THE EFFICACY OF MORPHINE ANALGESIA FOR PROCEDURAL PAIN IN PRETERM INFANTS: SUBJECTIVE JUDGEMENTS AS A MEASURE

by

BRUCE WILLIAM MCMURTRY

B.A. University of Calgary, 1992

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

Department of Psychology, Clinical Psychology Programme

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

August, 1997

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Psychology

The University of British Columbia
Vancouver, Canada

Date Sept. 2/97

DE-6 (2/88)

ABSTRACT

The purposes of this study were twofold. The first was to determine the efficacy of morphine analgesia in reducing procedural pain in preterm neonates. Efficacy in this context was considered a question of whether objectively measurable reductions in facial response to painful stimuli possess clinical and social validity. Subjective judgements made by untrained adults were employed to address the question of efficacy in three ways; 1) to determine whether the changes in facial expression achieved with morphine analgesia represented perceptible changes in pain levels, 2) to determine whether the morphine analgesia resulted in a judged reduction in the need for pain control intervention and 3) to determine whether judges believed morphine to provide an adequate level of pain reduction. The second purpose was to examine the effect of judges' empathic concern on their judgements of pain in this population. It was also expected that infant post-conceptional age would affect pain judgements due to a previously demonstrated attenuation of pain responses in post-conceptionally very young newborn infants.

The judges were 20 university students, a sample size providing a high level of inter-judge reliability. The patients were 50 preterm infants who had been administered morphine analgesia and were subjected to heel stick for blood testing as a part of their hospital treatment in a Neonatal Intensive Care Nursery. The judgement stimuli consisted of three ten-second video clips of each infant. Each of the three clips showed the infant either at rest (no pain), undergoing heel stick with no morphine analgesic (untreated pain)

or undergoing a heel stick with morphine (treated pain). The procedure was not visible in the video clips. Judges were divided into high and low empathic concern groups based on their scores on an empathic concern scale. They were asked to perform two judgement tasks. In the first, they were blind to the condition portrayed in each clip and clips were presented in random order. Judges rated the sensory intensity and unpleasantness of each clip and whether the infant portrayed required pain control intervention (yes or no). In the second, judges were presented with untreated and treated pain clips consecutively for each infant, and they were asked to rate the pain level of the infant in each clip on a visual analogue scale. They were also asked indicate whether the pain reduction achieved was adequate (yes or no).

The results confirmed that the effects of morphine analgesia on preterm infants undergoing painful procedures have social and clinical validity. Judges average ratings of sensory intensity and unpleasantness were significantly reduced when the infants were receiving morphine analgesia. The need for intervention was also significantly reduced with the administration of morphine analgesia. The results were more equivocal with respect to the adequacy of these changes. The general pattern of results suggested that morphine analgesia reduced the suffering experienced during heel stick to a level comparable to "at rest" ratings. However, judges believed even the infants at rest to be experiencing some pain. In the majority of cases (64%), infants were rated as receiving adequate pain reduction, but this percentage indicates that adequate reduction was not universally achieved, possibly due to a need for dosage adjustments. Contrary to

expectations, infant post-conceptional age had no consistent effect on judged pain responses or the rated adequacy of pain reduction.

Conforming with expectations, judge empathic concern had a significant effect on pain judgements, but the relationship was more complex than expected with high empathic concern judges providing higher ratings of sensory intensity and unpleasantness, but lower visual analogue ratings of pain. Overall, judges appeared to be quite sensitive to changes in pain related facial expression. Possible reasons for the lack of infant age related differences and for the complexity of judge empathic concern results, were discussed with attention paid to differences between the two rating tasks. The results of the present study suggest that morphine analgesia can reduce procedural pain in preterm infants in a clinically and socially meaningful way. The non-universality of these findings and the fact that efficacy and clinical significance are ultimately case by case determinations, suggest that the use of morphine analgesia is warranted on a general basis in this population, but individual responses need to be considered. The finding of apparent suffering even in the absence of an acute pain stimulus suggests that more attention needs to be focused on minimizing ongoing low-level suffering. The significant effects of judge empathic concern argues for closer attention to the characteristics of those who must assess pain in preterm infants and for the need to examine the effects of these characteristics on judgements made by different groups of judges (e.g. professional health care providers) and on other populations (older children and adults).

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGMENT

I am grateful for the assistance and support I received from a number of people. I would especially like to thank Dr. Ken Craig for his invaluable advice, guidance and confidence in my abilities. I would also like to thank my research committee, Drs. Eric Eich and Dimitri Papageorgis, for their comments and advice. In addition I would like to express my appreciation to Dr. Marilyn Hill for her assistance and moral support.

I would like to offer my gratitude to Shawn Reynolds for so generously sharing his statistical expertise and to Michelle Mathias, for her time, enthusiasm and diligence. I would also like to thank Shirley Louth, Christine Lilley, Cheryl Gilbert and Christine Chambers for their comments and support.

Finally I would like to thank everyone who provided less tangible but no less vital emotional support: my mom whose undying confidence and encouragement helped me through many crises and most of all, Renee, my wife, whose belief in me makes anything possible.

INTRODUCTION

General Background

Infants who are born early in gestation are typically very fragile and their survival is far from assured. The survival rate of these infants has dramatically increased due to advances in medical knowledge and practice (Fletcher, 1987). However, both intensive medical intervention and long term hospitalization are often required to stabilize preterm infants. Unfortunately, many of the procedures that comprise this improved care would be expected to cause pain (Barker & Rutter, 1994) and, ironically, the very fragility which necessitates special care may also render preterm infants less able to cope with procedural pain (Fitzgerald, Shaw, & MacIntosh, 1988).

While, historically, neonatal pain tended to be ignored and undertreated, there is a growing acceptance that neonates, including those born preterm, do experience pain and require intervention to reduce it (Fletcher, 1987). A variety of both immediate and longer term adverse consequences of untreated pain in neonates has been identified and there is near consensus that pain alleviation is an important goal, at least with regard to more invasive procedures such as surgery (McLaughlin, Hull, Edwards, Cramer, & Dewdney, 1993). However, hospitalized preterm infants must endure a vast array of minor procedures such as heel stick and intubation (Barker & Rutter, 1994) which are known to cause pain in older populations. The pain resulting from such procedures remains largely ignored (Baucher, May, & Coates, 1992). Although these procedures are relatively less

invasive, they are frequent and there are grounds to suggest that even minor procedures might have long term deleterious consequences (Taddio, Goldbach, Ipp, Stevens, & Koren, 1995).

Despite the apparent need, there are special challenges to the development and application of pharmacological interventions for pain reduction in preterm neonates, including altered drug dispositions and the difficulty of assessing treatment effectiveness (Scott et al. 1997). Renal and neurological function are immature in infants born preterm, and these factors may alter the safety and efficacy of analgesic medications commonly used with healthy infants (Scott et al., 1997). The determination of efficacy requires accurate assessment of pain before and after the intervention, as does appropriate clinical application. However, pain assessment in the preverbal infant is complicated by the impossibility of obtaining verbal self-report.

Fortunately, from birth humans possess a well developed repertoire of behaviours for expressing pain (Craig & Grunau, 1993). A variety of these behaviours have been examined as potential sources for pain assessment in preverbal infants, and among the most promising, in terms of availability and specificity, is facial expression. The Neonatal Facial Coding System (NFCS), an objectively coded instrument with demonstrated validity and reliability, has been developed specifically to measure the facial expression of pain in infants. Scott et al. employed the NFCS as a measure of pain and determined that morphine analgesia could safely produce a statistically significant

reduction in the pain expressions of preterm neonates of varying ages undergoing a heel stick procedure. While these results are promising, demonstrating the pharmacological safety of morphine analgesia, the question of efficacy is only partly addressed.

