritou (1992) investigated the decoding ability of educable mentally retarded and nonretarded controls. A short story was read to the children. The children identified a picture for how the person in the story felt.

The story was read again, and the children then produced the word for how the person felt. When these groups were matched for verbal mental age, those with the older mental age were more accurate than children with a younger mental age. Simon, Rosen, Grossman & Pratowski, (1995) observed that facial expression recognition was positively correlated with IQ for adults with mild to moderate mental retardation and was unrelated to measures of social skill. The facial recognition task was to identify which person, from a group of six photographs, exhibited a specific emotion. These results suggest that facial expression recognition is a cognitive process or is dependent on verbal ability as it is assessed by IQ tests.

Gosselin (1995) stated that although decoding is known to improve with development, little is known about how these improvements occur. He designed his study to investigate the types of decoding errors children committed. In his study, short stories were read to children of two age groups (5 - 6 year olds, and 7 - 8 year olds). They then selected the photo that best matched how the character in the story felt. There were no age differences in decoding happy expressions. However, the older children were more accurate for the remaining five expressions. Using knowledge of Facial Action Coding System (FACS; Ekman & Friesen, 1978) for the various expressions, he was able to conclude that improvements were due to a reduction in certain types of errors. In order to explain the change of errors, a very basic understanding of FACS is required. FACS is an objective, anatomically based system to describe the actions of the face. There are 44 codable facial action units (AU). Each of the universal facial expressions is characterized by a combination of AUs. There is some overlap of AUs across emotion expressions. That is, some emotional expressions contain the same action units (AU) as other expressions. For example, brow lowering is present in sadness, anger and fear, and raising of the upper lip occurs with both anger and disgust. Gosselin (1995) observed that with an increase in age, children became more sensitive to combinations of facial components that convey different emotions.

Consistent with the encoding research presented previously, a number of studies suggest that social variables influence decoding performance. For example, normal boys with unhappily married parents have been found to be significantly worse at decoding five facial expressions of emotion than were girls or boys with happily married parents (Shortt, Bush, McCabe, Gottman & Katz, 1994). In order to demonstrate the importance of social skills on nonverbal behaviour, Feldman et al., (1991) cited a number of studies that demonstrated that a variety of "disturbed" individuals (i.e., delinquents, abused children, psychiatric patients), did not decode photographs of facial expressions as accurately as those without disorders.

The facial expression recognition abilities of children with a specific disorder, Attention-Deficit Hyperactivity Disorder (ADHD), were investigated by Singh et al (1998). Children with ADHD were read two-sentence stories that included the target emotion word and asked to identify, from a selection of six photographs, the one that showed the emotion described. ADHD children were observed to have deficits in the ability to recognize the six universal emotion expressions, in comparison to norms for non-ADHD children (Singh, et al., 1998). It was suggested that ADHD children have social skill deficits that could be attributed to differences in social interactions in comparison to non-ADHD children.

Children's interactions with their peers can be assessed with peer ratings of specific behaviours or by sociometric methods. According to Feldman et al. (1991) sociometric measures are the most common method of assessing social competence[2]. They reviewed research that demonstrated that sociometric status was associated with decoding accuracy in the expected direction. For example, Walden and Field (1990) observed that sociometric preference scores were significantly related to facial expression discrimination scores. Discrimination was defined as the preschoolers' ability to match a standard face with one from a group of five.

Similarly, Philippot and Feldman (1990) examined the relationships among age, social competence and decoding of facial expressions by preschoolers. The task was to watch a videotaped scenario, with the main character's face blacked out, and then select an emotion face (happy, sad, or

---

[2] Although they do not state how sociometric methods are used to draw conclusions about social competence, it is possible to speculate on the relationship between social competence and sociometric status. One might assume that relatively popular children are more skilled in the types of behaviours necessary to establish and maintain positive interpersonal relationships. Conversely, disliked children may be likely to engage in social behaviours that are objectionable to their peers.

afraid) that was most appropriate. High social competence children were better decoders independent of age, sex or emotion. There was a sex by age interaction. Girls improved their performance between three and four years, whereas boys improved between four and five years. That is, boys were delayed by one year relative to girls. Although, this study was presented as one involving decoding, the children's task was not to decode a given facial expression, but rather to understand the scenario and select the facial expression that was most suitable for the given situation.

Beck and Feldman (1989) also suggested that decoding ability is related to social skills. Furthermore, they proposed that ability differences are primarily a result of learning processes. Therefore, they systematically attempted to increase competence in decoding of emotional expressions in children in grades five to seven. Improved competence was accomplished by providing feedback to one group and not providing it to the other. There were no sex differences of decoding, but there was a significant effect of feedback. Boys and girls in the feedback condition improved their accuracy at decoding different emotional expressions.

Two possible explanations for the relationship between social competence and decoding ability have been proposed. Children who are poor at discriminating nonverbal cues may miss important information about others that impairs their social competence. Conversely, those with poor social skills may not have the same opportunities, due to more limited social interactions, to acquire the ability to differentiate facial expressions. However, it is also possible that both skills are due to some third factor such as a concern for others. For example, Eisenberg et al. (1996) suggest that empathy or sympathy is related to social functioning.

In contrast to the above studies which propose that social variables are the primary influence on decoding ability, Kolb et al. (1992) suggest that improvements with age may reflect maturation of the frontal lobes. In their photograph-matching and cartoon-matching tasks, they observed improvement until approximately 14 years-of-age. This improvement was not gradual; it increased at about eight years-of-age and then again around 13 or 14 years. Adult patients with frontal lobe lesions performed the tasks with the same degree of accuracy as eight- to 13-year-olds. Because of the similarity in decoding accuracy for these two groups, they suggested that the development of decoding ability parallels frontal lobe maturity.

There are a few points that need to be considered in the evaluation of the emotion expression literature before summarizing it. First of all, there is a large body of literature that involves some aspect of development and emotional expression. Earlier literature (e.g., Odom & Lemond, 1972) focused on age differences in producing and discriminating expressions. Then, research examined encoding and decoding performance as a function of age in combination with other variables (e.g., social skill; Custrini & Feldman, 1989). More recently, there has been an emphasis on the effects of context on encoding and decoding of emotions (e.g., Russell, 1997). Therefore, the selection of studies described above was a representative sample of the literature that addressed the effect of age on decoding performance. The second comment concerns the fact that most tasks in the research are very language-laden. In order to perform well on the typical tasks, one must have good verbal skills or have a good understanding of emotion vocabulary. However, proficient nonverbal communication should not require a language-based test to assess it. Therefore, the verbal nature of the tasks may confound encoding and decoding performance. Finally, most studies use deliberate or posed expressions rather than spontaneous emotional expressions. There is evidence to suggest that these two types of expressions may have different characteristics (e.g., symmetry or latency; Hager & Ekman, 1997) and that different neurological systems control voluntary and involuntary expressions (Rinn, 1991). Furthermore, Rinn (1991) states that both systems tend to act simultaneously but with varying degrees of influence on the expression displayed. Display rules, which are learned, govern voluntary expressions. In sum, it would appear that the facial expression literature is somewhat difficult to interpret due to a variety of potential confounds (e.g., verbal tasks, the role of learning in posed expressions).

Despite the foregoing criticisms, the literature on the development of encoding and decoding of facial expressions suggests the following conclusions. First of all, both processes improve with increasing age. However, it is possible that different mechanisms may account for improvements in encoding and decoding. It has not yet been established at what age adult levels of performance are achieved. Few studies have been conducted with children older than ten years of age. Gender differences are not always observed, but when they are, girls tend to be more accurate than boys. Finally, more socially skilled individuals appear to be more accurate in their recognition of facial expressions of emotion.

Facial Expression of Pain

Encoding

A basic assumption underlying the preceding work is that there are universal facial expressions of emotion. Studies employ prototypical emotion expressions as the standard for judging decoding accuracy. The primary purpose of the present study was to examine developmental differences in the recognition of facial expressions of pain. In order to make judgments about the decoding of pain expression, there is also the assumption that there is a universal or prototypic facial expression of pain. The following section reviews evidence that analogous to the facial expressions of emotion, there is also a prototypic facial expression of pain.

A small number of North American research groups have used FACS (Ekman & Friesen, 1978) to describe the facial actions of individuals in pain. Of the 44 possible facial actions only a fraction of them have been found to be consistently related to acute pain. Prkachin and Craig and their colleagues (e.g., Prkachin & Mercer, 1989; Craig et al., 1991; Poole & Craig, 1992; Prkachin, Berzins & Mercer, 1994) have described a cluster of facial actions that are correlated with subjects' self-report of pain. These actions include tightening of the orbit (by raising the cheeks; FACS action units (AU 6 and 7), lowering the brow (AU4), wrinkling of the nose (AU9) or raising the upper lip (AU10). The eyes may close (AU43). Movements around the mouth are common, but are not reliably associated with specific painful stimuli or pain reports. The action units, associated muscles and facial movements are described in Appendix A. Prkachin (1992a) has described this pattern of facial actions as prototypical.

LeResche (1982) used 16 candid photographs to determine which facial actions occurred most frequently during pain. Her results also indicated that brow lowering, orbit tightening and eye closure were very common. In comparison to the videotapes of pain induced during clinical examinations, she observed less nose wrinkling and upper lip raising, whereas she observed a greater degree of mouth opening (AUs 25, 26, and 27) and lip stretching (AU20)

Further support for the notion that the components of the expression are universal are the findings on pain expressions in infants. Infants, premature and full-term, exhibit a pattern of facial actions that is very similar to that observed in adults. Even very premature infants exhibit the pattern of actions albeit less vigourously than older infants. Specifically these actions are - brow bulge, eye squeeze, naso-labial furrow and open mouth (Grunau, Johnston, & Craig, 1990; Johnston, Stevens, Craig & Grunau, 1993; Craig, Hadjistavropoulos, Grunau, &

Whitfield, 1994; Stevens, Johnston, & Horton, 1994; Johnston, Stevens, Yang & Horton, 1995).

It is evident that people have some ability to voluntarily control their display of pain. In conditions where there is experimental (Prkachin, 1992b) or clinical pain (Poole & Craig, 1992; Craig et al., 1991; Hadjistavropoulos, Craig, Hadjistavropoulos, and Poole, 1996; Galin & Thorn, 1993), participants have been able to mask and exaggerate their pain displays. FACS coding of the relevant action units demonstrated that there was less facial activity of pain in masking conditions and greater activity in the exaggerate conditions. Although there were subtle cues to pain intensity available, for the most part, it was possible to exert control over facial expressions of pain (Rinn, 1991).

Decoding

Adult observers are able to discriminate different intensities of pain from the relevant facial actions. That is, observers' ratings of pain are significantly accounted for by the FACS coded actions. In one of the earliest studies on observers' reactions, Patrick, Craig and Prkachin (1986) investigated the relationship between a model's pain tolerance, subjective pain reports and observers' ratings. In the first phase of the study, three groups of participants were videotaped while they received electric shocks of increasing intensity with a confederate present. The confederate modeled either tolerance or intolerance to the shocks or remained seated but was not a coparticipant. Following the ascending series of shocks, a random series was presented varying intensity level according to discomfort ratings from the previous series. In the second phase, untrained observers rated the level of discomfort they believed the participants were experiencing. Participants exposed to pain tolerant models reported a similar degree of pain as those exposed to an intolerant model even though they experienced more intense shock. Despite the comparable ratings of discomfort, their facial expressions displayed more activity associated with pain. Furthermore, the observers rated the subjects with the tolerant model as experiencing more pain which was consistent with the greater intensity of shock received.

In contrast to the study of laboratory pain described above, Prkachin et al. (1994) employed videotapes of clinical pain. Five judges were shown a videotape of patients with shoulder pain undergoing an examination of the shoulder by a physiotherapist. Patients and observers rated the patients' pain on the same scales. Facial expressions were measured using FACS (Ekman & Friesen, 1978). Their results indicated that observers' pain ratings were correlated with patients' pain ratings for severe pain but not when the pain was submaximal. Despite the variable degree of correlation between observers' and patients' ratings, the facial actions were consistently correlated with the patients' pain ratings. These results suggest that although the information necessary to draw inferences about the person's state was available to the observers, they did not make appropriate use of more subtle cues.

As indicated previously, facial expressions of pain are, to a certain degree, under voluntary control. Given that Craig et al. (1991) found that genuine pain expressions could be discriminated from suppressed and faked displays on the basis of FACS coded expressions, Poole and Craig (1992) examined how observers would rate dissimulated (i.e., masked or exaggerated) pain. Would they rate pain intensity according to the facial display or would they be able to identify the actual pain experienced? Results indicated that observers rated more pain in the genuine, suppressed and faked conditions than in the baseline (no pain) condition. The genuine condition differed significantly from both dissimulated conditions. Specifically, they rated more pain for the exaggerated faces and less pain was attributed to the suppressed faces relative to the genuine condition. However, the patients were not totally successful in deceiving the observers. At least some indication of pain was evident in their facial displays. For example, for the suppression condition, observers gave higher ratings of pain than in the baseline condition. The results suggest that observers' pain ratings were related to facial activity even when dissimulated. There were no gender differences in the amount of pain observers perceived in the various conditions.

In another study involving dissimulated pain, Hadjistavropoulos et al. (1996) asked participants to differentiate among no pain, genuine pain, masked pain and exaggerated pain of patients with low back pain. The videotapes of these four conditions were taken at the patients' clinic. In the pain conditions, patients were instructed to lift their legs off the table and to genuinely express, suppress or exaggerate the pain they were experiencing. The order of pain conditions was counterbalanced. Participants were informed that, for each patient, there were the four conditions. Their task was to classify the segments into one of the categories. They then rated the intensity of the patient's pain. No pain and exaggerated pain were most accurately classified by observers. However, masked and genuine pain were only correctly identified at a chance level. Therefore, it appears that even though patients were attempting to hide pain, certain cues were still present that were expressed. On the other hand, extreme pain expressions were more readily identified as such. Some facial cues (e.g., brow raising and lip corner pull) were more likely to occur with exaggerated faces and observers made systematic use of them to classify the video excerpts.

Prkachin and Craig (1995) suggest in their model that a "gain function" exists that affects an observer's likelihood of reporting pain (Figure 1). Some observers would report pain with minimal evidence whereas others would require considerable evidence before they would decide that someone is in pain. Overall, Prkachin (e.g., Prkachin et al., 1994; Prkachin et al., 1983) has found that observers demonstrate an underestimation bias. That is, observers tend to report that a person is experiencing less pain than he or she reports.

It has been suggested that experience with pain sufferers may influence the ratings by observers and therefore may affect the underestimation bias.

Subjects who lived with a family member suffering from chronic pain had a diminished underestimation bias (Prkachin, Solomon, Hwang & Mercer, 1995). On the other hand, nurses (von Baeyer, et al., 1984) and physiotherapy and occupational therapy students (Prkachin, et al., 1995) had an exaggerated underestimation bias. Taken together, these results imply that experience per se is not reliably related to bias. Rather, the nature of the observers' experience affects their bias. There is evidence that, in general, attributions for actors and observers are more similar if they like or identify with each other (Regan & Totten, 1975). Differences in attributions may explain why the family member would be more likely to provide ratings closer to the patients' than unfamiliar observers would. It has also been suggested that clinicians "distance" themselves from the suffering of their patients and this may result in ignoring or minimizing evidence of their patients' distress (Prkachin, et al., 1995).

Hadjistavropoulos et al. (1996) suggested that research should address the possibility that "there may be some individual difference variable that would make some judges better decoders than others" (p. 257). Nurturance has been found to be related to expressions of pain (Von Baeyer et al., 1984). Intuitively, it seems logical that empathy, a similar construct, would be related to sensitivity to pain expressions. Solomon (1995) used a multidimensional measure of empathy, the Interpersonal Reactivity Index (IRI; Davis, 1983), to determine which types of empathy were related to students' ratings of shoulder pain patients' ratings. Of the four types of empathy, fantasy was significantly related to observers' ratings and empathic concern was marginally significant (p = .07). Fantasy (F) is the tendency to identify with fictitious characters and empathic concern (EC) is the tendency to experience feelings of warmth, compassion and concern for others undergoing negative experiences.

Eisenberg et al. (1996) have used a simplified version of the IRI, adapted for children, in their study of the relationships among empathy/sympathy, social functioning and physiological responses to a videotape of a child in distress. In addition, they observed that children high in sympathy were also rated as more socially competent by their peers. In earlier work, Eisenberg and her colleagues have found that sympathy is related to helpfulness (Fabes, Eisenberg & Eisenbud, 1993), parental characteristics and coping behaviours (Eisenberg et al., 1991) and physiological responding (Fabes et al., 1993; Eisenberg, et al., 1989).

An early study by Zahn-Waxler, Radke-Yarrow and King (1979) examined the relationship between maternal behaviour and children's reactions to another child's distress. They found that children whose mothers used affective explanations when their 1 ½ - 2 ½ year old children caused or witnessed distress were significantly more likely to react prosocially (e.g. help or hug the other child). Distress included behaviours other than facial expression. Mothers who were rated as higher in empathic caregiving had children who were more likely to behave altruistically, and their actions were often accompanied by concerned emotional expressions. The authors concluded that maternal disciplinary techniques lay down the foundation for responsiveness to feelings of others. Although this study was not a decoding study per se, it does demonstrate that the ability to perceive distress may be present early in life and also that the early social environment influences sensitivity to others.

Chun, Turner and Romano (1993) discovered that children who lived with a parent with a chronic pain problem were rated by teachers as being less socially competent than children of healthy controls. Because social competence was related to pain exposure, it may also be related to sensitivity to pain expressions.

There have been no studies of the ability of children to decode facial expression of people in pain. However, Shih and von Baeyer (1994) ascertained that preschoolers were able to make gross distinctions between drawings of faces in pain.

Measurement of Pain with Children

The accurate assessment of pain in adults can be challenging, and probably even more so for children. Pain is a multidimensional experience; therefore, it should be assessed multidimensionally (Abu-Saad, 1994). The assessment of it must "conform to the communication capabilities of the suffering person, whether infant, verbal child, effectual adult, or incompetent adult" (Anand & Craig, 1996, p. 5). Given the foregoing statements, it is apparent that self-report measures, despite being the "gold standard" of assessment, should be supplemented with additional measures.

For the most part, knowledge about a child's pain is based on observation of the face and behaviour (Abu-Saad, 1994). However, self-report measures are also employed. The participants in this study were required to provide ratings of others' pain. In order to rate pain in others, it would seem logical that, at the very least, one should be able to provide a self-report of pain. Therefore, a brief description of self-report procedures will be presented.

There are three ways children have provided self- report of pain. The first is to give a verbal report of their experience. Studies by Abu-Saad (1994) and Wilkie et al. (1990) indicate that normal children over the age of about seven or eight are able to use different descriptors to characterize their pain. Although self-report is the 'gold standard' for adults, the limited verbal ability of younger children has precluded its use with them. Visual analogue scales (VAS) have been used with some success with children for rating the intensity of their pain (Abu-Saad, 1994). However, Bieri, et al. (1990) assert that these are not appropriate for the preoperational child for a variety of reasons (e.g., limited ability to understand that a line represents pain intensity). Therefore, they developed the Faces Pain Scale for children's self-report of pain severity. This 7-point scale includes line drawings, derived from children's drawings, of faces displaying "no pain" to "the most pain possible." The child is required to select the face that best illustrates a pain experience. Manne et al. (1992) used an approach similar to Bieri et al. (1990) except that the children used a 5-point scale with associated faces. Both studies indicated that this type of instrument could be used with children as young as three years of age. Unfortunately, a faces scale would be inappropriate in the present study because a 'face' should not be used to rate a facial expression. It would not be clear whether a face on the scale was being used to rate the intensity of pain or to match the patients' face. Therefore, because of the limitations of the VAS and faces scales it was necessary to use a verbal scale for rating pain intensity. Although a 7-point scale is more sensitive to variations in intensity, it was expected that 5-year-old children might have difficulties with such fine discriminations.

Present Study

The purpose of this study was to investigate age-related differences in the sensitivity to facial expressions of pain because, to date, there have been no reported studies that included children as observers of others' facial expressions of pain. In addition, response bias was examined to determine whether there were age-related differences in the willingness to report pain in others. Children of three different age groups were compared to young adults. Measures of self-perception and empathy were included as covariates to examine whether these individual difference characteristics influenced either sensitivity or bias in pain ratings. Sensitivity and response bias were addressed with a signal detection paradigm. Prkachin and his colleagues (Prkachin, Currie, & Craig, 1983; Prkachin & Craig, 1985; Prkachin, 1992) have demonstrated the utility of signal detection methods to analyze between group differences in sensitivity to pain expressions and response bias.

Hypotheses

Based on the literature reviewed, the following hypotheses and issues were examined.

1. If decoding pain expression is analogous to discriminating among different facial expressions, then decoding ability should increase with increasing age. However, an age effect might not be the same for sensitivity and response bias in judging pain expressions.

2. Participants scoring higher on social acceptance and empathy will be more accurate than those scoring lower in social acceptance.

3. Decoding studies do not reliably demonstrate differences between males and females. However, where there are differences, females are found to be more accurate than males. Therefore, if there are differences between males and females in this study, the latter will be more sensitive.

4. Finally, there will be a comparison among age groups of the most influential facial cues used when judging pain. That is, do children and adults rely on the same facial cues when they rate pain? This question regarding the influence of specific facial actions for participants of different ages is exploratory rather than one based on previous research. Therefore, there are no a priori hypotheses regarding the relative importance of facial cues for participants of different ages.

METHOD

Participants

A total of 134 children and adults from four age groups participated in this study. The three age groups of children were five- and six-year-olds (n = 33), eight- and nine- year-olds (n = 35), and eleven- and twelve-year-olds (n = 33). For convenience, these will be referred to as young, middle and older children, respectively. The number of boys and girls in each of the groups were as follows: young boys (n = 16), young girls (n = 17), middle boys (n = 17), middle girls (n = 18), older boys (n = 16) and older girls (n = 17). The group of young adults had 17 men and 16 women.

The children were solicited from a variety of sources; some were selected from an available school, some from summer day camps, others were by referral and the remainder from personal contacts. Therefore, it was a sample of convenience. The young adults were university students recruited from introductory psychology classes at UNBC, and they received course credit for their participation.

Materials

Participants viewed videotaped segments of the faces of patients undergoing an assessment of a shoulder injury by a physiotherapist at a sport medicine clinic. The assessment included ten range of motion and two accessory tests. The range of motion tests included active and passive abduction, flexion and internal and external rotation movements as well as two additional passive movements (strong and weak "lock and quadrant" tests). For the active tests, the patients moved their arms into position. In contrast, the patients had their arms moved by the therapist during the passive and quadrant tests. See Prkachin and Mercer (1989) for a more detailed description of the tests. The patients were standing, facing the camera, when the four active tests were performed. On the videotape, the patients' heads and shoulders were visible. For the remaining tests, the patients were lying on an examining table. The camera angle was from above and slightly to the side of the face. The shoulders were not usually visible for these tests.

The stimuli were carefully selected from among the entire series collected by Prkachin and Mercer (1989) to have certain properties. The intent was to construct a test tape displaying a series of facial expressions that could be specified, by independent criteria, as conveying no pain, some pain and a lot of pain. To do this, excerpts were selected to meet each category based on the measures of pain expression available from the studies of Prkachin and Mercer (1989) and Prkachin, Berzins and Mercer (1996). Each excerpt had been coded for the intensity (on a 5-point scale) and duration of four specific movements which have been shown to covary with pain: 1) brow lowering, 2) orbit tightening, 3) levator activity and 4) eye closure. A summary score, consisting of the sum of the products of the intensity and duration of each action was available. This score varied between 0 and 120. A videotape was constructed that included 30 excerpts which, according to the foregoing criteria, depicted no pain activity (a score of 0 on pain actions), minimal pain activity (a score ranging between 1.0 - 10.0), and moderate to strong pain (a score over 10). Therefore, there were a total of 90 target stimuli. In addition to having the pain index scores for the facial actions, the actual pain ratings by the patients were available for comparison purposes.

The stimuli were presented in a fixed randomized order varying the sex of the patient and the pain intensity of the facial expression. These target stimuli were presented for two seconds followed by a five second interstimulus interval. The trial number appeared on the screen before the presentation of the face. The first 30 seconds of the tape were blank, then there were 12 practice trials. These trials were followed by another 30 seconds of blank tape prior to the presentation of the 90 target trials. The total time of the tape was less than 13 minutes.

All participants viewed the tape on the same portable Genexxa TV/video machine. The participants were instructed to sit at a comfortable distance from the screen. In almost all cases the distance chosen was approximately 1 meter. One six-year-old girl, who said that she wanted to get up very close to see the faces better, chose to sit closer. One young man said that he moved back so that he could see the whole image on the screen better.

Participants also completed two questionnaires. The first was an empathy questionnaire and the second was a self-perception scale. There were a number of criteria employed in the selection of the instruments used for the assessment of empathy and social skill. The first was that a similar instrument be available for the wide range of ages observed in this project (Bryant, 1982). That is, it was deemed desirable to have parallel forms for adults and children to make across-age comparisons. In addition, the questionnaires should have acceptable psychometric properties. Thirdly, either the adult or children's questionnaire should have a demonstrated relationship to facial expression decoding ability in previous research.

In the case of empathy, domains of the IRI have been found to be related to decoding accuracy of facial expression of pain (Solomon, 1995). The IRI for adults also has excellent psychometric properties (Davis, 1983). Versions of the children's IRI have been employed in a variety of studies by Fabes and Eisenberg and their colleagues (e.g. Eisenberg, et al., 1989). Furthermore, the IRI measures different dimensions of empathy. It is probable that these dimensions are not equally associated with pain judgments. Therefore, the adults' and children's versions of the IRI were selected to assess empathy.

As indicated previously, social competence is operationalized in many different ways. Therefore, selecting a suitable instrument was somewhat more difficult. Social competence has been assessed with the CBCL (Achenbach & Edelbrock, 1983) in facial expression decoding research by Feldman and his colleagues (e.g., Feldman & Custrini, 1989). The CBCL does not have a parallel form for adults. Therefore, an instrument that has been found to be correlated with the CBCL or has been found to be related to other measures of social skill was sought. Harter's Self-Perception Profiles for children (SPPC) and college students (SPPCS), which measure a variety of dimensions of self-concept such as scholastic competence and global self-worth, were selected for the following reasons.

Despite the fact that none of the profiles have been employed in facial expression research, there is sufficient justification for the use of them to assess aspects of social competence via the social acceptance domain. Pope and Ward (1997) administered the CBCL and SPPC to children with craniofacial anomalies. Social competence was defined as the composite of the social acceptance score and the CBCL social competence scale score. Low perceived social acceptance on the SPPC is observed in more depressed than non-depressed students (e.g., Heath, & Weiner, 1996) and depressed children are more likely to be rated as lower in social competence as measured by the CBCL (Renouf, Kovacs, & Mukerji, 1997)

The social acceptance domain of the SPPC has been found to be related to other aspects of social behaviour as well. For example, both Boivin and Hymel (1997) and Austin and Joseph (1996) have found victimization to be associated with lower levels of social acceptance. Withdrawal and loneliness in the former study and bullying in the latter study, were also associated with low perceived acceptance.

Hymel, Bowker, and Woody (1993) administered a different self-concept instrument (Self-Description Questionnaire; Marsh, Smith, & Barnes, 1983; cited in Hymel, Bowker, and Woody, 1993) and compared the results with peer ratings. Overall, they claimed that average children, as defined by sociometric nominations, were fairly accurate in their self-perceptions. Therefore, the peer relations domain of self-concept corresponded to peer rated social competence.

Therefore, to assess empathy, the Interpersonal Reactivity Index (IRI; Davis, 1983) was administered to adults and the children's version of the IRI developed by Fabes et al. (1992) was completed by children. Harter's Self-Perception Profile for Children (SPPC; Harter, 1984) and Harter's Self-Perception Profile for College Students (SPPCS; Neeman & Harter, 1986) were administered with children and adults respectively. The children's IRI (Fabes, et al., 1992) and the SPPC (Harter, 1984) had been standardized with children in grade three by their respective developers. See Appendix B for the questionnaires.

The SPPC has 5 dimensions of self-concept (physical appearance, athletic competence, behavioural conduct, scholastic competence and social acceptance) as well as a measure of global self-worth. In the first version of the scale, the social scale was referred to as social competence. Harter (1984) has since reported that it measures a child's perception of his or her acceptance by peers. The SPPCS is composed of 12 dimensions plus the global self-worth measure. The categories on the children's version are broad whereas for the adults, categories are more refined. Therefore, the dimensions are similar but not exactly the same.

In contrast, the IRI scales are more directly comparable across ages. The adult version has four subscales, with 7 items per scale. These are: Empathic concern (EC), perspective taking (PT), personal distress (PD) and fantasy (F). For the children's IRI, the F scale was omitted and there were fewer items per scale. The original scale contained 4 items per scale resulting in a 12-item test. Fabes (1997) has recommended the use of a 10-item scale. The shorter scale is obtained by omitting two of the three negatively worded items (#4 & #8). The final version has 3 items for both EC and PD whereas PT still has 4 items. There is one negatively worded PT item remaining on the questionnaire. The children completed the 12 item version and analyses were conducted for both 12 and 10 item scales.

Procedure

Children. All of the children's sessions were conducted during the summer. Eighty children had their sessions in their own homes. For these children, consent was obtained at that time. The sessions for ten children were held in a familiar home - either a friend's, a relative's or a family daycare. In these cases, the parents agreed, in telephone conversations, to their children's involvement and allowed the adult present to provide written consent on the parents' behalf. The remainder of the children participated at their day camp. The day camps were located at two elementary schools. For these cases, a consent form signed by a parent had been returned prior to the experimental session.

The TV was set up in a convenient location in the home or in an area of the school free from distractions. Some demographic information was obtained from the parents and/or children. This information included the child's birthdate, parents' education and occupation, the child's school and programme (approximately one third of the children did not attend their neighbourhood schools), and whether the child or a primary caregiver experiences recurrent or chronic pain. For most of day camp children, information about parents' occupations and education were missing.

Parents were allowed to be present for the entire session if either they or their children preferred. Most parents chose to watch at least some of the tape with their children. They were informed that because one of the purposes of the study was to examine children's ratings of pain in others, they should avoid letting their child know how they would rate the patients' faces.

A brief description of the procedure was then given. The children were instructed to watch each face on the video, decide if the patient hurt during the movement, and then give their response. It was explained that there was no audio on the tape so they had to watch the faces closely to judge if the person was experiencing any pain with the movement. The five point rating scale was explained and shown (0= no pain, 1= maybe pain or can't tell, 2 = a little pain, 3 = moderate or a middle amount of pain and 4 = a lot of pain). A laminated sheet of paper with the rating scale was placed in front of the participants. Pilot testing indicated that children as young as 5½ years of age were able to use the 5 point scale when the researcher recorded the responses.

The practice trials were presented to assure that the participants watched and were able to rate the faces. For the one child who said he had difficulty with this task, a second opportunity to view the practice trials was given. The children viewed each target stimulus and then gave their response aloud. That is, pain ratings were recorded during the interstimulus interval by the researcher. The trial number on the screen was used to assure correspondence between the video and response sheet.

The two questionnaires were administered when the tape was finished. With one exception, the first questionnaire completed was the children's version of the IRI. The second questionnaire was Harter's SPPC. In all cases, the oldest children completed their questionnaires independently. Only rarely did an older child need assistance with a word. For all of the youngest children, the instructions and items were read for the children. They provided their answers aloud or, more commonly, by pointing to their choices. Most of the children in the middle group chose to record their answers themselves with the items being read to them. However, some preferred to do their questionnaires in private.

Adults. University students were contacted to arrange a mutually convenient time. When students arrived at the testing room they were given a description of the study as one comparing children and adults in their ratings of facial expressions of pain. Demographic information was then obtained. The procedure for the adults was the same as above with the following exception. After they viewed the tape and rated the pain on the 5-point scale, they watched the tape a second time rating the excerpts using the scale that was originally used by the patients in the study from which the excerpts were selected. The scale of affective pain intensity is a ratio-scaled verbal descriptor instrument to assess clinical pain (Heft, Gracely, Dubner & McGrath, 1980). See Appendix C for this rating scale. Students then completed the IRI (Davis, 1983), and Harter's SPPCS (Neeman & Harter, 1986).

RESULTS

Demographic Information

The age and sex composition for all of the groups is in Table 1. There were no significant differences in the ages of males and females. The majority of the children ($\underline{n}$ = 64) attended their neighbourhood schools. Thirty-three children attended alternate programmes provided by the school district; seven were registered in French immersion and 26 were enrolled in the Montessori programme. The remaining four children attended private Christian schools.

Table 1
Mean Ages, in years, for four age groups

Age Group	Males	Females	Total
Young Children	6.21	6.27	6.24
Middle Children	8.90	8.96	8.93
Older Children	11.84	11.95	11.90
Adults	21.11	23.11	22.08

To determine socioeconomic status (SES), the table by Blishen, Carroll, and Moore (1987) was used. This table is based on very precise job classification and income information. For the most part, the participants or their parents provided very general job descriptions or classifications. Therefore, the most general classification in a category was used for rating SES unless more precise employment or academic information had been given. Family SES was taken as the highest rated occupation in a family. For the large majority of university students, family SES was based on a parent's classification. Overall, this sample was above average (mean = 42.74; Blishen, et al., 1987) in SES (see Table 2). A 4 (age) X 2 (sex) ANOVA was computed to determine if there were age or sex differences in family SES. A significant main effect of age was found, $\underline{F}$ (3, 123) = 2.95, $\underline{p}$ = .04. The Tukey-HSD post-hoc analysis indicated

Table 2
Mean SES of participant's family

Age Group	SES
Young Children	57.37_{a}
Middle Children	54.66_{ab}
Older Children	55.55_{ab}
Adults	47.23_{b}

$_{a,\,b}$ numbers with the same subscripts are not significantly different at $\underline{p}$ = .05.

that the adults' family SES was significantly lower than that of the families of the young children (Table 2).