Efficacy has been defined as the "benefits derived from [treatment], its potency, its impact on clients, or its ability to make a difference in peoples' lives." (Jacobson & Truax (1991, p. 12). Treatment efficacy reflects the clinical validity of changes achieved and the practical effects of an intervention for the patient and/or others in their environment (Kazdin, 1992). While objective measures, such as the NFCS can be very sensitive to changes in the level of behavioural expression, they do not on their own indicate what those levels mean.

Pain is an inherently subjective experience and that experience cannot be directly inferred from scores on an objective measure. While measures such as NFCS gain in reliability and sensitivity by minimizing subjectivity, they leave unanswered whether a particular score or change can be interpreted as subjectively meaningful. A change in facial expression which is detectable may not correspond to a change in subjectively experienced pain level. From a clinical perspective, objective measures alone do not indicate when pain is severe enough to require intervention, or when the effects of an intervention have been adequate.

Self-report is sometimes considered the most direct channel for assessing the subjective experience of pain, precisely because it provides the sufferer's own appraisal

of the degree and meaning of the pain. However, self-report is impossible for preverbal infants. For a variety of reasons adult judgements may provide one of the best alternatives.

Infants are born largely helpless and therefore dependent upon caretakers. The immediate reactions of caretakers are predicated upon their understanding of the infant's expressions, but the communication of pain may also have broader consequences for the infant. The responses it elicits from the social milieu can also shape later experiences and expressions of pain (Craig, Lilley, & Gilbert, 1996). Therefore accurate communication of internal states is a necessary adaptive act (Craig, 1997).

Prkachin and Craig (1995) have developed a model depicting the communication of pain through facial expression. The Prkachin and Craig model encompass the sufferer's experience of pain, encoding of that experience in facial expression, and the subsequent decoding of that expression by judges. An instrument, such as the NFCS, focuses on measuring the vehicle of communication, facial expression. It allows a precise quantification of what the facial display contains, but understanding the underlying subjective experience requires inference of that experience. In essence, that inference has been removed in order to maximize the objectivity of scoring. Lost in such a measure is the process of inferring what a particular display means to the sufferer and also the effects the display is having on the observer.

It is likely that we are predisposed by evolution to be able to detect and decode the pain expressions of others (Fridlund, 1994). Part of the process of decoding those expressions and making a pain judgement involves inferring their meaning in terms of subjective experience. Determinations of efficacy require just such inferences. Subjective judgement may therefore provide a better index of efficacy than a more objective measure such as the NFCS. Judgements provide a measure of the effects of an intervention, not only on the individual experiencing pain but on those in the individual's social environment.

However, the subjectivity of such judgements and the fact that they play such a central role in infant pain suggests that they warrant close scrutiny. As might be expected, research shows that pain experience does not correlate perfectly with facial pain expression (Craig, 1980) and that objectively scored pain expressions only imperfectly predict judgements based upon them (Prkachin, Currie, & Craig, 1983).

Judgements, by their subjectivity alone, are bound to vary as a consequence of differences between judges. While the degree of inference implicit in judgements suggests they may permit more direct access to subjective experiences, this may come at the expense of greater vulnerability to individual bias. Since adult judgements play a central role in infant pain and particularly in preterm procedural pain, identifying sources of variability in those judgements is essential to improving the assessment and treatment of these infants.

This study examined the clinical and social validity of the behavioural effects of morphine analgesia in preterm neonates, as well the effects of empathic concern on pain judgements. Both of these issues were addressed from within a judgement study paradigm proposed by Rosenthal (1982), using adult observers' judgements of preterm infants' responses to a heel lancing procedure as the dependent measure.

The objective of the study was to determine whether morphine analgesia results in clinically and socially significant changes in the facial response of preterm infants to heel lancing. More specifically, the goals were: 1) to determine if changes in facial displays resulting from morphine analgesia were considered by judges to represent changes in pain level, 2) if those changes were seen to represent a decreased need for further intervention and 3) whether judges believed those changes represent adequate pain reduction.

Videotaped facial expressions of preterm neonates of varying gestational ages from the Scott et al. (1997) study were employed as the stimuli. Naive observers acted as judges. They were asked to rate the pain experience of infants in 3 conditions: 1) neutral (no pain), 2) heel stick with morphine analgesia, and 3) heel stick without morphine analgesia. The rating scale consisted of ratio scaled adjectives indicating qualitative levels of the apparent pain.

In addition, the judges were asked to indicate whether the infant portrayed in each clip required intervention to reduce it's pain and whether the pain reduction obtained with

morphine was adequate. The percentage of infants receiving adequate reduction of pain was calculated from the adequacy rating.

LITERATURE REVIEW

Procedural Pain in Preterm Neonates

Hospitalization of preterm infants

Infants born preterm are quite fragile and often require prolonged hospitalization, typically in neonatal intensive care units (NICUs). Invasive procedures are a common part of the medical care of these infants. In a recent survey, Barker and Rutter (1994) found that over 3000 procedures were conducted on a sample of 54 consecutive admissions to a neonatal intensive care unit, ranging in age from 23 to 41 weeks gestational age. The majority (74%) of procedures were conducted on the 30 % of the sample who were less than 31 weeks gestational age. The infants of youngest gestational age (23 weeks) underwent the highest number of procedures (488). Procedures conducted on the infants ranged from major surgery to less invasive but potentially painful procedures such as heel sticks for blood testing. Heel sticks were the most common procedure, comprising 56% of the total number of procedures conducted. The findings of this study probably are representative of care in most NICU settings.

The capacity of neonates to experience pain

Historical doubts

Historically, neonatal pain was largely ignored and under-treated relative to that of older children and adults (Fletcher, 1987). This arose largely from then common beliefs

that neonates did not experience or remember events which would precipitate pain in older children and adults (Anand & Hickey, 1987) and a lack of guidelines for appropriate pharmacological treatment of pain in this population (Fletcher, 1987). Since infants are known to process commonly used analgesic agents differently than older children and adults (Koren, Butt, Chin, Yanga, Soldin, & Tom, 1987; Lynn & Slatterly, 1987), leading to potentially serious side effects from inappropriate use, practitioners may often have chosen to err on the side of undertreatment. Procedures such as circumcision, heel stick, venipunctures, intubations, chest tube insertions and suctioning were routinely done without any consideration of the potential for pain (Fletcher, 1987). Although probably less commonly, surgical procedures without anaesthetic have also been documented (Fletcher, 1987).

Anand (1993) has outlined some of the reasons infants' capacity to experience pain has been discounted. Neural and cortical development begun during gestation continues after birth in infants carried to full term. The preterm infant, born early in gestation, is even less fully developed than the full term infant. Despite the recognition that neonates reacted behaviourally to nociceptive stimuli, early studies concluded that these responses were decorticate in nature, and therefore the infants could not perceive or localize pain (McGraw, 1941). From a theoretical perspective, it was argued that a high pain threshold protecting the infant from pain during childbirth would be adaptive (Bondy, 1984).

In recent years a converging body of evidence has been gathered in various fields which points to the fallacy of these views (Anand, 1993; Anand & Hickey, 1987).