Participants' Missed Ratings

A 4 (age) X 2 (sex) ANOVA was calculated to determine the effects of age and sex on the number of stimuli participants missed. The mean number of stimuli missed decreased with increasing age, $\underline{F}$ (3, 126) = 13.39, $\underline{p}$ = .001, eta squared = .24. The Tukey HSD test indicated that only the youngest children differed from the adults in the number of stimuli not rated. The missing data were not uniformly distributed among the children in the young and middle groups. In all further analyses with participants' ratings, the missing values were replaced with the average rating for that stimulus.

Sensitivity and Bias

The stimuli presented to participants can be conceived as representing three intensities of a signal (none, weak, strong) presented in the context of ambiguity. The rating scale used by participants can be conceived not only as an indication of apparent pain intensity, but also as a confidence scale of increasing certainty (i.e., one is unlikely to use the "strong pain" category unless one is confident that some pain was displayed). Data obtained in this manner can be analyzed using the methods of signal detection theory (McNichol, 1972) to characterize a person's performance in two ways. A measure of the participant's sensitivity to variations in the intensity of the stimulus can be calculated. Various indices of sensitivity are available, depending on the assumptions made about the task; however, they are all based on evaluating the observer's ability to discriminate between categories of stimuli. When the reference category against which responses to higher intensity stimuli contains none of the stimulus in question the measure obtained is conventionally interpreted as representing the observer's ability to detect the higher stimulus. In this study, the reference category included faces not displaying pain-related facial actions. A second measure, theoretically independent of the first, can also be calculated. This measure represents the observer's "bias," "response criterion" or tendency to be liberal or conservative in making ratings. Again, various indices are available to estimate these parameters, depending on the nature of the task.

To calculate indices of sensitivity and response bias, trials on which the patient was not making any pain-related facial actions were considered "noise". Trials on which the pain expression index varied from 1.0 to 10 were considered mild signals, whereas, trials on which the pain index exceeded 10 were considered strong.

The measure of sensitivity employed in the present study was P(A) (McNicol, 1972). P(A), for "proportion of area," was selected because it is a nonparametric measure of sensitivity that is ideally suited to the analysis of signal detection experiments involving the use of a rating scale. It is an estimate of the area of a unit square lying underneath a Receiver-Operating-Characteristic (ROC) curve formed when "hit" probabilities are plotted on the ordinate and "false alarm" probabilities on the abscissa. For each rating scale category, a "hit" is considered to be the use of that category to describe a "signal", whereas a "false alarm" is the use of that category to describe "noise". For all rating scale categories, except the most lax (i.e., "no pain"), hit probabilities, when calculated in this manner, will exceed false alarm probabilities if the stimuli being judged are distinguishable. It is conventional to accumulate the hit and false-alarm probabilities from the most stringent (in this case "4") to the most lax (in this case "0") rating scale category so that in practice the hit and false alarm probabilities ultimately sum to 1.0. Each rating scale category then becomes a point on the ROC curve. The area under that curve can then be estimated by application of the trapezoidal rule, yielding the measure P(A). When calculated this way, P(A) can vary from 0 to 1.0. A value of 0.5 represents chance performance. This reflects either an inability to distinguish two categories of stimuli or no sensitivity.

Two sensitivity measures were calculated in this manner for each participant. The first estimated participants' sensitivity to expression of mild pain and was calculated by considering mild pain expression to be signal and no pain expression noise. This measure will be referred to as P(A)m or alternatively, the sensitivity to or detectability of mild pain. The second estimated sensitivity to the expression of strong pain and was calculated by considering strong pain expression to be signals. This measure will be referred to as P(A)s or, alternatively, the detectability of strong pain.

The measure of response bias, "B''_D", was based on a nonparametric technique recommended by Donaldson (1992), based on application of the following formula to the data at each rating scale category:

$$B''_D = [(1 - p(\text{hit})) \times (1 - p(\text{FA})) - (p(\text{hit}) \times p(\text{FA}))] \;/\; [(1 - p(\text{hit})) \times (1 - p(\text{FA})) + (p(\text{hit}) \times p(\text{FA}))]$$

Measures of B''_D at each rating category were then averaged to yield a single measure of bias.

P(A)m and P(A)s were entered as repeated measures into a 2 (sex) X 4 (age) X 2 (pain intensity) repeated measures ANOVA (see Table 3 for means). Results indicated that there was a significant within-subject pain intensity effect, F (1, 126) = 291.91, p = .001, eta squared = .70. That is, the expression of strong pain was easier to detect than that of mild pain. There was also a significant main effect of age, F (3, 126) = 85.22, p = .001, eta squared = .67, indicating a developmental increase in sensitivity to detecting pain. These age and intensity effects are shown in Figure 2. Post-hoc analyses using the Tukey-HSD technique indicated that the young, middle and older children were significantly different from each other on both P(A)m and P(A)s; however, the older children were not significantly different from the adults. There were no significant interactions nor was there a main effect of sex.

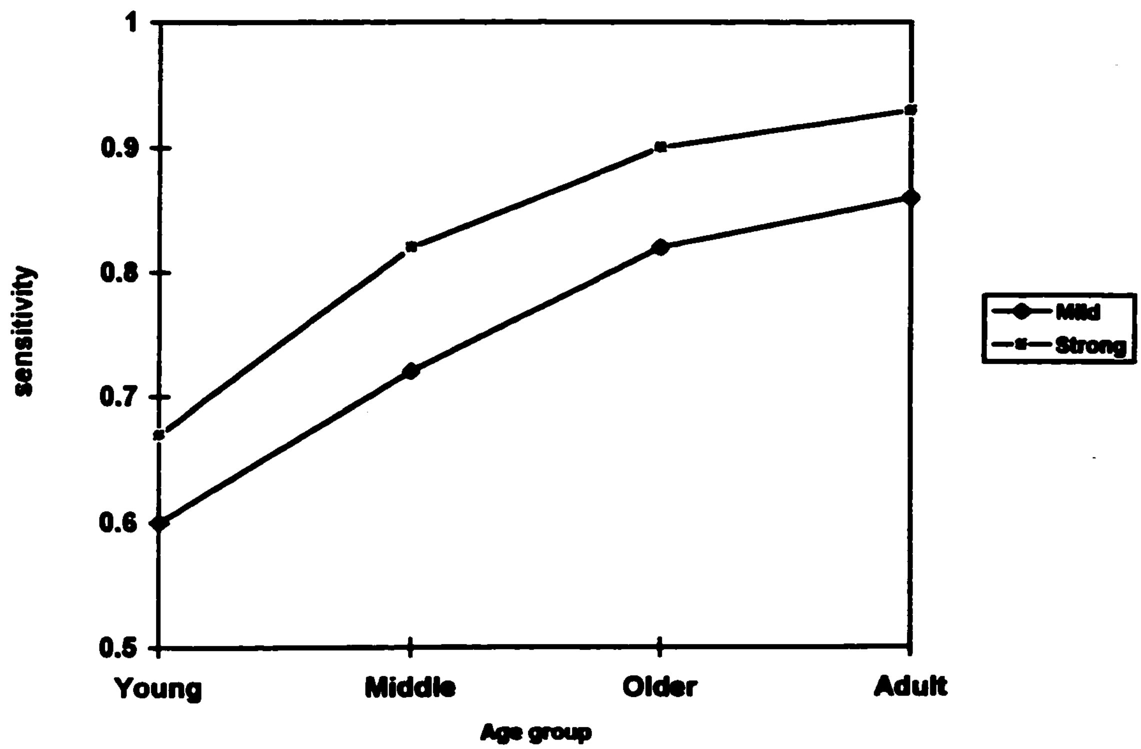
Table 3
Mean sensitivity and bias

Agegroup	N	Sensitivity		Bias	
		Mild	Strong	Mild	Strong
Young					
Male	16	0.58	0.65	0.39	0.34
Female	17	0.62	0.70	0.44	0.35
Middle					
Male	17	0.71	0.83	0.54	0.38
Female	18	0.74	0.82	0.48	0.37
Older					
Male	16	0.81	0.89	0.60	0.49
Female	17	0.84	0.90	0.46	0.35
Adult					
Male	17	0.87	0.93	0.43	0.31
Female	16	0.85	0.93	0.42	0.26

Bias reflects the tendency to be conservative or liberal in signal detection (Donaldson, 1992). Values above 0 indicate a conservative bias, whereas values below 0 reflect a tendency to be liberal. A person with conservative bias tends not to report a signal under conditions of uncertainty, whereas, as liberal bias indicates a greater tendency to report signals. As can be seen from Tables 3 and 4, the participants in this study were conservative in their ratings.

B''_D for both mild and strong pain intensities were also entered as repeated measures into a 2 (sex) X 4 (age) X 2 (pain intensity) repeated measures ANOVA. There were no significant between-subjects effects (age, sex

Figure 2. Sensitivity as a function of age



or age X sex interaction). However, there was a main effect of intensity on bias, F (1, 126) = 172.84, p = .001, eta squared = .58, as well as an age by intensity interaction, F (3, 126) = 3.02, p = .03, eta squared = .07. There was a greater degree of conservative bias for the mild intensity level. It also appears, from observation of the means (Table 4), that the middle and older children tended

Table 4
Mean bias scores for four age groups

Agegroup	N	Bias	
		Mild	Strong
Young Children	33	.42	.34
Middle Children	35	.51	.38
Old Children	33	.53	.42
Adults	33	.43	.29

to be more conservative, especially at the mild level. However, when separate 4 (age) X 2 (sex) ANOVAs were conducted for each intensity level, no significant effects resulted.

Recall that there was a main effect of age on sensitivity, and there was a significant difference in SES across age groups. In order to ascertain whether SES differences could account for the observed differences in sensitivity, correlations between them were computed for each group. As can be seen from Table 5, there were significant negative correlations between SES and sensitivity for both mild and strong pain for only the oldest children. That is, lower SES was associated with higher degrees of sensitivity to pain cues. In contrast, it was the adults who were from significantly lower SES families. Therefore, differences in SES across ages could not explain age- related differences in sensitivity. Nevertheless, other demographic differences may be able to account for differences in sensitivity. Specifically, the academic and occupational backgrounds of the participants' parents differed for the children and adults. The mothers and fathers of the children were more likely to have university degrees, 42% and 37%, respectively than the mothers (12%) and

Table 5
Correlations between SES and sensitivity and bias with level of significance (p) in parentheses

Agegroup	Sensitivity		Bias	
	Mild	Strong	Mild	Strong
Young Children	-.10	.10	-.02	-.05
	(.58)	(.59)	(.92)	(.77)
Middle Children	.05	-.01	.15	.23
	(.80)	(.99)	(.38)	(.20)
Older Children	-.36	-.42	-.03	.03
	(.04)	(.01)	(.89)	(.85)
Adults	.20	.24	.15	.19
	(.28)	(.20)	(.42)	(.31)

fathers (15%) of the adult students. In addition, almost half ($\underline{n}$ = 47) of the children's mothers and 15 of the fathers were employed in education, health, counselling or social work positions. In contrast, 8 mothers and 1 father of adults were employed in similar occupations. The percentage of parents with college diplomas or certification did not differ between groups for the mothers and father (between 27% and 30%). Therefore, occupational or educational differences between the parents may contribute to sensitivity.

As noted previously, bias and sensitivity are theoretically independent. Correlations among these measures were computed for the total sample (Table 6) as well as for the four age groups (Tables 7 through 10). In all cases, sensitivity to mild signals was highly correlated with sensitivity to strong cues ($\underline{r}$ ranged from .58 to .90, $\underline{p}$ = .001 for all correlations). The mild and strong bias intercorrelations were even greater ($\underline{r}$ = .91 to .98, $\underline{p}$ = .001). In contrast, the bias and sensitivity indices were not significantly correlated for any of the age groups. For the total sample, however, bias for the mild signal was significantly correlated with sensitivity to the strong signal ($r = .18$, $p = .04$). Overall, the above findings support the position that bias was empirically independent of sensitivity.

Table 6
Intercorrelations among sensitivity, bias and correlation measures for the total sample with significance level (p) in parentheses

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Sensitivity						
mild	1.00					
strong	.90 (.001)	1.00				
Bias						
mild	.11 (.21)	.18 (.04)	1.00			
strong	.09 (.33)	.060 (.50)	.95 (.001)	1.00		
Participant ratings						
self-report	.72 (.001)	.85 (.001)	.13 (.14)	.001 (.99)	1.00	
facial actions	.71 (.001)	.88 (.001)	.28 (.001)	.16 (.071)	.87 (.001)	1.00

Table 7

Intercorrelations among sensitivity, bias and correlation measures for young children with significance level (p) in parentheses

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Sensitivity						
mild	1.00					
strong	.74 (.001)	1.00				
Bias						
mild	.31 (.08)	.21 (.25)	1.00			
strong	.29 (.10)	.08 (.67)	.98 (.001)	1.00		
Participant ratings						
self-report	.46 (.01)	.79 (.001)	.15 (.42)	-.01 (.99)	1.00	
facial actions	.48 (.01)	.86 (.001)	.21 (.24)	.08 (.65)	.84 (.001)	1.00

Table 8
Intercorrelations among sensitivity, bias and correlation measures for middle children with significance level (p) in parentheses

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Sensitivity						
mild	1.00					
strong	.83 (.001)	1.00				
Bias						
mild	.07 (.70)	.28 (.11)	1.00			
strong	.05 (.76)	.15 (.41)	.95 (.001)	1.00		
Participant ratings						
self-report	.33 (.05)	.45 (.01)	.10 (.55)	.07 (.71)	1.00	
facial actions	.43 (.01)	.71 (.001)	.54 (.001)	.45 (.01)	.72 (.001)	1.00

Table 9
Intercorrelations among sensitivity, bias and correlation measures for old children with significance level (p) in parentheses

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Sensitivity						
mild	1.00					
strong	.58 (.001)	1.00				
Bias						
mild	-.16 (.37)	.20 (.28)	1.00			
strong	-.01 (.97)	.07 (.71)	.92 (.001)	1.00		
Participant ratings						
self-report	.24 (.19)	.51 (.002)	.21 (.25)	.05 (.68)	1.00	
facial actions	.23 (.19)	.76 (.001)	.47 (.01)	.33 (.06)	.500 (.01)	1.00

Table 10

Intercorrelations among sensitivity, bias and correlation measures for adults with significance level (p) in parentheses

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Sensitivity						
mild	1.00					
strong	.88 (.001)	1.00				
Bias						
mild	.04 (.82)	.18 (.33)	1.00			
strong	.19 (.29)	.18 (.32)	.91 (.001)	1.00		
Participant ratings						
self-report	.39 (.02)	.47 (.01)	-.12 (.51)	-.23 (.19).	1.00	
facial actions	.64 (.001)	.72 (.001)	.18 (.31)	.23 (.20)	.52 (.01)	1.00

Correlations between participant ratings and patient self report and facial actions

As another means of evaluating the sensitivity of observers to variations in patients' pain, the correlations between the participants' ratings and both the patients' self-report of pain (SR) and their facial actions (FACS) index were calculated. A separate correlation coefficient was computed over all 90 target trials for each participant. The means of these correlations for each group are presented in Table 11. A 4 (age) X 2 (sex) MANOVA was computed for the self-report and facial action correlations. Similar to the results for sensitivity, a main

Table 11
Mean correlations between participants' ratings and patient self-report and facial actions

Agegroup		N	Correlations with participant ratings	
			Patient Self Report	Patient Facial Actions
Young				
	Male	16	0.18	0.23
	Female	17	0.26	0.29
Middle				
	Male	17	0.40	0.44
	Female	18	0.41	0.44
Older				
	Male	16	0.46	0.50
	Female	17	0.45	0.48
Adult				
	Male	17	0.47	0.50
	Female	16	0.49	0.49

all $p < .05$
note: the correlation between patients' self-report and facial actions = 0.61.

effect of age was observed for both self-report F (3, 126) = 53.80, p = .001 and facial actions F (3, 126) = 37.38, p = .001. Tukey-HSD post-hoc analyses indicated that for the correlations between participant ratings and both SR and FACS, the youngest group of children differed from the other three groups (Table 12). The adults and older children also differed significantly from the middle children on SR correlations. Also note that the participants' ratings were more strongly correlated with the FACS scores than they were for the SR, t (133) = 14.78, p = .001.

Table 12
Mean correlations between participants' ratings and patients' self-report and facial actions

Agegroup	N	Correlations with participant ratings	
		Patient Self Report	Patient Facial Actions
Young Children	33	.22	.26
Middle Children	35	$.40_a$	$.44_a$
Older Children	33	$.45_{ab}$	$.49_a$
Adults	33	$.48_b$	$.50_a$

$_{a, b,}$ numbers with the same subscripts in a column are not significantly different at p = .05.

The SR and FACS correlations were then correlated with sensitivity and bias to ascertain the nature of relationship among all of these dependent measures. These computations were for the total sample and for the four age groups. For the most part, these correlations were significantly related to both sensitivity measures (Tables 6 through 10). Specifically, with one exception (i.e., older children for mild signals), both sensitivity measures were strongly correlated with the FACS correlations. Given that facial actions were rated, it is not surprising that the measure of sensitivity to facial actions is strongly associated. Similarly, the SR correlations were significantly related to sensitivity to both mild and strong pain cues for all participants with the exception of the older children rating the mild cues.

Bias was not significantly related to self-report for any groups (Tables 6 to 10). For facial action correlations, bias to mild signals was strongly and positively related for the middle and old children and the total sample. Only for the middle children was there a significant relationship for the strong signal. The older children and total sample correlations approached significance (p = .06 and p = .07, respectively). In all cases, the significant correlations between bias and participant correlations were positive; which indicates that as the association between participant ratings and patient scores increased so did the degree of conservativeness of the observer. Conversely, participants who were less conservative were found to have pain ratings that were less related to the patients' facial action scores.

Experience of and Exposure to pain

Participants were asked if they or a significant other experienced any chronic or recurrent pain problem. In the case of children, parents usually provided the answers. Neither experience of nor exposure to pain were significantly correlated with the other dependent measures computed. However, there was a trend for reporting chronic pain in a significant other to be related to the correlation between participants' and patients' ratings, $r = .15$, $p = .08$.

A 4 (age) X 2 (sex) MANOVA was computed with experience of and exposure to pain as the dependent variables. A significant main effect of age was observed for both experience $F(3, 120) = 2.81$, $p= .04$, eta squared = .07, and exposure to pain $F(3, 120) = 4.24$, $p = .01$, eta squared = .10. Post-hoc analyses indicated that adults had been exposed to more pain than either the young or middle children.

Self-perception and empathy analyses

All children in this study completed the full 12-item version of the IRI. Analyses were conducted for the scales with 12 items as well as with all of the negatively worded items removed. See Table 13 for the means for each age group. A 3 (age) X 2 (sex) MANOVA was computed for empathic concern (EC), perspective taking (PT), and personal distress (PD) with 3 and 4 items each. The analyses indicated that there were no significant age or sex differences on any of the scales. Because there were no systematic differences between ages or sexes for either 3 or 4 item scales, all further discussion of the children's IRI will be for the 10-item test.

The Self-perception scale data were entered into a 3 (age) X 2 (sex) MANOVA. There was a significant effect of sex on global self-worth, with boys

Table 13
Mean IRI subscale scores for children

	Males	Females	Total
EC (4 items)			
Young	11.50	11.65	11.58
Middle	12.53	12.33	12.43
Older	11.50	13.24	12.39
EC (3 items)*			
Young	10.31	10.18	10.24
Middle	10.82	10.50	10.65
Older	9.43	11.24	10.36
PT (4 items)*			
Young	10.00	11.29	10.67
Middle	11.94	11.06	11.49
Older	10.31	12.71	11.55
PT (3 items)			
Young	7.75	9.06	8.42
Middle	9.59	8.78	9.17
Older	7.88	9.71	8.81
PD (4 items)			
Young	10.88	11.88	11.39
Middle	11.88	11.33	11.60
Older	11.31	11.53	11.42
PD (3 items)*			
Young	7.75	8.47	8.12
Middle	8.64	8.33	8.49
Older	7.63	8.24	7.94

note: there are no significant agegroup, sex or agegroup X sex interactions for any of the scales
* scales that are retained on 10-item version

having higher scores (Table 14). Significant age differences were observed for physical appearance, athletic competence and global self-worth. In all cases, the youngest children had significantly higher scores than the oldest children; however, their scores were not significantly different from the scores of the middle group.

For physical appearance and GSW the middle children's scores did not differ significantly from either the older or younger group. Only for athletic competence did the oldest children have lower scores. Therefore, the young children did not respond in a manner that was systematically different from the middle group of children.

Table 14
Mean Self-perception subscale scores for children

		Males	Females	Total
Physical Appearance				
	Young	21.38	21.41	21.39_a
	Middle	21.23	19.61	20.40
	Older	19.68	18.05	8.85_b
Athletic Competence				
	Young	18.88	19.18	19.03_a
	Middle	20.47	17.28	18.83_a
	Older	16.50	15.88	16.18_b
Behavioural Conduct				
	Young	18.75	19.76	19.27
	Middle	17.76	19.78	18.80
	Older	18.06	18.59	18.33
Scholastic Competence				
	Young	18.88	19.12	19.00
	Middle	19.18	18.67	18.91
	Older	17.94	17.53	17.73
Social Acceptance				
	Young	19.25	19.47	19.36
	Middle	18.12	18.94	18.54
	Older	18.56	18.59	18.58
Global Self-Worth*				
	Young	22.56	20.71	21.61_a
	Middle	21.53	19.11	20.29
	Older	19.56	18.88	19.21_b

* boys had higher GSW, $\underline{F}$ (1, 95) = 5.569, $\underline{p}$ = .020.
$_{a,b}$ numbers with different subscripts, within a subscale, are significantly different at p = .05.

It is important to note that both of the children's questionnaires have been validated only on children as young as third grade. In this study, these questionnaires were administered to the youngest age group despite the lack of validation. Nevertheless, as noted above, there were no significant main effects of age on any of the subscales of the IRI and the six-year-old children did not respond differently from the nine-year-olds on the SPPC. In addition, the young children's scores were not different from the oldest children on three of the six subscales of the SPPC. Therefore, for the most part, the interview format employed for young children yielded comparable outcomes across different ages. Furthermore, the differences between the youngest and oldest children may not be attributable to the manner in which the instrument was administered but, may reflect differences in the children's self-concept.

The subscale scores of both questionnaires were correlated with all dependent measures for all participants. The results for adults and children are presented in Appendix D, Tables 1 to 8. In general, there were no systematic patterns across age groups. In fact, in many cases, even the direction of correlation differed between ages for the children. With the possible exception of PD, which was found to be related to the dependent measures, it is conceivable that the occasional significant effects are due to chance. Even in the case of PD, the results do not fall into a consistent pattern.

A comparison of two pain rating scales

The adults completed the rating task twice. The first time, they used the same five-point scale as the children used when they viewed the tape. The second time they used the affective rating scale used by the patients during the physiotherapy session. The scale of affective pain intensity is a ratio-scaled instrument; numerical values are associated with each of the verbal descriptors.

The data from the affective scale were dealt with in a number of different ways. First of all, an average difference or discrepancy score for each participant was calculated by subtracting the participants' rating on the affective scale from the patients' self-report value for each of the 90 stimuli. A single-sample t-test was conducted to determine whether the participants' ratings were different from the patients'. The results indicated that the patients' ratings were significantly higher than the participants', $\underline{t}$ (32) = 5.31, $\underline{p}$ = .001. Therefore, the adults in this study exhibited an underestimation bias, relative to the patients.

The difference score was then correlated with the bias and sensitivity indices. Bias for both stimulus intensities was strongly associated with the difference but sensitivity was not (Table 15). The SP and IRI subscale scores were also correlated with the difference between participant and patient ratings. As can be observed in Table 16, PD was positively correlated with discrepancy. That is, as personal distress scores increase, the greater is the degree of underestimation. Notice that this parallels the PD data for bias (Table 10). The results for the correlations between SP subscales and difference scores are also

Table 15
Correlations among the difference between participants' and patients' ratings and sensitivity and bias measures (p)

		participant - patient difference
Sensitivity		
	mild	.10 (.59)
	strong	.11 (.53)
Bias		
	mild	.54 (.001)
	strong	.49 (.01)

Table 16
Correlations among the difference between participants' and patients' ratings and IRI scales (p)

	participant - patient difference
Empathic Concern	-.29 (.10)
Perspective Taking	-.14 (.43)
Personal Distress	.42 (.02)
Fantasy	.12 (.51)

comparable to the findings for bias (Appendix D, Table 1). This consistency suggests that bias and discrepancy are measuring similar constructs.

The final way the affective scores were analyzed was to correlate them with the other five-point rating scale as well as with the patients' SR and FACS values. Thus, there were two measures of participants' pain ratings, a patient's self-report of pain and a measure of pain-related facial action. All of these measures were then intercorrelated. The mean values for these correlations are presented in Table 17. Single-sample t-tests were conducted on different pairs of means (Table 18). The following results will refer to comparison pairs as indicated on Table 18. The participants' two pain ratings were the most highly correlated measures (_r_ = .68). This correlation was significantly greater than any other correlation calculated (comparisons 1 and 2). Both of the participants' ratings scales were more highly correlated with patients' facial actions than with

Table 17
Mean correlations for self-report, facial action and 2 rating scales

	5-point scale	affective scale	patient's self-report	patient's facial actions
5-point scale	1.00			
affective scale	.68	1.00		
self-report	.48	.42	1.00	
facial actions	.50	.45	.61	1.00

Table 18
T-Tests to compare mean correlations for adults (mean r)

Comparison pair	t-value	p
1. 5-point & affective (.68) vs 5-point & SR (.48)	10.36	.001
2. 5-point & affective (.68) vs 5-point & FACS (.50)	-10.03	.001
3. 5-point & SR (.48) vs 5-point & FACS (.50)	1.66	.11
4. affective & SR (.42) vs affective & FACS (.45)	2.32	.03
5. 5-point & FACS (.50) vs SR & FACS (.61)	-15.06	.001
6. affective & FACS (.45) vs SR & FACS (.61)	7.76	.001
7. 5-point & SR (.48) vs affective & SR (.42)	3.73	.001
8. 5-point & FACS (.50) vs affective & FACS (.45)	1.95	.06

note: df = 32; FACS = patients' facial action score on pain index; SR = patients' self-report of pain; 5-point = participants' categorical rating scale; affective = participants' 17 point scale (i.e., the same scale as patients' SR) the patients' SR but was significant only for the affective scale (comparisons 3 and 4). That is, the ratings participants made, with either rating format, were more highly correlated with the FACSs than with the patients' subjective report. The correlation between the patients' facial actions and self-report of pain (r = .61) was greater than the FACS correlations with either the five-point (r = .50) or the affective scale (r = .45) (comparisons 5 and 6). The five-point rating scale was significantly more correlated with patient SR (r = .48) than was the affective scale (r = .42), t = 3.73, p = .001. The comparison of rating scales with FACS approached significance; the 5-point scale ratings were more related to the FACS score than were the affective scale ratings (comparison 8).

Action unit importance

The final, exploratory purpose of this study was to examine whether or not children of different ages and adults use the same facial cues to rate others' pain. The index used to categorize pain into three intensity levels included measurements of brow lowering (AU4), orbital contraction (AUorb = AU6 + AU7), levator contraction (AUlev = AU9 + AU10) and eye closing (AU43). This index did not include mouth opening. Previous research has indicated that mouth opening may occur during intense pain (LeResche, 1982) and may be a salient cue to pain intensity. Therefore, measurements of opening of the mouth were included in the following analyses.

To determine the relative importance of the different facial actions to pain ratings hierarchical regression analyses were computed. In all cases, the first step was to remove the variance accounted for by the participants and the patients. The second step was to enter the five facial actions together as predictors of participant ratings. The regression analyses (Tables 19 through 23) display only the second step of the analyses. Therefore, the R^2 reported are the R^2 change for the second step.

When all stimuli were combined, a pattern emerged. Note that brow lowering and mouth opening were most influential predictors of participants' ratings and that levator contraction was always negatively related to participants'

Table 19
Regression analysis for action units predicting participants' ratings for the total sample

Facial Action	B	SE B	β	t	p
brow lowering	.075	.003	.292	27.82	.001
orbital contraction	.014	.001	.100	10.08	.001
levator contraction	-.016	.004	-.041	-3.86	.001
mouth opening	.577	.020	.246	29.25	.001
eye closing	.014	.002	.047	4.83	.001

$R^2 = .24$, $F(5, 12053) = 760.83$, $p < .001$

Table 20
Regression analysis for action units predicting participants' ratings for the young children

Facial Action	B	SE B	β	t	p
brow lowering	.058	.006	.211	9.11	.001
orbital contraction	.001	.003	.003	.12	.91
levator contraction	-.006	.010	-.014	-.61	.54
mouth opening	.441	.047	.175	9.42	.001
eye closing	.006	.006	.022	1.06	.29

$R^2 = .10$, $F(5, 2962) = 65.45$, $p < .001$

Table 21
Regression analysis for action units predicting participants' ratings for the middle children.

Facial Action	B	SE B	β	t	p
brow lowering	.078	.005	.305	15.14	.001
orbital contraction	.016	.003	.115	6.03	.001
levator contraction	-.020	.008	-.051	-2.50	.01
mouth opening	.595	.038	.256	15.84	.001
eye closing	.016	.004	.066	3.54	.001

$R^2 = .27$, $F(5, 3142) = 235.20$, $p < .001$

Table 22
Regression analysis for action units predicting participants' ratings for the older children

Variable	B	SE B	β	t	p
brow lowering	.086	.005	.353	17.94	.001
orbital contraction	.018	.002	.141	7.62	.001
levator contraction	-.020	.008	-.051	-2.58	.01
mouth opening	.652	.035	.294	18.66	.001
eye closing	.007	.004	.029	1.59	.11

$R^2 = .34$, $F(5, 2962) = 310.00$, $p < .001$

Table 23
Regression analysis for action units predicting participants' ratings for the adults

Variable	B	SE B	β	t	p
brow lowering	.078	.005	.313	15.68	.001
orbital contraction	.020	.002	.152	8.08	.001
levator contraction	-.020	.008	-.050	-2.48	.01
mouth opening	.619	.036	.271	16.97	.001
eye closing	.017	.004	.071	3.87	.001

$R^2 = .32$, $F(5, 2962) = 282.62$, $p < .001$

ratings. For the total sample, adults and middle children, all action units significantly predicted ratings. The older children's ratings were predicted by all actions except eye closure. The young children's ratings were only predicted by brow lowering and mouth opening.

DISCUSSION

In the first part of this section, the results directly associated with the hypotheses will be summarized briefly and in the order they were initially presented. Then, there will be a discussion of the use of the term "accuracy". Accuracy can be interpreted in different ways depending on the type of data collected. A more detailed discussion of the results will follow. Finally, more general topics that may bear on the observations will be discussed.

Summary of results

The first hypothesis predicted that decoding ability would increase with age. In terms of discrimination, it was expected that sensitivity would improve with age. However, a specific effect of age on bias was not predicted. The results of this study did support the hypothesis that sensitivity would increase with age. Sensitivity increased throughout childhood, and the adults' sensitivity appeared to be higher than the oldest children, although the difference was not statistically significant. With respect to bias, overall, the participants in this study demonstrated a conservative bias. The effect of age on response bias, however, was less clear. Nine- and twelve-year-old children were more conservatively biased than the other two age groups at the mild pain intensity level.

It was hypothesized that participants scoring higher on social acceptance and empathy would be more accurate than those scoring lower on these variables. There were no systematic relationships between any of the personality measures and sensitivity or bias. However, the personal distress dimension of empathy was related to bias in the opposite direction for older children and adults.

Previous research on the effects of gender on decoding ability has been inconclusive. Studies have either found no gender effects (e.g., Tremblay et al., 1987) or they have found that females exhibit superior performance (Philippot & Feldman, 1990). In general, the results of this study were in the predicted direction, but none were significant.

Finally, there was a comparison among age groups of the most influential facial cues used when judging facial displays of pain. There were no specific predictions proposed regarding the relative importance of facial cues for participants of different ages. Brow lowering and opening the mouth were the most influential cues for all ages of participants. These were the only significant facial actions predictive of young children's pain ratings. Adults and middle children relied on all cues to rate pain. Eye closure was not a significant predictor for older children. With the exception of the youngest children, levator contraction was a significant negative predictor of pain ratings.

Judgment accuracy

Typical decoding studies analyze data in terms of accuracy. In this study, accuracy of observers' ratings was of interest. Therefore, the first part of this section will deal with the different ways in which the term "accuracy" can be interpreted. Accuracy is defined in typical recognition experiments as the agreement (proportion or frequency) among observers or with the experimenter's judgment of what is the correct emotion for the given situation (e.g., Kirouac & Dore, 1985; Walden & Field, 1990). The definition of accuracy is comparable in some pain studies. For example, Galin and Thorn (1993) and Hadjisdtovropoulos et al. (1996) defined accuracy as the frequency of correct labeling of video excerpts according to categories (e.g., no pain, genuine, or dissimulated).

In this study, accuracy can be interpreted in three ways. In SDT, P(A) (i.e. discriminability or sensitivity) is a measure of accuracy in interpreting the presence of a signal (hit) relative to false alarms. However, accuracy can also be the degree of relationship (i.e., correlation) between observers' pain ratings and patients' ratings of pain experienced. Finally, accuracy could also be reflected in the discrepancy between patients' self-reports and observers' ratings. Given that the term 'accuracy' has three possible interpretations, the use of it is ambiguous. Therefore, in the subsequent sections, the three will be referred to by the specific labels employed in the results section: P(A) or sensitivity, patient-participant correlations and discrepancy between ratings.