Evidence for pain in infancy

From a theoretical perspective, emotions appear at a stage in development when they would be adaptive and the capacity to react to tissue damage would aid in survival and therefore be adaptive from birth (Owens, 1984).

Anand and Hickey (1987) have reviewed a large body of evidence indicating that neonates, even those born preterm, do posses the necessary anatomical and functional requirements for pain perception. Despite the fact that neonates are born with less than fully developed nervous systems, neural pathways and the cortical and subcortical centres necessary for pain perception appear to be well developed by late gestation.

Further suggesting a capacity to experience pain, both full term and preterm neonates have been demonstrated to show a number of physiological and endocrine responses to invasive stimuli, including changes in cardiovascular function, transcutaneous partial pressure of oxygen, and palmar sweating (Anand & Hickey 1987). Anand et al. (1987) demonstrated that infants born at various post-conceptional ages all mounted a potentially dangerous hormonal and metabolic stress response to the physiological insult of surgery, which resulted in increased morbidity and mortality. The magnitude of this response was up to 3 times greater than that seen in adults, although of shorter duration. The very fragility which often necessitates hospitalization of preterm infants suggests that they may

face more risk from untreated pain than their full term counterparts. These stress responses have been reduced in full term neonates and eliminated in preterm neonates (Anand, Sippell, & Aynsley-Green, 1987) with the use of appropriate anaesthetics. While much of Anand and colleagues work has focused on surgical pain, even more minor procedures have potentially harmful results (Taddio Goldbach, Ipp, Stevens, & Koren, 1995).

Infants show a variety of behavioural responses to painful stimuli (reflex withdrawal, crying, and facial grimaces). The similarity of many of these responses to those of older children and adults reporting pain has been used in support of the contention that neonates can feel pain (Owens, 1984).

When subjected to a painful stimulus, adults and children show a reflexive withdrawal of the affected limb. The strength of this response, named the cutaneous flexor withdrawal response (CFWR), has been shown to correlate with various aspects of pain perception in adults (Janko & Trontelj, 1983; Willer, 1985). Preterm infants (25 to 34.5 weeks post-conceptional age) have also been demonstrated to show CFWR to stimulation of the sole of the foot with Von Frey Hairs (Fitzgerald, Shaw & MacIntosh, 1988). Fitzgerald and colleagues found that the threshold for the CFWR response in preterm subjects was very low, increasing with the post-conceptional age of the infants.

In addition to these reflexive movements, neonates of all post-conceptional ages respond to painful stimuli with a wide variety of other behaviours such as facial

expressions, crying, and other body movements which closely parallel those seen in adults and older children (Craig & Grunau, 1993). However, the vigour of these responses tends to be attenuated in infants of younger post-conceptional age (Craig, Whitfield, Grunau, Linton, & Hadjistavropoulos, 1993). This attenuated behavioural response may suggest a limited capacity to experience pain but alternatively it may be attributable to the need for fragile preterm infants to conserve precious energy reserves necessary for survival or development of an immature musculoskeletal system (Craig & Grunau, 1993). Consistent with this interpretation Fitzgerald and MacIntosh (1989) have argued, based on their attenuated CFWR threshold, "that preterm infants, if anything may be supersensitive to painful stimuli". However, CFWR do not appear to be specific to painful stimuli in neonates (Fitzgerald, Shaw, & MacIntosh, 1988) and there is no reason to believe that they provide a better index of experience than do the behavioural responses. The discordance of findings from studies examining different response domains indicates that each can contribute different but important information about the pain experience of preterm infants. It is also apparent that the experience of preterm infants might differ from that of full term infants and that further research is required to fully elucidate any differences.

Long term consequences of pain

There also appear to be long term consequences to untreated pain in neonates and these vary with the post-conceptional age of the infant. For example, in contrast to the

older infants (above 32 week post-conceptional age) in the Fitzgerald, Shaw, and MacIntosh (1988) study who showed habituation of the CFWR, the younger infants (below 32 weeks post-conceptional age) tended to become sensitized with repeated administration of the noxious stimulus. In a later study, Fitzgerald, Millard, and MacIntosh (1989) found that this hypersensitivity could be reversed with the application of a topical anaesthetic cream (EMLA).

There is evidence that such hypersensitivity can last for extended periods. Taddio, Goldbach, Ipp, Stevens, and Koren (1995) found that full term infants who had been circumcised near birth showed exaggerated behavioural responses to immunisation inoculation injections 4 months later.

Also suggesting long term consequences from untreated pain, but contrasting with findings of hypersensitivity, Grunau, Whitfield, Petrie, and Fryer (1991) found that parents of toddlers who had been extremely low birth weight (ELBW: <1000g) infants hospitalized in early life, reported these children to be relatively insensitive to pain. Since these ELBW infants represent children who have extensive pain experience early in infancy, the authors argue that differences in prior pain experience explained the results. Further complicating interpretation of these effects, Grunau, Whitfield, Petrie, and Fryer (1994), found that 4.5 year old children who had been ELBW neonates were higher in somatization, the tendency to report somatic symptoms in the absence of clear organic aetiology, than a matched sample of children who had been carried to full term. Some of

the ELBW children had somatization scores falling above the cut-off for clinical concern and this subgroup could be differentiated from other ELBW children on the basis of various psychosocial and personality measures. This suggests that the relationship between early pain experiences and later somatization is a relatively complex one. Thus, while it would be difficult to argue for the presence of explicit memory for pain in infancy (Craig & Grunau, 1993), the foregoing indicate that early pain experience does have an effect on future pain behaviours.

The evidence for long term consequences to untreated pain argue for the continuity of pain experience from infancy into childhood and for the importance of understanding and minimizing procedural pain in infants.

Current views and practices

There appears to be a growing consensus that neonates are sensitive to pain and attitudes to treatment may be changing as well (Frank, Lund, & Fanaroff, 1986; Fletcher, 1987). In a survey of board-certified neonatal-perinatal physicians, McLaughlin, Hull, Edwards, Cramer, and Dewdney (1993) found that nearly all physicians acknowledged that neonates could perceive pain and more than 75% reported always using anaesthesia for surgical procedures in their practice. A large proportion (87%) reported that attitudes toward pain management in neonates has changed recently. Likewise a survey of head nurses of Canadian neonatal intensive care units revealed that over 85% of the units used analgesia for post operative pain management following cardiac and major surgery

(Fernandez & Rees, 1994). The lack of unanimity amongst the healthcare providers sampled in these studies concerning neonatal pain or the need to provide analgesic interventions is noteworthy.

Despite changing attitudes and practices with regard to major procedures, there appears to be continuing reluctance on the part of practitioners to use analgesia for more minor procedures (Fletcher, 1987). Frank, Lund, & Faranoff (1986) found that 77% of NICU nurses surveyed thought pain medication was under-utilized in neonates. In the Fernandez and Rees (1994) study, 35% of the nurses reported that analgesia was not used routinely following minor surgery. The use of analgesics or anaesthetics for procedural pain was rare in the units sampled. In addition, opiods even when used were not used optimally. Similarly, Baucher, May, and Coates (1992), found that Directors of 38 paediatric intensive care units (PICUs) and 31 neonatal intensive care units (NICUs) reported that analgesics were infrequently used in their units for a variety of minor medical procedures including venipunctures, such as heel stick. For the majority of procedures, a significantly greater number of PICUs than NICUs were reported to use analgesia. Since neonates, including those born preterm, are typically treated in NICUs rather than PICUs, these results would suggest that such patients receive less analgesic medication than older infants and children.

One factor potentially underlying the relative undertreatment of neonates is an attenuated behavioural response to painful procedures (Baucher, May, & Coates, 1992).

The robustness of pain response declines with younger gestational age and therefore the tendency to underestimation and undertreatment of pain may be exacerbated in preterm infants. For example, Shapiro (1991) found that preterm neonates were judged by nurses to be experiencing less pain than full term infants exposed to identical stimulus. The discordance between the obvious need for treatment of procedural pain and the perceptions of observers points to the importance of understanding how adults assess the presence and severity of pain in preterm neonates.

In addition to assessment difficulties, the complexity and unpredictability of dose response to analgesics, particularly in preterm infants, probably contributes to a reluctance to treat pain (Greeley, Boyd, & Kern, 1993). There is a lack of research and clinical experience on the optimal use of analgesics in preterm neonates (Scott et al., 1997). Morphine is one of the most commonly used analgesic agents in neonatal populations and there are some early indications that it can be used safely on preterm infants. However, evidence for its efficacy is limited (Scott et al., 1997). The task of assessment is also central to the task of establishing the efficacy and optimal use of pain control interventions (McGrath, 1987).

Assessment of Pain in Preterm Neonates

General considerations

Accurate assessment is essential to effective pain management (Beyer & Wells, 1989, Porter, 1993) and for research on treatment effectiveness (Kazdin, 1992). Relative to many other species, the human infant is born early in development and is therefore dependent upon adults for care, including pain management. The actions of adult caretakers and therefore the successful management of pain are predicated on their assessments of the infant's pain. From a broader perspective, these actions are important in providing the social context for the infant's pain experience.

The role of the physical and social context in determining the nature of the current pain episode and in shaping features of future pain through learning is central to current bio-psycho-social conceptualizations of pain (Craig, Lilley, & Gilbert, 1996). Bio-psycho-social formulations such as the gate control theory of Melzack and Wall (1965) suggest that pain is a multidimensional experience, which is subject to modulation at many levels by various environmental and experiential variables (Craig & Grunau, 1993).

The impact of context and learning on pain has been well documented in older populations (Fordyce, 1986). The pain of the infant is likewise subject to the biological and environmental context in which it occurs (Craig & Grunau, 1993) A study by Craig and Grunau (1987), demonstrating the biological contextual variable of behavioural

sleep/wake state to have an effect on infant pain expression indicates that at least the biological substrates for the influence of context exist in early infancy.

While it has been cogently argued that because of limited prior pain experience the expressions of neonates are likely to be relatively free of learned components, these early expressions themselves provide the foundation for learning (Craig & Grunau, 1993). The social context provided by caretakers and clinicians will help to shape and define pain as the individual develops from infancy through childhood to adulthood (Craig, Lilley, & Gilbert, 1996). Therefore, assessment as the basis for caretaker and clinician interventions has an impact extending beyond the immediate care the infant receives.

The terms measurement and assessment are sometimes used interchangeably (McGrath & Unruh, 1987), since both refer to determinations about the presence and severity of pain, but distinctions can be made. Measurement is the quantification of various aspects of an experience on some metric (Beyer & Wells, 1989). McGrath and Unruh (1987) suggest that assessment is a broader term, encompassing not only measurement but also the choice of domains to measure, the indices to employ and the provision of meaning to the data collected through consideration of individual and contextual variables. Pain assessment, by this account, refers to the task of inferring the pain experience from the available information, whereas measurement is the more basic process of quantifying that data.

It is assessment which ultimately determines the pain management an infant receives, since providing care entails broad decisions about what behaviours indicate pain, how they should be quantified, and ultimately, their interpretation, but measurement remains the essential core of the task.

In the assessment of pain, measures may range from relatively objective scoring of responses based on a priori specified criteria to subjective judgements of pain based on unspecified, idiosyncratic criteria. The latter is actually more akin to assessment since it is less restricted in domain and latitude for interpretation. Fortunately, despite the complexity of assessing pain in preverbal infants (Pidgeon, McGrath, Lawrence, & MacMurray, 1989), the variety of physiological and behavioural responses mentioned in the preceding section provide a number of potential avenues for pain measurement.

Measures of pain in preterm neonates

Physiological responses

A variety of physiological indices have been employed in assessment of infant pain, primarily because they have been demonstrated to accompany pain in adults (Porter, 1993). Parameters which have been demonstrated to vary in response to pain stimuli such as heel lance or circumcision include increases in heart rate resulting from activation of the sympathetic nervous system (e.g. Booth, McGrath, Brigham, Frewen, & Whittall, 1989; Field & Goldson, 1984; Johnston & Strada, 1986) and decreases in respiratory rate

(Booth, McGrath, Brigham, Frewen, & Whittal, 1989; Field & Goldson, 1984). Decreased oxygen saturation (High & Gorski, 1985) has also been noted, as have rapid fluctuations in intercranial pressure (Durand, Sangha, Cabal, Hoppenbrouwers, & Hodgman, 1989) and decreases in cardiac vagal tone reflecting parasympathetic nervous system activity (Porter, Porges, & Marshall, 1988).

While there has been less research on physiological responses in preterm infants, those that have been investigated parallel the changes seen in full term neonates (Stevens & Johnston, 1994). Preterm infants subjected to heel lance show increases in mean heart rate (Craig, Whitfield, Grunau, Linton, & Hadjistavropoulos, 1993), increases in intracranial pressure (Stevens, Johnston & Horton, 1993), and decreases in oxygen saturation (Craig, Whitfield, Grunau, Linton, & Hadjistavropoulos, 1993). However, there is considerable variability in the physiological responses to tissue insult among preterm neonates (McIntosh, VanVeen, & Bramayer, 1993), likely due to ongoing changes in the nervous system (Porter, 1993).

Physiological indices have other limitations as tools for pain assessment (Craig & Grunau, 1993). Physiological responses to tissue damage are most clearly associated with a general stress response (Gunnar et al., 1984) rather than being specific to pain. While all pain is distressing, distress does not necessarily indicate pain. Therefore, physiological measures alone are not sufficient for pain assessment and alternative measures are required.

Behavioural measures

Neonates also possess a varied behavioural repertoire for expressing pain. Behavioural cues which have been associated with pain in neonates include crying, facial expression, and body movements (Craig & Grunau, 1993). Various studies suggest that these behavioural responses are more consistent and specific to pain in infants than autonomic activity (Craig & Grunau, 1993; Grunau & Craig, 1987; Craig et al., 1993; Johnston et al., 1993).

Cry is generally the initial behavioural response to pain in infants (Johnston & Strada, 1986) and it is also the most salient for observers (Lester, 1984). Because of this, cry would seem to serve as a "biological siren" in eliciting care from adults (Craig & Grunau, 1993). Specific features of cry which can be detected with spectral analysis may provide caregivers with information as to the presence and severity of painful distress (Zeskind & Marshal, 1988). Parameters of pain cry which have been investigated include temporal characteristics and acoustic features using spectral analysis. The temporal characteristics of pain have been shown to vary with the presumed painfulness of a procedure and the acoustic features of cry, analysed with spectral analysis, have been shown to change in response to circumcision versus handling (Porter, 1986).

Cry is more readily available than physiological indices, and appears to contain important information with respect to the presence and severity of pain. However, infants show considerable individual difference in their cries in response to similar painful

stimuli (Johnston & Strada, 1986). In addition preterm infants may not be capable of sustained crying (Wolff, 1987). Finally, specific defining features of pain cry have not been found and no prototypic pain cry has been defined (Fuller, 1991; Johnston, 1993). A more basic problem in the exclusive use of cry as an index of pain in preterm neonates is that many of these infants are intubated for ventilation and therefore unable to cry (Craig & Grunau, 1993).