Sensitivity

Possibly the most important and clearest finding was that sensitivity to facial expression of pain did increase with increasing age, with the adults approaching maximum sensitivity for the strong signal. This finding is consistent with literature on the recognition of facial expression of emotions reviewed in the first chapter (e.g., Gosselin, 1995; Kirouac & Dore, 1985). Note, however, that the experimental tasks are different. In the previous research, participants discriminated among different emotional expressions, whereas in this study, participants discriminated different intensities of the same type of expression.

The significant improvement in sensitivity with age during childhood could be due to a variety of factors. It is possible that with increasing age, children become more sensitive to pain cues simply as a function of having more experience with or exposure to others in pain. Although the results of this study indicated that adults were more likely than young children to report having a family member with a chronic or recurrent pain problem, there was no increasing tendency to report this with age during childhood. In all likelihood the question was not specific enough to determine whether exposure to pain is age-related.

A second possible reason for improved performance between six and twelve years of age may be due to an increased ability to discriminate the most important facial cues. Gosselin (1995) observed that between the ages of five and eight, children became better at making finer discriminations between different facial expressions. The ability to make finer discriminations with age is also consistent with the results of the regression analyses in this study which indicated that younger children did not use all of the cues available to them to make their ratings.

A related possibility is that pain decoding may parallel cognitive development. However, there were a few six-year-olds who were more sensitive than the least sensitive adults and old children. It is unlikely that those young children display other cognitive skills equivalent to children in grade six or introductory psychology students.

It has been suggested that the ability to discriminate facial expressions coincides with development of the frontal lobes. That is, decoding performance will continue to improve until the frontal lobes mature at approximately 14 years of age (Kolb et al., 1992). Therefore, the improvement throughout childhood would be consistent with this prediction.

The results of this study demonstrate that sensitivity to facial expression of pain improves during childhood. However, from this study, it is not possible to establish which of the above explanations is most viable nor is it possible to disentangle the relative contribution of experiential, cognitive and maturational variables to the development of sensitivity.

The observed lack of a significant difference in sensitivity between older children and adults has two viable explanations. First of all, it is possible that children reach adult levels of pain discrimination by late childhood. It should be noted here that adult performance was not perfect, but it was very high. As indicated previously, there is some debate about whether adult decoding ability is attained prior to adolescence or whether improvement continues beyond that age. Kolb et al. (1992) suggested that recognition performance should improve until it reaches the maximum level at approximately 14 years of age.

The second reason for no significant difference between the older children and adults concerns demographic differences between the groups.

SES differed among groups with only the youngest children being significantly different from the adults; but, SES was significantly correlated with sensitivity only for older children. Therefore, SES data did not parallel the differences observed in sensitivity. Furthermore, SES was not consistently related to any of the other dependent variables. Although SES, per se, could not account for the differences in sensitivity between adults and children, the adults' parents did differ from the children's parents in two major ways. First of all, the children had a higher proportion of their parents in health, education and counselling occupations. In addition, a higher percentage of the children's parents had obtained a university degree. Therefore, it is possible that specific educational or occupational differences could have contributed to relationships between age and sensitivity. That is, the adults might have been significantly different in sensitivity if they had similar family backgrounds. In a decoding study with adults, Kirouac and Dore (1985) observed that significant education interactions reflected a "complex interplay" of factors that influence judgment accuracy.

Bias

Overall, participants' ratings were biased in the conservative direction. A conservative bias indicates that participants were unlikely to report pain unless they were sure that there was some degree of pain present. On the other hand, a liberal bias would have indicated that participants were inclined to report pain if uncertain. Given that it was a clinic situation, and the participants knew that all of the patients had injured shoulders, one might expect a report of pain when uncertain rather than the opposite, at least among adults.

There were no main effects of age or gender on bias. However, there was an interaction between stimulus intensity level and age. The middle and old children were more biased at the mild level. When the stimuli were mild, these children were less likely to report that the patient was experiencing pain. It is difficult to speculate about why children of these ages would be different from younger children and adults.

The adults underestimated pain when using the same affective scale as the patients. This observers' underestimation bias of pain has been reported in the past (e.g., Prkachin, Berzins & Mercer, 1994). Although this has been referred to as an underestimation bias in the literature, it is not bias in the signal detection tradition. An underestimation bias reflects a difference between raters' judgments, whereas in SDT, bias requires the calculation of hit and false alarm probabilities.

The bias measures, based on the 5-point categorical scale, were most correlated with the discrepancy between patients' self-report and the participants' ratings on the same affective scale. Although the adults underestimated the patients' pain, it was the discrepancy in ratings that was most correlated with response bias. Although bias, as computed in this study is different than bias when evaluated as a measure of discrepancy, there appears to be some similarity in the concepts. For example, a comparable pattern of results occurred for the correlation between the personality measures and discrepancy or bias. In fact, depending on whether or not a researcher is interested in accuracy of ratings or response tendency in the signal detection tradition, either method of reporting bias could be appropriate.

Personality variables

The focus of this study was on the perception of pain in others; personality variables were considered supplementary or ancillary. Nevertheless, as demonstrated in the Introduction, there is evidence in the literature that justifies the expectation that measures of personality or personal characteristics would be related to nonverbal decoding ability.

To assess social competence and empathy with comparable measures for such disparate age groups is difficult. Other techniques (e.g. sociometrics, role playing) were simply unavailable due to financial, time and seasonal constraints. The additional time necessary to collect such social skill data was not warranted given the secondary nature of this information. Furthermore, most sessions were conducted in children's homes during summer vacation. Therefore, in order to keep the intrusion to a minimum, two fairly brief questionnaires were administered.

With few exceptions, the personality variables examined in this study were not significantly related to the participants' perception of facial expression of pain. The lack of interrelationships is especially conspicuous for the sensitivity measures and for the correlations between participant ratings and patient SR and FACS which were most directly associated with age. It is interesting that the few significant effects involving personality measures tended to be associated with bias. The foregoing results taken together suggest that bias and sensitivity reflect different aspects of stimulus judgments. This independence of bias and sensitivity will be discussed further in the signal detection section below.

The observation that few self-perception or empathy domains were systematically associated with the dependent measures warrants discussion. There were few domain scores that were related to dependent variables across different age groups, especially among children. Even by excluding data from the young children (given that it is debatable that their questionnaire data be included), there were not many consistencies in findings involving self-reported personal characteristics between the other two groups. This general lack of consistency does not lend itself to easy interpretation. It is, of course, possible that children of these ages are very different from each other. But, other explanations of the general lack of patterned responses exist. Perhaps the children's sample was so homogeneous that there was little variability in SPPC and IRI scores. It is also possible that the small number of significant correlations is due to chance.

Empathy. For the adults in this study, only one dimension of empathy was correlated significantly with only one dependent measure. Specifically, PD was positively correlated with bias. The more empathic distress an adult reported, the less likely the observer was to report pain (i.e., conservative bias). The results differed for each age group of children. The PD scores for youngest children were positively associated with the correlations between their pain ratings and the patients'. The oldest children exhibited the greatest degree of relationship between empathy domains and bias: All of the empathy scores were negatively correlated with bias. The more empathy reported, the less conservative was the response tendency. Given that the oldest children and adults were most similar on the dependent variables, it is curious that the PD relationship with bias is in the opposite direction for the oldest children and adults. Is it possible that adults with high PD scores have developed a defensive reaction to another person in pain?

In contrast to other studies of empathy (e.g., Bryant, 1982), there were no significant age or gender differences on any of the dimensions assessed by the IRI. The IRI is a multidimensional instrument that measures a general self-reported response style when others are upset or need help. The affective and cognitive responses may be different as well. That is, one can understand what emotion another person may be experiencing without vicariously experiencing it oneself. The use of emotion labels is a cognitive skill that is highly dependent on verbal ability. Furthermore, the ability to vicariously experience another's state may be independent of such a cognitive skill, as suggested by Davis (1983). Therefore, the fact that empathy is usually operationalized by way of verbal methods may be a confound. Many researchers have indicated that the results of a paper and pencil test may not be accurate predictors of behaviour that is presumably dependent on the construct assessed. Consequently, there is a need for convergent validity of SR data. Eisenberg and her colleagues (e.g., Fabes, Eisenberg & Eisenbud, 1993) have begun to employ physiological measures (e.g., heart rate) to assess the correspondence between bodily responses to described or viewed situations and empathy/sympathy questionnaire results. According to Bryant (1982), a more thorough understanding of empathy will be gained by employing multiple measures: questionnaire data, physiological responses, facial expressions and behavioural responses such as helping.

Social Competence. In this study, social acceptance was not associated with any dependent measure of decoding performance. There are many possible explanations for this finding. The first and simplest is that social acceptance, as measured by the Self-Perception Profiles, may not be related to sensitivity to pain expressions.

As described in the introduction, Feldman and his colleagues found that social competence, as measured by the CBCL, was related to decoding ability. Social competence may be related to decoding accuracy in typical tasks but this task was not typical nor were the dependent measures. For example, P(A) is considered to be a measure of accuracy but is not equivalent to the proportion of correctly identified faces in an emotion decoding study. Perhaps social competence is related to discriminating among universal facial expressions when choosing among them, but may not be related to discriminating differing intensities of the same emotional expression.

Social competence is assessed in a variety of ways (Berndt & Burgy, 1995) that depend on different assumptions about it: sociometrics (e.g., Walden & Field, 1990; Rosenblum & Olson, 1997), ratings by peers (Parke et al., 1997), teachers (e.g., Rosenblum & Olson, 1997) or parents; and behavioural observation (e.g., Rosenblum & Olson, 1997). Obviously, social competence as a social skill construct is either poorly or inconsistently defined.

The SPPC was chosen for this study because it has been observed that social acceptance, on the SPPC, is related to other indicators of social competence (e.g., sociometric status; Hymel, et al., 199). Furthermore, Pope and Ward (1997) included social acceptance as part of a composite measure of social competence.

In all likelihood social acceptance and social competence are related or overlap but not synonymous. Therefore, it may not have been appropriate to use social acceptance as an indicator of social competence.

Self-perception domains. The social acceptance dimension of the self-perception profiles was the reason the Self-Perception Profiles were selected. The complete profiles were administered. To administer only the items of interest could possibly have invalidated the questionnaire. Although the social acceptance scale was of primary interest, analyses were conducted for all domains. The self-perception domains were, for the most part, not significantly associated with sensitivity or participant-patient correlations. However, some of them were correlated with bias. Given that few correlations were significant and nothing in the literature indicated they would be associated with decoding, they will not be addressed further.

Gender

With one exception (i.e., GSW), there were no gender effects observed in this study. Although most of the findings were in the expected direction, none reached significance. This result is consistent with Missaghi-Lakshman and Whissell (1991) who found no differences between boys and girls in decoding and with Tremblay et al, (1987) who also reported no gender effects. But, it is in contrast to the findings of Shortt et al. (1994) who did observe effects of gender on recognition ability. Phillipot and Feldman (1989) had found an age by gender interaction with boys' decoding performance delayed relative to girls. Therefore, previous research could be considered inconclusive.

There are a few possible explanations for the lack of effect of gender. Males and females may not differ in their perception of facial expression of pain. It is also possible that the nature of the task in this study was sufficiently different that the results of other decoding research are not relevant. For example, in most other decoding research, stories or situations are read with characters described or videos displayed. The participants then give an emotion label or choose one from those provided. The signal detection method used here provides the type of expression and the participant indicates the intensity (from none to a lot) of one type of expression. Therefore, males and females may respond differently on the typical decoding task than on a signal detection task.

It is also possible that males and females may have responded differently in past research because of the effects of the experimental situation or materials. Gender effects in past research may be attributable to differences in reliance on or use of the experimental context itself. For example, there may be cues in the instructions, descriptions or videos used that differ in salience for boys and girls. If experimental information is more salient to one group relative to another, then the ratings of that information may differ. The fact that the context for all excerpts used in this study was the same (i.e., a clinic) may have eliminated any effects.

In other decoding research, the time to complete the task is usually longer than the 13 minutes required in this study. Girls may find decoding tasks more interesting and may be able to attend to these tasks for longer periods of time. As support for this possibility, only one woman and no girls commented that the task in this study was boring. A few of the youngest girls asked if they could see more faces. In contrast, quite a few boys, almost all in the 9-year-old group, said it was boring and had more trouble paying attention. Even though it may have been less interesting to some of the boys, they were able to attend to the whole tape. If this phenomenon existed in a more demanding or time-consuming study, it is possible poorer performance could be attributed to lack of interest or attention rather than poorer discrimination ability.

There are two related additional possible reasons for the lack of gender effects. The first is that the children's sample was biased in that a high percentage of the children's parents had university degrees. It is possible that more highly educated people socialize their children to be more androgynous. Secondly, parents now may be raising their children differently than when some of the earliest work was conducted. Consequently, boys' and girls' socialization experiences may be more similar.

<u>Signal Detection Theory Analyses for Pain Judgment</u>

Ellermeier (1997) indicates that signal detection methodology is well-suited to studies investigating differences in pain judgment. Using this approach, he reanalyzed data that had demonstrated differences between men and women in their ratings of pain. He hypothesized that the apparent differences were not due to a difference in sensation, rather, they reflected a difference in willingness to report pain (i.e., a response bias). Therefore, he suggested treating pain ratings (i.e., category ratings) as confidence ratings rather than as sensation magnitude. Two parameters could then be derived. The first would reflect sensory discrimination (i.e., sensitivity) and the other would reflect how the participants used the response categories (i.e., bias). His reanalysis confirmed that men and women were equally "sensitive" but there were sex differences in bias with different stimulus intensities. Kemperman et al.

(1997) also indicated that signal detection theory can be useful to distinguish sensory and psychological factors in the report of pain by different patient groups. If signal detection methods are suitable for the analysis of pain rating data for the person experiencing pain then, they should also be appropriate to analyze observers' ratings.

See, et al., (1997) evaluated various measures of bias to determine which is the best under certain conditions. Given their analysis criteria (e.g., independence of bias and sensitivity, effect of nonperceptual manipulations, etc.), they recommend that, for nonparametic data, B_D is the most effective bias measure. Furthermore, sensitivity and bias measures should be paired depending on their categorization as parametric or nonparametric. Therefore, A' or P(A) is the appropriate (i.e., nonparametric) measure of sensitivity to correspond to B_D. P(A) should be affected by signal salience, a perceptual manipulation whereas B_D should be more influenced by nonperceptual factors (e.g., personality variables).

In this study, mild and strong intensity pain expressions were the differences in the salience of the signal. The fact that there was a significant effect of intensity of facial expression, demonstrates that the faces were perceived as different.

Also according to signal detection theory, sensitivity should not be affected by nonperceptual manipulations or variables. It is logical, therefore, that personality variables were essentially unrelated to the perceptual measure but were more likely to be associated with response bias.

Methodological Issues

Selection of dependent measures. The above discussion demonstrates the suitability of signal detection methods for pain judgment studies. Nevertheless, additional or alternate measures would be appropriate depending on the purposes of a study. In this study, P(A) and participant-patient correlations appeared to reflect similar aspects of expression judgment. A significant effect of age was observed for both. Similarly, they had parallel patterns of relationships to the empathy and self-perception dimensions. As suggested previously, bias and discrepancy also appeared to reflect similar aspects of response tendency.

Selection of rating scale. From the results with adults, it is suggested that a 5-point scale be employed for most purposes. In comparison to the affective rating scale, the 5-point scale was more closely correlated with patients' self reported pain and somewhat more correlated with facial actions. Furthermore, the 5-point scale was highly correlated with the participants' ratings on the affective scale. On the other hand, the use of the affective scale would be more suitable if discrepancy scores were of primary interest.

Facial actions

Prkachin and Mercer (1989) created the pain index used in this study from a combination of 4 movements of the middle and upper face. Mouth activity was excluded from the index because it was not reliably associated with self-reports of pain. Other investigators have found that a variety of facial actions around the mouth may occur during pain. For example, LeResche (1982) very clearly found that mouth activity was an aspect of the pain expressions she coded. It is possible that the photographs employed in her study were taken in situations that elicited a more intense pain response or emotions or motivational states in addition to pain. Despite the fact that the patients with the shoulder injuries reported intolerable pain for some movements, it is unlikely that the therapist-induced pain would be as severe as childbirth or leg surgery without anesthesia. Therefore, indices of facial pain in clinic settings may not include all actions that could be present during more intense pain. In support of this position is the observation that exaggerated pain expressions also tend to include facial actions involving the mouth (Galin & Thorn, 1993)

In this study mouth opening and brow lowering were the facial cues most predictive of pain ratings by observers. In general, the youngest children used fewer facial actions than participants of other ages to make judgments of pain. This finding is analogous to Gosselin's (1995) results that indicated that older children made better use of all of the facial actions present in an emotional display. Younger children may either attend to fewer facial cues or use fewer cues reliably.