Body movement and postural changes have also been associated with pain in infants. These include the reflex withdrawal response employed by Fitzgerald et al. (1989), which has been systematically measured with the use of videotape and photogrametric analyses by Franck (1986). However these reflex responses also occur in infants in response to innocuous stimuli and therefore cannot be considered specific to pain in this population. Other body movements have been reported to vary in infants in response to pain (e.g., Craig et al. 1984). The Infant Body Coding System (IBCS; Craig, Whitfield, Grunau, Linton, & Hadjistavropoulos, 1993) scores movement of the hands, feet, arms, legs, head, and torso of the infant as present or absent to provide an overall index of body movement. Craig and colleagues (Craig, Whitfield, Grunau, Linton, & Hadjistavropoulos, 1993) used the system to measure the response of 56 full term and preterm infants to heel lancing and found that activity was greater in a preparatory swab procedure than during rest and greater still during heel lance. However, observer judgements of distress being experienced by term and preterm infants subjected to heel lance were predicted by body movement but to a lesser extent than facial activity (Hadjistavropoulos, Craig, Grunau,

Hadjistavropoulos & Whitfield, , in press). Body movements are less intense in preterm infants than in full term infants, but do not vary with post-conceptional age among those born preterm (Craig, Whitfield, Grunau, Linton, & Hadjistavropoulos, 1993).

Of the behavioural indicators, facial expression appears to be the most specific to pain in infants (Craig & Grunau, 1992). Johnston and Strada (1986) found that the facial response to pain was more consistent across infants than was cry pattern, heart rate, or body movement.

The validity of facial expression as an indicator of pain in adults is well demonstrated. Facial expression changes when individuals are exposed to stimuli known to be painful and those expressions are moderately related to self-reported severity of pain (LeResche & Dworkin, 1988). Facial expression related to pain can be identified by observers, and the ratings of those observers vary with the severity of noxious stimuli (Patrick, Craig, & Prkachin, 1986).

Facial expression is a particularly important channel for interpersonal communication (Ekman, 1982) and infants are relatively well equipped from birth to communicate through facial expression. The musculature of the neonate is relatively well developed, probably to make feeding possible, but also providing the plasticity necessary for communication of complex internal states (Craig, Hadjistavropoulos, Grunau, & Whitfield, 1994). Facial expression is also readily available to caretakers, and the facial

response of infants to noxious stimuli resembles that of adults subjected to acute pain (Craig & Patrick, 1985; Patrick, Craig, & Prkachin, 1986). Healthy full term neonates' facial responses to painful stimuli such as heel lance (Grunau & Craig, 1987), or hypodermic needle sticks and injections (Grunau et al, 1990) have been shown to be relatively consistent, both across infants and across procedures. This response can be differentiated from responses to non-invasive procedures (Grunau et al, 1990). This response can also be used effectively by adults judging pain in infants, both full term (Craig, Grunau, & Aquan-Assee, 1988; Hadjistavropoulos, Craig, Grunau, & Johnson, 1994) and preterm (Hadjistavropoulos, et al., 1994) and is decreased by procedures designed to minimize pain in the full term infant (McCrory, 1990).

Facial activity has been demonstrated to discriminate between noxious and other intrusive events better than bodily activity (Craig, Whitfield, Grunau, & Hadjistavropoulos, 1993) or cry (Hadjistavropoulos, Craig, Grunau, & Johnston, 1994).

There appears to be less individual difference in the facial expression of pain among infants than among adults (Grunau & Craig, 1987). However, Craig, Whitfield, Grunau, Linton, and Hadjistavropoulos (1993) found that facial measures in response to heel lance for blood testing varied according to the gestational age of the infant. Full term infants showed significantly greater facial movement than preterm infants during baseline, swab, lance, and recovery periods. The degree of reactivity of infants to the heel swab vs. baseline and heel lance vs. heel swab also varied by gestational age with older gestational

aged infants showing greater reactivity. The very youngest infants (25-27 weeks gestation) in that study showed minimal facial response to any event. Johnston, Stevens Craig, and Grunau, (1993) similarly found that pain responses were attenuated in infants of younger gestational age.

Attempting to take advantage of the considerable information inherent in facial displays, objective and reliable systems have been developed which allow for detailed analysis of facial expressions in infants.

Objective measurement of facial expression

The most comprehensive system employed for the coding of facial pain expressions is the Facial Action Coding System (FACS), developed by Ekman and Friesen (1978a, 1978b). FACS is an anatomically based, comprehensive, objective system for describing facial movements, designed to facilitate the study of emotional states. The system has been demonstrated to provide reliable and valid descriptions of adult facial expression including pain responses (Craig, Prkachin, & Grunau, 1992), but without modification it is not applicable for use on infants, due to differences in facial morphology (Oster & Rosenstein, 1993).

Oster and Rosenstein (1993) have suggested modifications to the FACS which allow for its use in infant populations and called their system the Baby FACS. Like the FACS, the Baby FACS is anatomically based and capable of capturing all possible variations in

visible facial activity. While this capability provides considerable flexibility, it contributes complexity to learning and applying the system. Attaining proficiency in the FACS, a necessary prerequisite to learning the modifications which comprise the Baby FACS, takes approximately 100 hours (Ekman & Friesen, 1978b). Fortunately, a relatively limited number of facial actions appear to be related to the expression of pain in neonates (Grunau & Craig, 1987).

Capitalizing on the relatively specific prototypical pain expression of infants, Grunau and Craig (1990) have developed the Neonatal Facial Action Coding System (NFCS) from the FACS to specifically study pain in neonates. The NFCS is restricted to scoring of ten items as compared to the 46 facial actions scorable in baby FACS. Five of the NFCS scorable facial movements have consistently been found to be associated with pain in healthy full term new-borns subjected to heel lancing (Grunau & Craig, 1987) and hypodermic needle sticks (Grunau et al. 1990). Two of the NFCS items, taut tongue and tongue protrusion, examine tongue activity, items not appearing in FACS. However, tongue show and movement have been integrated into the latest version of Baby FACS (Oster & Rosenstein, 1993). Contrasting with the clearly defined muscular basis for FACS items, the NFCS items are defined for their prescriptive usefulness. The NFCS has been used widely in the study of pain expressions of infants (Grunau & Craig, 1987)including those born preterm (Craig, Whitfield, Grunau, Linton, & Hadjistavropoulos (1993).

Craig, Hadjistavropoulos, and Grunau (1994) compared the two systems in order to determine if FACS captured any aspects of the infant pain expression which were missed or distorted by NFCS. Trained coders attained similarly high reliability using either system (89% NFCS and 88% FACS). While NFCS and FACS score facial actions somewhat differently, overall composite measures were calculated by summing the individual actions for each child, weighted by a principle components analysis derived coefficient. The composite scores derived from the two systems correlated highly (r=.89). Both systems were sensitive to changes in expression accompanying the invasive event, and descriptive features of the facial expression derived from each were similar.

The authors concluded that the high correlations between the systems "strongly support the supposition that the NFCS captures the facial display during a painful event identified with FACS, but using a less complex and more economical system." They also point out that the NFCS includes taut tongue, a variable missing in FACS but which appears to be particularly important in the infants' facial expression of pain and is not observed with adults. The most recent version of the FACS adapted for infants (Baby FACS; Oster & Rosenstein, 1993) has added scoring for tongue position and shape. This action may reflect the importance of the mouth and tongue in feeding, an argument the authors supported by previous observations that infants demonstrate a high incidence of tongue protrusion in non-noxious, control conditions (e.g., Grunau et al., 1990).

The developers of NFCS suggest that "[b]ecause of its ease of scoring, the NFCS has potential as a clinical tool that may help to identify and manage clinical pain" (Craig, Hadjistavropoulos, Grunau, & Whitfield, 1993, p. 316), but that potential has yet to be realized. Integration of objective scoring systems into clinical practice may be facilitated by a better understanding of the way preterm neonatal pain is currently assessed.

Subjective judgement and the assessment of pain

Subjective judgements can play many roles in pain assessment, including substituting for objective measures. Observers' judgements, based on psychological units of measure (cf. Rosenthal, 1987) (e.g. how much pain is this person experiencing ?) are commonly employed as a measure of pain. Of course, judgement also plays a limited role in scoring most other measures such as the NFCS. As described by Ekman (1982), the difference between judgement in the broader sense and what he calls the measurement of a sign vehicle (objective coding such as NFCS) lies in what the observer or judge does in each process. In the first case, judges make inferences about something, while in the latter they describe the surface of the behaviour. Observers in the measurement of sign vehicles "operate like machines" and to the greatest extent possible inferences are removed from the process (Ekman, 1982). The inference involved in the broader type of judgement involves determination of what cues in the display are important, how they are weighed, and how they relate to the phenomenon of interest.

Both the measurement of sign vehicles and subjective judgements as measures may answer the same question but provide different information (Ekman, 1982). Both provide an indication of the presence and level of a particular state or trait. The judgement approach can tell us if people are capable of detecting the state or trait from particular cues. The measurement of sign vehicles approach can tell us exactly what the available cues are, or what cues differentiate two different displays. In the case of pain measurement, the measurement of sign vehicles approach requires an inference, hopefully based upon empirical validation, that the cues measured actually reflect pain. In the use of subjective judgements based on psychological units of measure, this step is actually integrated into the judgement. The judgement inherently subsumes judges' interpretations as to the meaning of cues. Comparison of sign vehicles measures and subjective judgements can provide an indication of what degree judges employ the information available in a display.

In the task of assessment, judgement may play an even broader role. Subjective judgement may be used to determine the choice of measures and their interpretation. While empirically validated, more objective measures and statistical determinations of significance are generally used in research, in clinical practice, and in some research applications, subjective judgements based on idiosyncratically defined criteria remain the predominant form of behavioural pain measurement and ultimately assessment. The clinician uses subjective judgement to determine the choice of domains to attend to, what cues are relevant within that domain, the presence or absence and/or level of those cues,

and what they mean within the particular context where pain is being assessed. The complexity of this task is largely determined by the time and resources available to the clinician as well as the conclusions she/he must draw.

The clinician is called upon to make numerous assessments daily, in real time, and must make treatment-relevant decisions about pain levels in an individual patient. Pain measures which require recording and detailed coding are less amenable to this type of use as are statistical analyses. There has been at least one successful attempt to use adaptations of NFCS coding in clinical pain assessment (Stevens, Johnston, Petryshen & Taddio, in press); however, this type of use appears to remain rare. Objective measures such as the NFCS are even less likely to be employed by caretakers who are not health care professionals, such as parents. It is these everyday caretakers who become the primary caregivers after hospital discharge and who will therefore play the greatest role in providing context for the infants' pain. For that reason, judgements of both nurses and naive observers are important.

Facial expression in pain judgement

Since subjective judgements play such a central role in clinical and caregiver assessment, it is important to understand how they are made. In order to maintain a reasonably narrow scope, the proposed study and current discussion will be limited to judgements as a substitute for pain measures, with an a priori specified domain of facial activity providing the cues. Restricting access in this way limits the social and clinical

validity of judgements to the extent that judges are not free to choose the domain attended to. However, even judgements restricted in this way do involve a degree of inference which is missing from measures such as the NFCS. Judges must make subjective decisions about the importance and weighting of particular cues within the display. In addition, since the stimuli provided to observers in a judgement task cannot be entirely restricted to the domain of interest (e.g. mechanical apparatus may be visible), judges have access to cues which are not captured by objective measures such as the NFCS. The choice of facial expression as the set of cues provided for the proposed study is based on the existence of an objective measure to which judgements can be compared and on the importance of facial cues in pain assessment.

Studies suggest that observers can use facial information in making judgements of pain in neonates and that, in comparison to cry, facial expression is a relatively greater source of information (Craig, Grunau, & Aquan-Assee, 1988). A survey by Owens (1986) suggests that adults also employ facial responses in judging pain in children. He found that compared to cry and limb movements, nurses' ratings of pain in older children immediately post surgery were most highly correlated with facial reactions. Similarly, Hadjistavropoulos, Craig, Grunau, Hadjistavropoulos and Whitfield (in press) found that facial expression and body movement both predicted judgements by undergraduate nursing students asked to rate the affective discomfort experienced by preterm and full term neonates in response to a heel lance. Facial activity accounted for the most unique

variance in ratings after other variables (body activity, procedural information, and developmental status) had been extracted.

In contrast, a survey of neonatal intensive care nurses asked as to the sources of information they use to judge various levels of pain in their charges, but not provided with specific responses, reported that facial expression was used somewhat but far from extensively (Pigeon, McGrath, Lawrence, & MacMurray, 1989). Only a small percentage of respondents reported using facial expression of infants to indicate either mild, moderate or severe pain. Interestingly, a pink face was reported as indicative of no pain. The authors suggest that given the subtlety of changes in facial expression in this population, nurses may not differentiate among them. However, these results were based upon self-report and may not reflect the actual use of facial cues in clinical practice. Even if facial expression is not regularly employed by adults in assessing infant pain, it would appear likely from the other studies that they are capable of using it if specifically asked. Given the considerable amount of pain-related information available in facial expression, it is important to determine to what extent adults can and do use facial cues.

Subjective judgement and social significance in pain assessment

Despite the tendency for judgements based upon psychological units of measurement to have a lower reliability than more objective measures (Ekman, 1982), they represent valuable information and may have greater validity due to their higher degree of social meaning (Rosenthal, 1987). Subjective judgements based upon psychological units of

measurement are often used as criteria in validating more objective systems (e.g. Grunau & Craig, 1987). Subjective judgements of pain also tend to be relatively reliable, at least between observers (Hadjistavropoulos et al. , in press).

In discussing pain measurement, it is tempting to view the pain behaviour of the infant as a socially isolated phenomenon. However, when discussion moves to subjective judgement, the relevance of broader social factors becomes more salient. Judges bring to the situation various characteristics which may influence their perception or interpretation of the infant's behavioural responses.

Because of their inherent social significance, subjective judgements based upon psychological units of measure may serve as a valuable index of clinical significance or efficacy. In reference to psychotherapy research, but equally applicable to pain treatment, clinical significance has been defined by Kazdin as "the practical value or importance of the effect of an intervention... to the patient or others" (Kazdin, 1992, p. 349). Jacobson and Traux (1991) define efficacy similarly, as "the benefits derived from [treatment], its potency, its impact on clients, or its ability to make a difference in people's lives."