This study did not ask subjects to discriminate pain from other facial expressions. The prototypical pain expression shares some facial actions with other prototypical emotion expressions (e.g., brow lowering occurs with both pain and sadness[3]; Gosselin, 1995). Gosselin (1995) claimed that age-related differences in performance can be attributed to improvements in discriminating the critical action units associated with an emotion. Sullivan, Kilpatrick and MacDonald (1995) state that in the course of normal daily interactions, children are exposed to facial expressions that may include components of mixed emotions. That is, expressions often are not pure or do not match the prototype of an emotional expression. Together these lead to the suggestion that it would be informative to examine the ability to discriminate pain expressions from other facial expressions. From that, it would be possible to determine which facial action units, alone or in combination, were critical for making a judgment of pain.

Future Directions

Future directions will be addressed from three main perspectives: developmental and methodological topics, and the pain communication model of Prkachin and Craig (1995).

Developmental Issues. The present study demonstrated that the age of the observer was an important factor in accuracy of pain judgments. Accuracy here refers to both sensitivity to differing intensities of pain expression as well as the degree of relationship between participant and patient pain ratings.

It is recommended that a wider age range of observers be studied to further clarify the relationship between development and accuracy. The

[3] There were occasions, with the youngest children, that they provided an emotion label rather than indicate a degree of pain. One girl in particular said that most of the pain faces were sad but it didn't look like they hurt during the movement.

responses of adolescents or children younger than kindergarten could be investigated. For example, preschoolers' sensitivity could be assessed with a simpler format (e.g., Did it hurt? yes/no or no, a little, a lot). Although this study demonstrated an age-related change in sensitivity to pain expressions, it was only possible to speculate as to why this occurred. Was it due to more experience with pain or with a variety of other facial expressions, general cognitive development or more specific neurological development? In order to examine the relationship between the development of pain expression decoding ability and any of these variables, future research will need to be precise in specifying or measuring the predicted correlates.

Older adults may perform differently than university students. Development continues throughout the lifespan. Therefore, adults of different ages or those with different experiences may interpret pain displays in diverse ways.

The children's sample was fairly homogeneous. Therefore, a more heterogeneous sample of children may yield different results (i.e., more variable SES, parents' education or occupation). Alternatively, a sample could be selected on the basis of specific criteria. For example, children with high versus low scores on a variable of interest could be chosen. The variable of interest could be one that would be expected to be related to experiencing or being exposed to pain (e.g., sport participation or health status). It is also possible that social skills variables are related to pain expression decoding and different measures could be employed.

The oldest children and adults had PD and bias correlations that were in the opposite direction. An investigation of adolescents could further examine the relationship between these variables. The first suggestion is to examine adolescents' responses on both the adults' and children's versions of the IRI. Although the children's items were designed to parallel the adults', differences could be attributed to the items on the scale. It is conceivable that adolescents and adults do respond differently to others in distress. Adolescents may not have very well-established strategies for responding to distress. It is also possible that it is only the self-report responses that are different not the actual response to them. This would suggest that measures of convergent validity (e.g., physiological variables such as heart rate) would be useful to assess the relationship between empathic distress and physiological response when evaluating decoding ability.

Some of the action units present in pain expressions are also observed in other facial expressions of negative emotions. Camras, Sullivan and Michel (1993) indicated that the pain-distress expression is very similar to the anger expression. Similarly, components of pain are observed in sadness and disgust. Therefore, in contrast to the present study in which the intensity of one expression was discriminated, a future study could examine the discriminability of pain from other expressions. First of all, it was assumed when undertaking this study that findings from the general emotion expression literature applied to pain expressions. This assumption may have been faulty. Pain decoding may develop differently than it does for other expressions. Secondly, discriminating between expressions may be different than discriminating between different intensities of the same expression. By examining pain in comparison to other expressions, one would gain a clearer understanding of the critical action units observers use when making judgments.

Methodological Issues. The youngest children performed at significantly better than chance and the adults approached maximum sensitivity. Although pilot testing indicated that five year olds were able to use the 5-point rating scale it may have been too precise for them. Perhaps a 3- or 4-point scale would have yielded higher sensitivity scores. In addition, future research should employ a Faces scale, as described in the introduction, to determine if children could use it to rate others' pain or whether they would treat it as a matching task.

Future research could further investigate the relationships among the different dependent measures employed and determine under which conditions they would be most appropriate. Although discrepancy and B_D are significantly correlated, the latter is a better measure of response bias than discrepancy. However, response bias may be interpreted differently depending on the position of a point relative to the ROC curve. In some circumstances, the average difference between patients' and observers' scores may be the preferred measure. For two reasons, it is probably easier for most readers to understand.

It is a simple mathematical function (i.e., subtraction) and it is expressed in units of the ratio scale that directly correspond to the ratings made by both observers and patients.

Empathy could be investigated from another perspective, or in combination with the IRI. For example, a physiological measure such as skin conductance, heart rate or facial EMG could be employed during the presentation of excerpts (McHugo & Smith, 1996). The use of physiological measures would be especially important as a way to validate the PD scale because it is intended to address self-reported distress. The use of a physiological measure would permit an examination of whether or not the self report of distress is related to physiological arousal.

Although social acceptance was not related to any of the dependent variables, that does not mean that social competence or some other social skill is not. Facial expression encoding and decoding have been found to be related to a variety of social variables. Also, facial expressions are important for interpersonal communication. Furthermore, the expression of pain or distress could be a more potent cue to act for observers (e.g., to protect oneself from a threat, to help relieve another's suffering). Therefore, future research should employ a different measure of social skill (e.g., sociometric status) to better understand how social variables could influence the sensitivity to facial expressions of pain.

The pain communication model. This study focused primarily on the decoding (C) component of the model (Prkachin & Craig, 1995); therefore, the majority of comments below will be limited to this aspect. Research on this model is beginning to address which characteristics of the observer and of the person being observed influence pain judgments. As indicated above, the age of the observer was an important factor in determining sensitivity to differing intensities of pain expression. Observers' age was also associated with the correlations among participant ratings and patient self-report of pain and patient facial actions.

Matheson (1997) found that her participants were "predisposed" to see more pain in the faces of elderly patients than in younger patients. This study focused on the age of the observer rather than age of the patient. Future studies could examine differences in patients and observers of a wider age range. For example, children may rate the pain of other children differently than that of adults. Conversely, older adults, who may have witnessed or experienced a greater variety of painful situations may also differ in their responses. Their differences could be reflected in terms of sensitivity or response bias.

It is likely that people with diverse pain experiences decode pain expressions differently; which leads to the suggestion that participants for future research could be selected on the basis of their experience of or exposure to pain. The videotape for this study had patients with shoulder injuries. It is possible that an observer with a similar injury history may respond differently than one without such experience. This could extend to other painful conditions as well; such as arthritis, headaches, back pain, cancer.

Hadjistavropoulos et al. (1996) suggested that future research should address the possibility that "there may be some individual difference variable that would make some judges better decoders than others" (p.257). Although the results of the present study suggest that the PD domain of empathy may be associated with response bias independent of accuracy, there were no self-perception or empathy variables that were systematically related to sensitivity. Nevertheless, other individual difference variables (e.g., alexithymia, depression, anxiety, etc.) may be predictive of decoding performance.

Summary and Conclusions

In general, the results of this study demonstrate that the ability to recognize facial expression of pain, and more specifically, to discriminate different intensities of painful expression, improves with increasing age. Despite the fact that the method employed in this study to assess decoding performance differed from that typically used in facial expression of emotion research, the results are consistent with previous research concerning the effect of age on the ability to recognize different emotional expressions. Given that the SDT approach used in this study yielded parallel findings in the general facial expression literature, this suggests that SDT methods could be more widely employed in facial expression research.

The results of this study also highlight the importance of using a variety of dependent measures to evaluate observers' performance. Specifically, judgment accuracy can be assessed by different means, and therefore, can be interpreted differently. For example, although sensitivity, correspondence between raters and discrepancy can all be measures of accuracy, they afford different interpretations of observers' performance.

The videotape created for this study is ideal for examining characteristics of observers that influence pain ratings. The tape could be employed in a variety of settings with many different participant populations. In addition, it is fairly brief and yet, with 90 excerpts of three intensity levels, allows the collection of a wealth of data for analysis.

Finally, it should be apparent that the signal detection approach employed in this study is especially well-suited to the study of pain. As previously indicated in this paper, this method is useful to examine the independent effects of sensitivity (i.e., pain intensity) and bias (i.e., the willingness to report pain) for the person experiencing pain. The present study demonstrated the utility of this approach to evaluate observers' judgments. It is documented in the literature that patients and research participants are able to suppress or exaggerate pain expressions and also that observers may expect dissimulation (e.g., Poole & Craig, 1992). Therefore, the SDT approach would allow investigators to dissociate intensity and bias when examining any participant's pain ratings.

REFERENCES

Abu-Saad, H.H. (1994). Pain in children: Developing a programme of research. Disability and Rehabilitation, 16, 45-50.

Achenbach, T.M. & Edelbrock, C.S. (1982). Manual for the child behavior checklist and child behavior profile. Burlington, CT: Department of Child Psychiatry, University of Vermont.

Anand, K.J.S. & Craig, K.D. (1996). New perspectives on the definition of pain. Pain, 67, 3 - 6.

Austin, S., & Joseph, S. (1996). Assessment of bully/victim problems in 8 to 11 year-olds. British Journal of Educational Psychology 66, 447-456.

Beck, L. & Feldman, R.S. (1989). Enhancing children's decoding of facial expression. Journal of Nonverbal Behavior, 13, 269-278.

Berndt, T.J. & Burgy, L. (1996) Social self-concept. In B.A. Braken, et al. Handbook of self-concept: Developmental, social, and clinical considerations. New York, NY: John Wiley & Sons.

Bieri, D., Reeve, R.A., Champion, G.D., Addicoat, L., & Ziegler, J.B. (1990). The Faces Pain Scale for the self-assessment of the severity of pain experienced by children: Development, initial validation, and preliminary investigation for ratio scale properties. Pain, 41, 139-150.

Blishen, B.R., Carroll, W.K., & Moore, C. (1987). The 1981 socioeconomic index for occupations in Canada. Canadian Review of Sociology and Anthropology, 24, 465-488.

Boivin, M., & Hymel, S. (1997). Peer experiences and social self-perceptions: A sequential model. Developmental Psychology, 33, 135- 145.

Bryant, B.K. (1982). An index of empathy for children and adolescents. Child Development, 53, 413-425.

Bugental, D.E. Kaswan, J.W., & Love, L.R. (1970). Perception of contradictory meanings conveyed by verbal and nonverbal channels. Journal of Personality and Social Psychology, 16, 647- 655.

Camras, L.A., Sullivan, J., & Michel, G. (1993). Do infants express discrete emotions? Adult judgments of facial, vocal, and body actions. Journal of Nonverbal Behavior, 17, 171-186.

Chun, D.Y., Turner, J.A., & Romano, J.M. (1993). Children of chronic pain patients: Risk factors for maladjustment. Pain, 52, 311-317.

Craig, K.D., Hadjistavropoulos, H.D., Grunau, R.V.E. & Whitfield, M.F. (1994) A comparison of two measures of facial activity during pain in the newborn child. Journal of Pediatric Psychology, 19, 305-318.

Craig, K.D., Hyde, S.A. & Patrick, C.J. (1991). Genuine, suppressed, and faked facial behavior during exacerbation of chronic low back pain. Pain, 46, 161-172.

Custrini, R.J. & Feldman, R.S. (1989). Children's social competence and nonverbal encoding and decoding of emotions. Journal of Clinical Child Psychology, 18, 336-342.

Davis, M.H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. Journal of Personality and Social Psychology, 44, 113-126.

Donaldson, W. (1992). Measuring recognition memory. Journal of Experimental Psychology: General, 121, 275-277.

Egan, G. (1986). The skilled helper: A systematic approach to effective helping (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Eisenberg, N., Fabes, R.A., Murphy, B., Karbon, M., Smith, M., & Maszk, P. (1996). The relations of children's dispositional empathy-related responding to their emotionality, regulation, and social functioning. Developmental Psychology, 32, 195-209.

Eisenberg, N., Fabes, R.A., Schaller, M., Carlo, G., & Miller, P.A. (1991). The relations of parental characteristics and practices to children's vicarious emotional responding. Child Development, 62, 1393-1408.

Eisenberg, N., Fabes, R.A., Schaller, M. & Miller, P.A. (1989). Sympathy and personal distress: Development, gender differences, and interrelations of indexes. In N. Eisenberg (Ed.), Empathy and related emotional responses. San Francisco, CA: Jossey-Bass.

Ekman, P. (1977). Biological and cultural contribution to body and facial movement. In J. Blacking (Ed.), The anthropology of the body. London: Academic.

Ekman, P., & Friesen, W. (1978). Facial action coding system. Palo Alto, CA: Consulting Psychologists Press.

Ellermeier, W. (1997). On separating pain from the willingness to report it. Behavioral and Brain Sciences, 20, 448-449.

Fabes, R.A. (1997) Personal communication.

Fabes, R.A., Eisenberg, N., & Eisenbud, L. (1993). Behavior and physiological correlates of children's reactions to others in distress. Developmental Psychology, 29, 655-663.

Feldman, R.S., White, J.B., & Lobato, D. (1982). Social Skills and nonverbal behavior. In R.S. Feldman (Ed.), Development of nonverbal behavior in children. New York: Springer-Verlag.

Feldman, R.S., Philippot, P., & Custrini, R.J. (1991). Social competence and nonverbal behavior. In R.S. Feldman and B. Rime (Eds.). Fundamentals of nonverbal behavior. New York: Cambridge University Press.

Galin, K.E., & Thorn, B.E. (1993). Unmasking Pain: Detection of deception in facial expressions. Journal of Social and Clinical Psychology, 12, 182-197.

Gosselin, P. (1995). Le developpement de la reconnaissance des expression faciales des emotions chez l'enfant. Canadian Journal of Behavioural Science, 27, 107-119.

Grunau, R.V.E., Johnson, C.C., & Craig, K.D., (1990). Neonatal facial and cry responses to invasive and non-invasive procedures. Pain, 42, 295-305.

Hadjistavropoulos, H.D. Craig, K.D., Hadjistavropoulos, T., & Poole (1996). Subjective judgments of deception in pain expression: Accuracy and errors. Pain, 65, 251-258.

Hager, J.C., & Ekman, P. (1997). The assymetry of facial actions is inconsistent with models of hemispheric specialization. In P. Ekman & E.L. Rosenberg (Eds.). What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Codings System (FACS). New York: Oxford University Press.

Harter, S. (1985) Manual for the Self-Perception Profile for Children. Denver, CO: University of Denver.

Heath, N.L., & Wiener, J. (1996). Depression and nonacademic self-perceptions in children with and without learning disabilities. Learning Disability Quarterly, 19, 34-44.

Heft, M.W., Gracely, R.H., Dubner, R., & McGrath (1980). A validation model for verbal descriptor scaling of human clinical pain. Pain, 9, 363-373.

Hymel. S., Bowker, A., & Woody, E. (1993). Aggressive versus withdrawn unpopular children: Variations in peer, teacher and self-perceptions in multiple domains. Child Development, 64, 879- 896.

Johnston, C.C., Stevens, B.J., Craig, K.D., & Grunau, R.V.E., (1993). Developmental changes in pain expression in premature, full-term, two- and four-month-old infants. Pain, 52, 201-208.

Johnston, C.C., Stevens, B.J., Yang, F., & Horton, L. (1995). Differential response to pain by very premature neonates. Pain, 61, 471-479.

Kemperman, I., Russ, M.J., Clark, W.C., Kakuma, T., Zanine, E., & Harrison, K. (1997). Psychiatry Research, 70, 175-183.

Kirouac, G. & Dore, F.Y. (1985). Accuracy of the judgment of facial expression of emotions as a function of sex and level of education. Journal of Nonverbal Behavior, 9, 3-7.

Kolb, B., Wilson, B., & Taylor, L. (1992). Developmental changes in the recognition and comprehension of facial expression: Implications for frontal lobe function. Brain and Cognition, 20, 74-84.

LeResche, L. (1982). Facial expression in pain: A study of candid photographs. Jounal of Nonverbal Behavior, 7, 46-56.

Levensen, (1994). Emotional Control: Variations and consequences. In P. Ekman & R. J. Davidson (Eds.). The nature of emotion: Fundamental questions. New York, NY: Oxford University Press.

Manne, S.L. Jacobsen, P.B., & Redd, W.H. (1992). Assessment of acute pediatric pain: Do child self-report, parent ratings, and nurse ratings measure the same phenomenon? Pain, 48, 45-52.

Matheson, D. H. (1997). The painful truth: Interpretation of facial expressions of pain in older adults. Journal of Nonverbal Behavior, 21, 223-238.

McHugo, G.J., & Smith, C.A. (1996). The power of faces: A review of John T. Lanzetta's research on facial expression and emotion. Motivation and Emotion, 20, 85-120.

Missaghi-Lakshman, M. & Whissell, C. (1991). Children's understanding of facial expression of emotion: II. Drawing of emotion-faces. Perceptual and Motor Skills, 72, 1228-1230.

Neeman, J. & Harter, S. (1986). Manual for the Self-Perception Profile for College Students. Denver, CO: University of Denver.