In the one study specifically addressing the effectiveness of analgesic medications in preterm infants, Scott et al (1997) employed NFCS coding as the measure of morphine analgesia. The NFCS scores revealed that morphine analgesia resulted in a statistically significant reduction in pain related facial expression during heel stick. The clinical significance of these findings remains to be demonstrated.

Subjective evaluations of the patient or others is one clinically significant measure suggested by Kazdin (1992). In the case of most preterm infants, two groups of adults are likely to fulfil these criteria. In hospitals, nurses are responsible for the majority of pain assessment and treatment decisions (Coven, 1980; Frank, 1986). In addition, NICU nurses must administer painful procedures to premature infants on a daily basis (Beaver, 1987). Once the infant leaves the hospital parents generally become the primary caretakers and thus will be responsible for the majority of day to day pain assessment and management. Importantly, because of their respective roles, these two groups will also serve as the primary sources of environmental feedback and context for the infant's pain reactions.

Pain judgement as the outcome of nonverbal communication.

It is possible to view the assessment of pain through subjective judgement as part of a process of interpersonal communication. This perspective emphasises the social nature of pain assessment and the various factors which might influence observers' judgements (Craig, Lilley, & Gilbert, 1996; Prkachin & Craig, 1995).

Rosenthal (1987) has discussed in detail how judgement studies can be employed to examine various aspects of the process of nonverbal communication and their relationships. Numerous studies of this type have demonstrated the susceptibility of interpersonal judgements in general to a variety of factors, beyond actual sender state (Rosenthal, Hall, DiMatteo, & Archer, 1979). It would appear likely that similar forces

beyond the actual pain experience might act upon judgements of preterm neonatal pain. Despite their potential importance as measures of neonatal pain and their common use clinically, subjective judgements appear to be less than optimal in terms of use of available information based as they are upon idiosyncratic criteria and perceptual skills (Prkachin & Craig, 1995).

Employing a judgement study paradigm, Prkachin, Berzins, and Mercer (1994), examined the relationship between observer judgements, FACS coded facial expressions, and self-reports of patients with shoulder pain and found these correlated when pain was severe. With less severe pain however, the statistical relationship between observer ratings and self-reports disappeared, despite the fact that self-report still correlated highly with FACS coded expression. Similar studies have yet to be conducted with neonates but would be of obvious importance. However, the inability of infants to provide self-report would prevent analyses of judge accuracy with respect to subjective experience. Despite this limitation, it remains possible and important to analyze relative sensitivity to facial cues among judges.

Given the predominance of subjective judgement in clinical and everyday pain assessment, and the relative importance of facial information in interpersonal communication, it is important to establish how judges make use of the information available in facial displays of pain by preterm neonates. It is equally important to

determine the impact of factors extraneous to the infant's actual experience which impact on those judgements.

The Communication of Pain Through facial expression

An interpersonal communication model of pain

Prkachin and Craig (1995) have outlined a model of pain communication applicable to individuals of any age, which serves as a convenient heuristic for examining the complexities of judging neonatal pain from facial expression. The Prkachin and Craig model describes pain communication as a process encompassing three sequential aspects: 1) the experience of pain, 2) the encoding of that experience as behaviour, and 3) the subsequent response of the observer including the decoding of that behaviour.

The judgements observers make about another person's pain are influenced by a variety of factors extrinsic to the painful stimulus itself. This model is based on a broader conceptualisation of nonverbal communication proposed by Rosenthal (1982), who has written extensively on each stage and relationships among them. Consistent with Prkachin and Craig (1995), the next section of the literature review will be divided into three sections. Each section will outline research and hypotheses specific to pain communication between preterm neonates and adults, focusing on factors at each stage which might impact on pain judgements of those adults. This section will also summarise

relevant information from previous sections and indicate how the proposed study will address the questions raised therein.

The experience of pain in preterm infants: encoder state

The first stage of pain communication is the actual experience of pain. The severity and characteristics of a painful experience may be influenced by various factors internal or external to the sufferer. In the case of adults, there is considerable evidence that the experience of pain may be only partly determined by characteristics of the stimuli, and that it may be influenced by numerous psychological characteristics of the sufferer (Melzack & Wall, 1983). Because of this complexity, unidimensional models of pain which viewed it as a sensory experience reflecting the intensity of the noxious stimuli, have been largely replaced by more complex models which allow for distinctions between various aspects of the pain experience (Fernandez & Turk, 1992). It might reasonably be assumed that the pain experience of the neonate is likewise multidimensional, although the dimensions are likely less differentiated due to the neonates' limited cognitive capacities and relatively limited history of learning opportunities (Craig & Grunau, 1993).

One of the most common distinctions made is between the sensory and affective or unpleasantness dimensions of pain (Fernandez & Turk, 1992; Gracely, McGrath, & Dubner, 1978). This distinction is explicit in the definition of pain accepted by the International Association for the Study of Pain (Mersky & Bogduk, 1994) which includes

the statement that pain is "an unpleasant sensory and emotional experience...". The sensory dimension of pain reflects the intensity, quality, duration, and locus of the pain sensation, while the affective dimension reflects cognitive and motivational variables which are influenced by psychological factors. It is possible to differentiate between different types of pain in adults based on different patterns of pain response (Melzack, 1975). These different dimensions may be detectable by observers. Prkachin, Bezrins, and Mercer (1994) found that self-reported sensory and affective dimensions of pain were significantly related to certain facial actions in adult pain patients asked to perform a painful shoulder movement. There are no data directly supporting the existence of an affective component in infant pain, but the possibility that it occurs in at least a rudimentary form is supported by evidence suggesting emotional capacities in infants (Johnston, 1993), the vigorous autonomic reactions described above, and the considerable degree of continuity over development in other aspects of pain.

As indicated in the Prkachin and Craig model, the experience of pain may be modified by both intrinsic and extrinsic factors. Among the intrinsic factors which might be expected to affect the pain of preterm neonates is the level of neurologic development previously discussed. While the complexity of the pattern of differences in pain response between preterm and full term infants makes conclusions about the relative severity of pain experience difficult, it is clear that preterm infants do experience pain. Other intrinsic factors which have been hypothesized to affect the pain experience of preterm neonates include sex, temperament, sleep state and possibly fatigue, hunger and satiation

and health status (Craig & Grunau, 1993). Sex would seem to be a potentially important variable since there have been suggestions of greater nervous system maturity and sensitivity to tactile stimulation in neonatal girls (Reisman, 1987). Surprisingly, Grunau and Craig, (1987) found that boys were actually quicker to display facial responses to heel lance than were girls. It is unclear how this reduced latency of expression in male infants relates to differences in experience, but it might be expected to influence judgements.

Extrinsic factors discussed by Prkachin and Craig (1995) comprise any exogenous influence on the pain experience. Of primary interest in the case of preterm infants is the potential to reduce pain through environmental interventions such as analgesic drugs. The study by Scott et al (1997) discussed earlier suggested that morphine analgesia can safely and effectively reduce facial response to painful procedures, but that study failed to address issues of clinical significance. Within the context of a judgement study those issues can be addressed. Specifically, observers in this study were asked to judge the severity of pain being experienced by infants with and without exposure to morphine analgesia and to indicate whether the pain reduction achieved with morphine analgesia was adequate for each infant.

Studies of adults indicate that analgesics may act selectively upon either the sensory or affective aspects of pain, so it is important to examine both in determining medication efficacy. Judges in this study were asked to rate the infant's pain on a verbal descriptor

scale of sensory intensity and a scale of affective discomfort (Gracely, Dubner, & McGrath, 1979).