Nelson C.A., & de Haan, M. (1996). Neural correlates of infants' visual responsiveness to facial expression of emotion. Developmental Psychobiology, 29, 577-595.

Odom, R. D., & Lemond, C.M. (1972). Developmental differences in the perception and production of facial expressions. Child Development, 43, 359-369.

Parke, R.D., O'Neil, R. Spitzer, S., Isley, S. Welsh, M., Wang, S., Lee, J., Strand, C., Cupp, R. (1997). A longitudinal assessment of sociometric stability and the behavioral correlates of children's social acceptance. Merrill-Palmer Quarterly, 43, 635-662.

Patrick, C.J., Craig, K.D., & Prkachin, K.M. (1986). Observer judgments of acute pain: Facial action determinants. Journal of Personality and Social Psychology, 50, 1291-1298.

Philippot, P. & Feldman, R.S. (1990). Age and social competence in preschoolers' decoding of facial expression. British Journal of Social Psychology, 29, 43 - 54.

Philippot, P., Feldman, R.S., & McGee, G. (1992) Nonverbal behavioral skills in an educational context: Typical and atypical populations. In R.S. Feldman (Ed.), Applications of nonverbal behavioral theories and research. Hillsdale, NJ: Lawrence Erlbaum Associates.

Poole, G.D. & Craig, K.D. (1992). Judgments of genuine, suppressed, and faked facial expressions of pain. Journal of Personality and Social Psychology, 63, 797-805.

Pope, A.W., & Ward, J. (1997). Factors associated wiht peer social competence in preadolescents with craniofacial anomalies. Journal of Pediatric Psychology, 22, 455-469.

Profyt, L. & Whissell, C. (1991). Children's understanding of facial expression of emotion: I. Voluntary creation of emotion-faces. Perceptual and Motor Skills, 73, 199-202.

Prkachin, K.M. (1992a). The consistency of facial expressions of pain: A comparison across modalities. Pain, 51, 297-306.

Prkachin, K.M. (1992b). Dissociating spontaneous and deliberate expressions of pain: Signal detection analyses. Pain, 51, 297- 306.

Prkachin, K.M., Berzins, S., & Mercer, S.R. (1994). Encoding and decoding of pain expressions: a judgement study. Pain, 58, 253-259.

Prkachin, K.M. & Craig, K.D. (1985). Influencing nonverbal expressions of pain: Signal detection analyses. Pain, 21, 399-409.

Prkachin, K.M. & Craig, K. D. (1995). Expressing pain: The communication and interpretation of facial pain signals. Journal of Nonverbal Behavior, 19, 191-205.

Prkachin, K.M., Currie, N.A., & Craig, K.D. (1983). Judging nonverbal expressions of pain. Canadian Journal of Behaviour Science, 15, 409-421.

Prkachin, K.M. & Mercer, S.R. (1989). Pain expression in patients with shoulder pathology: Validity, properties and relationship to sickness impact. Pain, 39, 257-265.

Prkachin, K.M., Solomon, P.S., Hwang, T. & Mercer, S.R. (1995) Does experience affect judgements of pain behaviour? Evidence from relatives of pain patients and health-care providers. Unpublished manuscript.

Regan, D.T. & Totten, J. (1975). Empathy and attribution: Turning observers into actors. Journal of Personality and Social Psychology, 32, 850-856.

Renouf, A.G. , Kovacs, M., & Mukerji, P. (1997). Relationship of depressive, conduct, and comorbid disorders and social functioning in childhood. Journal of the American Adademy of Child and Adolescent Psychiatry, 36, 998-1004.

Riggio (1992). Social interaction skills and nonverbal behavior. In R.S. Feldman (Ed.), Applications of nonverbal behavioral theories and research. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rinn, W.E. (1991). The neuropsychology of facial expression. In R.S. Feldman & B. Rime, Fundamentals of Nonverbal Behavior. NY: Cambridge University Press.

Rosenblum, K.L., & Olson, S.L. (1997). Assessment of peer neglect in the preschool years: A short-term longitudinal study. Journal of Clinical Child Psychology, 26, 424-432.

Rosenthal, R. (1982) Conducting judgment studies. In K. Scherer and P. Ekman (Eds.), Handbook of methods in nonverbal behavior research. New York: Cambridge University Press.

Russell, J.A. (1995). Facial expressions of emotion: What lies beyond minimal universality? Psychological Bulletin, 118, 379-391.

Russell, J.A. (1997). Reading emotions from and into faces: Resurrecting a dimensional-contextual perspective. In J.A. Russell & J.M. Fernandez-Dols (Eds.), The psychology of facial expression. New York: Cambridge University Press.

See, J.E., Warm, J.S., Demaber, W.N., & Howe, S. R. (1997). Vigilance and signal detection theory: An empirical evaluation of five measures of response bias. Human Factors, 39, 14-29.

Shih, A.R. & von Baeyer, C.L. (1994) Preschool children's seriation of pain faces and happy faces in the affective facial scale. Psychological Reports, 74, 659-665.

Shortt, J.W., Bush, L.K., McCabe, J.L.R. Gottman, J.M., & Katz, L.F. (1994) Children's physiological responses while producing facial expressions of emotions. Merrill-Palmer Quarterly, 40, 40-59.

Shweder, R.A. (1994). "You're not sick, you're just in love": Emotion as an interpretative system. In P. Ekman & R. J. Davidson, The nature of emotion: Fundamental questions. New York, NY: Oxford University Press.

Singh, S.D., Ellis, C.R., Winton, A.S.W., Singh, N. N., Leung, J.P. Oswald, D.P. (1998). Recognition of facial expessions of emotion by children with attention-deficit hyperactivity disorder. Behavior Modification, 22, 128-142.

Solomon, P.E. (1995). Enhancing sensitivity to pain expressions. Unpublished Doctoral Dissertation, University of Waterloo.

Stevens, B.J., Johnston, C.C., & Horton, L. (1994). Factors that influence the behavioral pain responses of premature infants. Pain, 59, 101-109.

Sullivan, L.A., Kirkpatrick, S.W., & MacDonald. (1995). Interpretations of facial expressions of emotion by sexually abused and non-abused girls. Journal of Child Sexual Abuse, 4, 45-61.

Tremblay, C., Kirouac, G., & Dore, F. (1987). The recognition of adults' and children's facial expression of emotion. Journal of Psychology, 121, 341-350.

Tucker, J.S., & Riggio, R.E. (1988). The role of social skills in encoding posed and spontaneous facial expressions. Journal of Nonverbal Behavior, 12, 87-97.

Von Baeyer, C.L., Johnson, M.E., & McMillan, J.J. (1984). Consequences of nonverbal expression of pain: Patient distress and observer concern. Social Science & Medicine, 19, 1319-1324.

Walden. T.A., & Field, T.M. (1990). Preschool children's social competence and production and discrimination of affective expressions. British Journal of Developmental Psychology, 8, 65-76.

Walker-Andrews, A.S. & Lennon, E. (1991). Infants' discrimination of vocal expressions: Contributions of auditory and visual information. Infant Behavior and Development, 14, 131-142.

Wilkie, D.J., Holzemer, W.L., Tesler, M.D., Ward, J.A., Paul, S.M., & Savedra, M. C. (1990). Measuring pain quality: Validity and reliability of children's and adolescents' pain language. Pain, 41, 151-159.

Xeromeritou, A. (1992). The ability to encode facial and emotional expressions by educable mentally retarded and nonretarded children. The Journal of Psychology, 126, 571-584.

Zahn-Waxler,C., Raddke-Yarrow, M., & King, R.A. (1979). Childrearing and children's prosocial initiations toward victims of distress. Child Development, 50, 319-330.

Zaidel, S.F. & Mehrabian, A. (1969). The ability to communicate and infer positive and negative attitudes facially and vocally, Journal of Experimental Research in Personality, 3, 233-241.

Zimbardo, P.G. & Leippe, M. (1991). The psychology of attitude change and social influence. New York, NY: McGraw-Hill.

Zuckerman, M., Lipets, M.S., Koivumka, J.H., & Rosenthal, R. (1975). Encoding and decoding nonverbal cues of emotion. Journal of Personality and Social Psychology, 32, 1068-1076.

Appendix A

Pain-related Facial Action Units

Appendix A

Facial Action Units

Action Unit	Description	Muscular Basis
AU4	Brow Lowerer	Corrugator Procerus
AU6	Cheek Raiser	Orbicularis oculi, outer portion
AU7	Lid Raiser	Orbicularis oculi, inner portion
AU9	Nose Wrinkler	Levator labii superioris, alaque nasi
AU10	Upper Lip Raiser	Levator labii superioris
AU43	Eyes Close	Orbicularis oculi

Appendix B

Questionnaires

Children's Thoughts & Feelings Questionnaire (IRI)

The following statements are about your thoughts and feelings in a variety of situations. For each item, indicate how well it describes you by choosing the appropriate number on the scale. When you have decided on your answer, circle the number that best describes you. READ EACH ITEM CAREFULLY BEFORE RESPONDING. Answer as honestly and accurately as you can.

<u>ANSWER SCALE</u>

1	2	3	4	5
Not at all like me.	A little bit like me.	Kind of like me.	A lot like me.	Always like me.

1. I often feel sorry for people who don't have the things I have.	1	2	3	4	5
2. I get scared when I see an accident.	1	2	3	4	5
3. I have a hard time understanding why other people do the things they do.	1	2	3	4	5
4. Sometimes I don't feel very sorry for other people when they are having problems.	1	2	3	4	5
5. I sometimes feel like I don't know what to do when someone gets real upset.	1	2	3	4	5
6. When I see someone being picked on, I feel kind of sorry for them.	1	2	3	4	5
7. I sometimes try to understand my friends better by imagining how they think about things.	1	2	3	4	5
8. When I see someone get hurt, I usually remain calm.	1	2	3	4	5
9. When I see someone who is very sad or upset, I feel like I want to go somewhere else.	1	2	3	4	5
10. When I'm mad at someone, I don't waste much time trying to understand how they feel.	1	2	3	4	5
11. I often feel sorry for other children who are sad or in trouble.	1	2	3	4	5
12. I try to understand how other kids feel before I decide what to say to them.	1	2	3	4	5

Interpersonal Reactivity Index

The following statements inquire about your thoughts and feelings in a variety of situations. For each item, indicate how well it describes you by choosing the appropriate letter on the scale at the top of the page: A, B, C, D or E. When you have decided on your answer, fill in the letter on the answer sheet next to the item number. READ EACH ITEM CAREFULLY BEFORE RESPONDING. Answer as honestly as you can. Thank you.

ANSWER SCALE:

A	B	C	D	E
DOES NOT DESCRIBE ME VERY WELL				DESCRIBES ME VERY WELL

Answer	Item
____	1. I daydream and fantasize, with some regularity, about things that might happen to me.
____	2. I often have tender, concerned feelings for people less fortunate than me.
____	3. I sometimes find it difficult to see things from the "other guy's" point of view.
____	4. Sometimes I don't feel very sorry for other people when they are having problems.
____	5. I really get involved with the feelings of the characters in a novel.
____	6. In an emergency situations, I feel apprehensive and ill-at-ease.
____	7. I am usually objective when I watch a movie or play, and don't often get completely caught up in it.
____	8. I try to look at everybody's side of a disagreement before I make a decision.
____	9. When I see someone being taken advantage of, I feel kind of protective towards them.

____ 10. I sometimes feel helpless when I am in the middle of a very emotional situation.

____ 11. I sometimes try to understand my friends better by imagining how things look from their perspective.

____ 12. Becoming extremely involved in a good book or movie is somewhat rare for me.

____ 13. When I see someone get hurt I tend to remain calm.

____ 14. Other people's misfortunes do not usually disturb me a great deal.

____ 15. If I'm sure I'm right about something, I don't waste much time listening to other people's arguments.

____ 16. After seeing a play or movie, I have felt as though I were one of the characters.

____ 17. Being in a tense emotional situation scares me.

____ 18. When I see someone being treated unfairly, I sometimes don't feel very much pity for them.

____ 19. I am usually pretty effective in dealing with emergencies.

____ 20. I am often quite touched by things that I see happen.

____ 21. I believe that there are two sides to every question and try to look at them both.

____ 22. I would describe myself as a pretty soft-hearted person.

____ 23. When I watch a good movie, I can very easily put myself in the place of a leading character.

____ 24. I tend to lose control during emergencies.

____ 25. When I'm upset at someone, I usually try to "put myself in his shoes" for a while.

____ 26. When I am reading an interesting story or novel, I imagine how I would feel if the events in the story were happening to me.

____ 27. When I see someone who badly needs help in an emergency, I go to pieces.

____ 28. Before criticizing somebody, I try to imagine how I would feel if I were in their place.

What I am Like
(Self-Perception Profile for Children)

Name: ______________________ Age: ________ Boy or Girl

Sample Sentence

	Really True for me	Sort of True for me				Sort of True for me	Really True for me
a.	☐	☐	Some kids would rather play outdoors in their spare time	But	Other kids would rather watch T.V.	☐	☐
1.	☐	☐	Some kids feel that they are *very good* at their school work	But	Other kids worry about whether they can do the school work assigned to them	☐	☐
2.	☐	☐	Some kids find it *hard* to make friends	But	For other kids it's pretty *easy*.	☐	☐
3.	☐	☐	Some kids do very *well* at all kinds of sports	But	Others *don't* feel that they are very good when it comes to sports.	☐	☐
4.	☐	☐	Some kids are *happy* with the way they look	But	Other kids are *not* happy with the way they look.	☐	☐
5.	☐	☐	Some kids often do not like the way they *behave*	But	Other kids usually *like* the way they behave.	☐	☐
6.	☐	☐	Some kids often get *mad* at themselves	But	Other kids are pretty *pleased* with themselves.	☐	☐
7.	☐	☐	Some kids feel like they are *just as smart* as other kids their age	But	Other kids aren't so sure and *wonder* if they are as smart.	☐	☐
8.	☐	☐	Some kids have *a lot* of friends	But	Other kids *don't* have very many friends.	☐	☐

	Really True for me	Sort of True for me				Sort of True for me	Really True for me
9.	☐	☐	Some kids wish they could be a lot better at sports	But	Other kids feel they are good enough at sports.	☐	☐
10.	☐	☐	Some kids are *happy* with their height and weight	But	Other kids wish their height or weight were *different*.	☐	☐
11.	☐	☐	Some kids usually do the *right* thing	But	Other kids often *don't* do the right thing.	☐	☐
12.	☐	☐	Some kids *don't* like the way they are leading their life	But	Other kids *do* like the way they are leading their life.	☐	☐
13.	☐	☐	Some kids are pretty *slow* in finishing their school work	But	Other kids can do their school work *quickly*.	☐	☐
14.	☐	☐	Some kids are kind of *hard* to like	But	Other kids are really *easy* to like.	☐	☐
15.	☐	☐	Some kids think they could do *well* at just about any new outdoor activity they haven't tried before	But	Other kids are afraid they might *not* do well at outdoor things they haven't ever tried.	☐	☐
16.	☐	☐	Some kids wish their body was *different*	But	Other kids *like* their body the way it is.	☐	☐
17.	☐	☐	Some kids usually *act* the way they know they are *supposed* to	But	Other kids often *don't* act the way they are supposed to .	☐	☐
18.	☐	☐	Some kids are *happy* with themselves most of the time.	But	Other kids are often *not* happy with themselves.	☐	☐
19.	☐	☐	Some kids often *forget* what they learn	But	Other kids can remember things *easily*.	☐	☐

	Really True for me	Sort of True for me				Sort of True for me	Really True for me
20.	☐	☐	Some kids are always doing things with a *lot* of kids	But	Other kids usually do things by *themselves*.	☐	☐
21.	☐	☐	Some kids feel that they are *better* than others their age at sports	But	Other kids *don't* feel they can play as well.	☐	☐
22.	☐	☐	Some kids wish their physical appearance was *different*	But	Other kids *like* their physical appearance the way it is.	☐	☐
23.	☐	☐	Some kids usually get in *trouble* because of things they do	But	Other kids usually *don't* do things that get them in trouble	☐	☐
24.	☐	☐	Some kids *like* the kind of *person* they are	But	Other kids often wish they were someone else.	☐	☐
25.	☐	☐	Some kids do *very well* at their classwork	But	Other kids *don't* do very well at their classwork.	☐	☐
26.	☐	☐	Some kids wish that more kids liked them	But	Other kids feel that most kids *do* like them.	☐	☐
27.	☐	☐	In games and sports some kids usually *watch* instead of play	But	Other kids *play* rather than just watch.	☐	☐
28.	☐	☐	Some kids wish something about their face or hair looked *different*	But	Other kids *like* their face and hair the way the are.	☐	☐
29.	☐	☐	Some kids do things they know they *shouldn't* do	But	Other kids *hardly ever* do things they know they shouldn't do	☐	☐
30.	☐	☐	Some kids are very *happy* being the way they are	But	Other kids wish they were *different*.	☐	☐

	Really True for me	Sort of True for me				Sort of True for me	Really True for me
31.	☐	☐	Some kids have *trouble* figuring out the answers in school	But	Other kids almost *always* can figure out the answers.	☐	☐
32.	☐	☐	Some kids are *popular* with others their age	But	Other kids are *not* very popular.	☐	☐
33.	☐	☐	Some kids *don't* do well at new outdoor games	But	Other kids are good at new games right away.	☐	☐
34.	☐	☐	Some kids think that they are attractive or good looking	But	Other kids think that they are *not* very attractive or good looking.	☐	☐
35.	☐	☐	Some kids are usually very *kind* to others	But	Other kids *wish* they would be *kinder* to others.	☐	☐
36.	☐	☐	Some kids *aren't* very happy with the way they do a lot of things	But	Other kids think the way they do things is *fine*.	☐	☐

What I am Like

(Self-Perception Profile for College Students)

Name: ______________________________ Age: ________ Male or Female

The following are statements which allow students to describe themselves. There are no right or wrong answe since students differ markedly. Please read the entire sentence across. *First* decide which one of the two parts each statement *best* describes you; then go to that side of the statement and check whether that is just *sort of* tru for you or *really* true for you. You will just check ONE of the four boxes for each statement. Think about what yo are like in the university environment as you read and answer each one.

	Really True for me	Sort of True for me				Sort of True for me	Really True for me
1.	☐	☐	Some students like the kind of person they are	BUT	Other students wish that they were different.	☐	☐
2.	☐	☐	Some students are *not* proud of the work they do on their job	BUT	Other students are very proud of the work they do on their job.	☐	☐
3.	☐	☐	Some students feel confident that they are mastering their coursework	BUT	Other students do not feel so confident.	☐	☐
4.	☐	☐	Some students are not satisfied with their social skills	BUT	Other students think their social skills are just fine.	☐	☐
5.	☐	☐	Some students are not happy with the way they look	BUT	Other students are happy with the way they look.	☐	☐
6.	☐	☐	Some students like the way they act when they are around their parents	BUT	Other students wish they acted differently around their parents.	☐	☐
7.	☐	☐	Some students get kind of lonely because they don't really have a close friend to share things with	BUT	Other students don't usually get too lonely because they do have a close friend to share things with.	☐	☐
8.	☐	☐	Some students feel like they are just as smart or smarter than other students	BUT	Other students wonder if they are as smart.	☐	☐
9.	☐	☐	Some students often question the morality of their behaviour	BUT	Other students feel their behaviour is usually moral.	☐	☐
10.	☐	☐	Some students feel that people they like romantically will be attracted to them	BUT	Other students worry about whether people they like romantically will be attracted to them.	☐	☐

	Really True for me	Sort of True for me				Sort of True for me	Really True for me
11.	☐	☐	When some students do something sort of stupid that later appears very funny, they find it hard to laugh at themselves.	BUT	When other students do something sort of stupid that later appears very funny, they can easily laugh at themselves.	☐	☐
12.	☐	☐	Some students feel they are just as creative or even more so than other students	BUT	Other students wonder if they are as creative.	☐	☐
13.	☐	☐	Some students feel they could do well at just about any new athletic activity they haven't tried before	BUT	Other students are afraid they might not do well at athletic activities they haven't ever tried.	☐	☐
14.	☐	☐	Some students are often disappointed with themselves	BUT	Other students are usually quite pleased with themselves.	☐	☐
15.	☐	☐	Some students feel they are very good at their job	BUT	Other students worry about whether they can do their job.	☐	☐
16.	☐	☐	Some students do very well at their studies	BUT	Other students don't do very well at their studies.	☐	☐
17.	☐	☐	Some students find it hard to make new friends	BUT	Other students are able to make new friends easily.	☐	☐
18.	☐	☐	Some students are happy with their height and weight	BUT	Other students wish their height or weight was different.	☐	☐
19.	☐	☐	Some students find it hard to act naturally when they are around their parents	BUT	Other students find it easy to act naturally around their parents.	☐	☐
20.	☐	☐	Some students are able to make close friends they can really trust.	BUT	Other students find it hard to make close friends they can really trust.	☐	☐
21.	☐	☐	Some students do not feel they are very mentally able	BUT	Other students feel that they are very mentally able.	☐	☐
22.	☐	☐	Some students usually do what is morally right	BUT	Other students sometimes don't do what they know is morally right.	☐	☐
23.	☐	☐	Some students find it hard to establish romantic relationships	BUT	Other students don't have difficulty establishing romantic relationships.	☐	☐
24.	☐	☐	Some students don't mind being kidded by their friends	BUT	Other students are bothered when friends kid them.	☐	☐

	Really True for me	Sort of True for me				Sort of True for me	Really True for me
25.	☐	☐	Some students worry that they are not as creative or inventive as other people	BUT	Other students feel they are very creative and inventive.	☐	☐
26.	☐	☐	Some students don't feel they are very athletic	BUT	Other students do feel they are athletic.	☐	☐
27.	☐	☐	Some students usually like themselves as a person	BUT	Other students often don't like themselves as a person.	☐	☐
28.	☐	☐	Some students feel confident about their ability to do a new job	BUT	Other students worry about whether they can do a new job they haven't tried before.	☐	☐
29.	☐	☐	Some students have trouble figuring out homework assignments	BUT	Other students rarely have trouble with their homework assignments.	☐	☐
30.	☐	☐	Some students like the way they interact with other people	BUT	Other students wish their interactions with other people were different.	☐	☐
31.	☐	☐	Some students wish their body was different	BUT	Other students like their body the way it is.	☐	☐
32.	☐	☐	Some students feel comfortable being themselves around their parents	BUT	Other students have difficulty being themselves around their parents.	☐	☐
33.	☐	☐	Some students don't have a close friend they can share their personal thoughts and feelings with	BUT	Other students do have a friend who is close enough for them to share thoughts that are really personal	☐	☐
34.	☐	☐	Some students feel they are just as bright or brighter than most people	BUT	Other students wonder if they are as bright.	☐	☐
35.	☐	☐	Some students would like to be a better person morally	BUT	Other students think they are quite moral.	☐	☐
36.	☐	☐	Some students have the ability to develop romantic relationships	BUT	Other students do not find it easy to develop romantic relationships.	☐	☐
37.	☐	☐	Some students have a hard time laughing at the ridiculous or silly things they do	BUT	Other students find it easy to laugh at themselves.	☐	☐
38.	☐	☐	Some students do not feel that they are very inventive	BUT	Other students feel that they are very inventive.	☐	☐
39.	☐	☐	Some students feel they are better than others at sports	BUT	Other students don't feel they can play as well.	☐	☐

	Really True for me	Sort of True for me				Sort of True for me	Really True for me
40.	☐	☐	Some students really like the way they are leading their lives	BUT	Other students often don't like they like the way they are leading their lives.	☐	☐
41.	☐	☐	Some students are not satisfied with the way they do their job	BUT	Other students are quite satisfied with the way they do their job.	☐	☐
42.	☐	☐	Some students sometimes do not feel intellectually competent at their studies	BUT	Other students usually do feel intellectually competent at their studies.	☐	☐
43.	☐	☐	Some students feel that they are socially accepted by other people	BUT	Other students wish more people accepted them.	☐	☐
44.	☐	☐	Some students like their physical appearance the way it is	BUT	Other students do not like their physical appearance.	☐	☐
45.	☐	☐	Some students find that they are unable to get along with their parents	BUT	Other students get along with their parents quite well.	☐	☐
46.	☐	☐	Some students are able to make really close friends	BUT	Other students find it hard to make really close friends.	☐	☐
47.	☐	☐	Some students would really rather be different	BUT	Other students are very happy being the way they are.	☐	☐
48.	☐	☐	Some students question whether they are very intelligent	BUT	Other students feel they are intelligent.	☐	☐
49.	☐	☐	Some students live up to their own moral standards	BUT	Other students have trouble living up to their moral standards.	☐	☐
50.	☐	☐	Some students worry that when they like someone romantically, that person won't like them back	BUT	Other students feel that when they are romantically interested in someone, that person will like them back.	☐	☐
51.	☐	☐	Some students can really laugh at certain things they do	BUT	Other students have a hard time laughing at themselves.	☐	☐
52.	☐	☐	Some students feel they have a lot of original ideas	BUT	Other students question whether their ideas are very original.	☐	☐
53.	☐	☐	Some students don't do well at activities requiring physical skill	BUT	Other students are good at activities requiring physical skill.	☐	☐

54.	☐	☐	**Some students are often dissatisfied with themselves**	**BUT**	**Other students are usually satisfied with themselves.**	☐	☐

Appendix C

Pain Descriptors

Pain Descriptors

A - Undetectable

B - Detectable, not unpleasant

C - Slightly Unpleasant

D - Slightly Annoying

E - Annoying

F - Unpleasant

G - Slightly Distressing

H - Slightly Miserable

I - Very Annoying

J - Distressing

K - Very Unpleasant

L - Miserable

M - Very Distressing

N - Slightly Intolerable

O - Very Miserable

P - Intolerable

Q - Very Intolerable

Appendix D

Supplementary Tables

Table D1

Correlations among college students' self-perception subscales and sensitivity, bias and participants' correlations with patients' self-report and facial actions with level of significance (p) in parentheses

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Physical Appearance	.001 (.999)	-.026 (.89)	-.13 (.46)	-.11 (.53)	-.50 (.80)	.01 (.97)
Athletic Competence	.14 (.45)	.11 (.55)	-.34 (.05)	-.35 (.05)	.06 (.75)	-.10 (.58)
Creativity	-.09 (.60)	-.09 (.62)	-.11 (.53)	-.09 (.60)	.05 (.76)	.01 (.97)
Relationships with Friends	-.16 (.38)	-.08 (.67)	-.12 (.52)	-.24 (.18)	-.10 (.58)	-.19) (.29)
Humor	-.02 (.91)	.04 (.84)	-.06 (.76)	-.09 (.64)	-.26 (.15)	-.08 (.65)
Intellectual Competence	-.14 (.44)	-.15 (.39)	-.20 (.26)	-.25 (.16)	.07 (.72)	-.16 (.38)
Job Competence	.06 (.75)	.05 (.77)	-.42 (.02)	-.45 (.01)	.19 (.29)	.08 (.67)
Morality	-.09 (.63)	-.06 (.76)	-.22 (.21)	-.34 (.05)	.14 (.42)	-.07 (.68)
Behaviour with Parents	-.01 (.97)	-.09 (.61)	-.25 (.16)	-.28 (.11)	.01 (.77)	-.18 (.33)
Romantic Relationships	.18 (.32)	.13 (.47)	-.42 (.01)	-.38 (.03)	.11 (.54)	.02 (.91)
Scholastic Competence	-.01 (.97)	-.03 (.87)	-.27 (.13)	-.30 (.09)	.07 (.71)	.01 (.94)
Social Acceptance	.08 (.67)	.05 (.77)	-.24 (.18)	-.29 (.10)	-.02 (.92)	-.10 (.58)
Global Self-Worth	-.08 (.66)	-.07 (.69)	-.24 (.17)	-.32 (.07)	.06 (.76)	-.08 (.67)

Table D2
Correlations among adults' IRI subscales and sensitivity, bias and participants' correlations with patients' self-report and facial actions (p)

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Empathic Concern	-.24 (.18)	-.06 (.74)	-.03 (.86)	-.09 (.61)	-.10 (.57)	-.07 (.72)
Perspective Taking	-.32 (.07)	-.22 (.22)	-.13 (.48)	-.12 (.53)	-.10 (.58)	-.17 (.34)
Personal Distress	-.17 (.35)	-.03 (.85)	.38 (.03)	.34 (.06)	-.001 (.99)	.04 (.82)
Fantasy	-.03 (.87)	.001 (.999)	-.03 (.89)	.06 (.76)	-.12 (.51)	.05 (.80)

Table D3
Correlations among young children's self-perception subscales and sensitivity, bias and participants' correlations with patients' self-report and facial actions (p)

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Physical Appearance	.07 (.72)	-.03 (.88)	.38 (.03)	.42 (.01)	-.10 (.58)	.001 (.997)
Athletic Competence	.29 (.10)	.37 (.04)	-.19 (.30)	-.22 (.23)	.26 (.14)	.31 (.08)
Behavioural Conduct	-.04 (.85)	.18 (.32)	.36 (.04)	.30 (.09)	.33 (.06)	.31 (.08)
Scholastic Competence	.09 (.62)	.05 (.77)	.09 (.62)	.09 (.61)	.15 (.41)	.11 (.55)
Social Acceptance	-.10 (.57)	-.17 (.34)	-.10 (.58)	-.08 (.66)	-.12 (.51)	-.11 (.53)
Global Self-Worth	-.13 (.48)	-.19 (.30)	-.01 (.94)	.02 (.91)	-.07 (.69)	-.20 (.28)

Table D4
Correlations among young children's IRI subscales and sensitivity, bias and participants' correlations with patients' self-report and facial actions (p)

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Empathic Concern EC	.04 (.81)	.09 (.63)	.01 (.96)	-.01 (.98)	.03 (.86)	.01 (.96)
Perspective Taking PT	.12 (.51)	.07 (.69)	-.14 (.43)	-.16 (.38)	-.03 (.86)	.04 (.83)
Personal Distress PD	.08 (.65)	.25 (.16)	.30 (.09)	.22 (.23)	.34 (.05)	.35 (.05)

Table D5
Correlations among middle children's self-perception subscales and sensitivity, bias and participants' correlations with patients' self-report and facial actions (p)

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Physical Appearance	-.05 (.79)	-.11 (.54)	.03 (.86)	.04 (.82)	-.30 (.08)	-.22 (.20)
Athletic Competence	-.24 (.17)	-.21 (.23)	.06 (.72)	.03 (.87)	-.27 (.12)	-.23 (.18)
Behavioural Conduct	.18 (.32)	.16 (.37)	.23 (.18)	.21 (.22)	-.09 (.61)	.07 (.68)
Scholastic Competence	.18 (.31)	.05 (.79)	.001 (.998)	.03 (.86)	-.04 (.81)	-.01 (.98)
Social Acceptance	.04 (.83)	-.07 (.69)	-.13 (.45)	-.10 (.56)	-.06 (.98)	-.11 (.51)
Global Self-Worth	.01 (.94)	-.03 (.85)	.24 (.17)	.22 (.21)	-.32 (.06)	-.158 (.366)

Table D6

Correlations among middle children's IRI subscales and sensitivity, bias and participants' correlations with patients' self-report and facial actions (p)

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Empathic Concern EC	.06 (.75)	-.02 (.90)	-.25 (.14)	-.35 (.04)	-.05 (.79)	-.15 (.39)
Perspective Taking PT	.07 (.68)	.23 (.19)	.06 (.74)	.01 (.98)	.15 (.38)	.33 (.05)
Personal Distress PD	-.02 (.93)	-.02 (.92)	-.07 (.70)	-.06 (.72)	.15 (.39)	-.09 (.60)

Table D7

Correlations among old children's self-perception subscales and sensitivity, bias and participants' correlations with patients' self-report and facial actions (p)

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Physical Appearance	-.22 (.22)	-.11 (.53)	-.04 (.81)	.11 (.54)	.07 (.68)	-.04 (.83)
Athletic Competence	-.22 (.23)	-.22 (.21)	-.32 (.07)	-.38 (.03)	.01 (.98)	-.15 (.41)
Behavioural Conduct	.23 (.20)	-.10 (.58)	-.31 (.08)	-.21 (.25)	-.03 (.89)	-.13 (.47)
Scholastic Competence	.24 (.18)	-.001 (.996)	-.25 (.16)	-.17 (.33)	.10 (.59)	-.09 (.64)
Social Acceptance	.16 (.38)	.09 (.64)	-.22 (.23)	-.17 (.35)	.07 (.71)	.11 (.55)
Global Self-Worth	.11 (.54)	-.04 (.82)	-.17 (.35)	-.17 (.35)	.12 (.52)	-.05 (.79)

Table D8

Correlations among old children's IRI subscales and sensitivity, bias and participants' correlations with patients' self-report and facial actions (p)

	Sensitivity		Bias		Correlations with participant ratings	
	mild	strong	mild	strong	patient self-report	patient facial actions
Empathic Concern	.08	-.11	-.32	-.25	-.30	-.08
EC	(.65)	(.55)	(.07)	(.17)	(.09)	(.68)
Perspective Taking	.13	-.12	-.41	-.43	-.11	-.29
PT	(.49)	(.52)	(.02)	(.01)	(.53)	(.11)
Personal Distress	.08	.08	-.48	-.54	.13	.002
PD	(.65)	(.65)	(.01)	(.001)	(.46)	(.99)

Table D9

Correlations among the difference between participants' and patients' ratings and self-perception scales (p)

	participant - patient difference
Physical Appearance	-.16 (.38)
Athletic Competence	-.25 (.17)
Creativity	-.11 (.54)
Friendship	-.34 (.05)
Humor	.05 (.78)
Intellectual Ability	-.16 (.38)
Job Competence	-.47 (.01)
Morality	-.24 (18)
Behaviour with Parents	-.32 (.07)
Romantic Relationships	-.37 (.03)
Scholastic Competence	-.30 (.10)
Social Acceptance	-.24 (18)
Global Self-Worth	-.31 (.08)