A threshold line (symbolised as "T") displayed in the experience section of the Prkachin and Craig model was included to account for the potential for pain to be present without being expressed in the face. In this model, pain must reach a certain level before it will become evident in the face, and this level will vary between individuals. This component is of considerable interest to the current discussion, since it might explain the confusing pattern of response variability across different measurement domains between preterm and full-term infants. It may explain why preterm infants have an attenuated facial response to noxious stimuli when compared to full term infants when their greater reflexive and stress response would indicate that they are experiencing at least as much pain. Craig and Grunau (1993) have explained this tendency among preterm infants for attenuated behavioural display as resulting from a need to conserve precious energy reserves necessary for survival, and immaturity of facial musculature.

Facial encoding of pain

The Prkachin and Craig (1995) model portrays facial expression as the consequence of activation of a central motor program which leads to discrete changes in facial expression.

A prototypical facial expression of pain has been described in children and infants (Craig, Hadjistavropoulos, Grunau, & Whitfield, 1994). Izzard, Huebner, Reisser, McGuiness, and Dougherty (1980) described the reactions of a sample of 1-19 month-old infants to needle injection as consisting of "brows down and together, nasal root broadened and bulged, eyes tightly closed, and the mouth angular and squarish" (p. 39). This expression could be differentiated from that of anger where the eyes were wide and staring. A similar pattern of facial expressions in response to pain in infants has been described using the NFCS. Five of the ten NFCS facial movements have been shown to consistently be related to pain (Grunau & Craig, 1987; Grunau, Johnston, & Craig, 1990). These include brow lowering, eyes squeezed, deepening of the nasolabial furrow, open lips and mouth, and a taut cupped tongue. Possibly due to differences in facial morphology (Craig, Hadjistavropoulos, & Grunau, 1994), certain features of the infant facial expression do not correspond to the typical adult expression. The adult and infant prototypical expressions are similar in the upper face, but may differ in the region of the mouth (LeResche & Dworkin, 1984). This difference in the mouth region may be due to the fact that neurophysiological innervation of the upper and lower face differ, and the relatively greater plasticity in the lower face of infants which would be expected because of the necessity for rooting and feeding (Craig et al, 1994).

These same five movements were found to characterise the facial responses of preterm infants to heel lance in the Scott et al study. Those actions differentiated between morphine and no morphine conditions as well as between baseline, lance, swab, and

recovery events. Infants of greater post-conceptional age (32-45 weeks) showed significantly greater levels of activity than those of lower post-conceptional age (24-32 weeks). Craig, Whitfield, Grunau, Linton, and Hadjistavropoulos (1993) also found these actions to differentiate between baseline, lance, swab, and recovery events in all but the youngest (25-27 weeks gestational age) preterm infants subjected to heel lance. Another study, which examined the facial expressions of preterm and full term infants soon after birth and those of two month old infants subjected to invasive events, also indicated similar facial expression across all groups with magnitude increasing with both post-conceptional and postnatal age (Johnston, Stevens, Craig, & Grunau, 1993). However, these researchers found that premature infants showed significantly more horizontal mouth stretch than full term 2-month and 4-month olds but less taut tongue than full term new-borns.

Decoder judgements of pain based on facial expression

The importance of pain expressions as a social behaviour in infancy has already been outlined. They serve to solicit aid for the infant, and the way they are perceived and interpreted by adult observers probably helps to define and shape later pain experience throughout the life span. In the third stage of pain communication the display must be decoded by observers.

From the preceding section, it is clear that facial expressions of all but the youngest preterm infants transmit considerable information about pain. Observers are capable of

using this information quite successfully to ascertain pain, but the degree to which nurses report doing so in clinical practice is actually quite low (Pigeon, McGrath, Lawrence, & MacMurray, 1989). There has been no research on the effectiveness of use of facial information by adults judging infant pain, but adult observers appear quite successful in using available facial information in judging the pain of other adults. Observers can detect different levels of painful stimuli and the effects of psychosocial influences on facial expression (Patrick, Craig, & Prkachin, 1986) and can differentiate between expressions of pain and other emotions.

Unfortunately, observers can have difficulty detecting the actual pain experience from this facial information. In a study of adult judgements of other adults, Prkachin et al. (1994) found that only when pain was severe; did self-reports of pain, FACS coded facial expressions, and observers' judgements of the pain correlate highly. When the pain was less extreme, patients' reports correlated with significantly FACS scores, but not with observer ratings. Observer's ratings in general were systematically 50 to 80% lower than those of the patients.

The operation of an underestimation bias could be compounded in judgements of preterm infants, whose facial pain displays are attenuated relative to full term infants. Hadjistavropoulos, Craig, Grunau, and Johnston (1994) found that adults judged infants undergoing injections to be experiencing relatively low levels of pain. As noted, Shapiro (1991) found that nurses tended to rate the pain of preterm infants lower than that of full

term infants exposed to the same stimuli. Taken together these results suggest that preterm infants, particularly those of lower gestational age, may be particularly likely to have their pain underestimated.

In addition to this more general tendency to underestimate pain in others, individual differences between observers might influence their estimates of pain in others. Prkachin and Craig (1995) review evidence that clinical experience will likely lead to greater underestimation of pain. Clinicians who routinely worked with pain sufferers tended to show a greater underestimation of pain being expressed on videotape than did observers with little experience (Prkachin, Solomon, Hwang, & Mercer, 1995). Von Baeyer, Johnston, and McMillan (1984) found that experienced nurses were less likely to correctly identify a facial display as expressing pain than were inexperienced nurses. Prkachin and Craig suggest that these tendencies may be due to selection of nurses with such characteristics or due to relative insensitivity developed over time.

The potential for other biases due to other characteristics of the sufferer or observer are also discussed by Prkachin and Craig. A particular judge factor which deserves consideration is that of empathic concern. Empathic concern reflects the tendency of an observer to feel "warmth, compassion and concern for others undergoing negative experiences" (Davis, 1980). Empathic concern can be reliably measured (Davis, 1980) and it would seem likely that such a tendency would influence the sensitivity of observers to facial expressions of pain. Empathic concern is considered an "other oriented"

emotional reaction to suffering in others (Davis, 1983). It appears to be related to emotional vulnerability, selflessness and concern for others (Davis, 1983). As might be expected, empathic concern was also highly correlated with perspective taking, the tendency to spontaneously adopt the psychological point of view of others, the former representing an emotional reaction and the latter more of a cognitive process.

Summary and Rationale for the Study

Morphine analgesia is one of the most common forms of pain treatment in infants but questions about safety and effectiveness in preterm infants remain. The current study obtained observer ratings of various aspects of infant pain as a measure of the effectiveness of morphine analgesia. Because of their level of inference, these ratings provided a measure of social and clinical validity of the effects of morphine analgesia.

The second goal was to examine the impact of judge empathy on various pain ratings. Both goals were achieved using the same judgement tasks, within one experiment. Judges were presented with three facial expressions of each infant including baseline (at rest), untreated pain (no morphine/heel lance) and treated pain (morphine/ heel lance).

Hypotheses

Based on the literature review, it was hypothesized that;

1) Observers, blind to infant treatment condition, would rate infants undergoing untreated pain as experiencing significantly more pain than the same infants undergoing treated pain or at rest.

2) Observers, blind to infant treatment condition, would be more likely to rate infants undergoing untreated pain as requiring more pain control intervention than the same infants undergoing treated pain or at rest.

3) Observer ratings of pain would significantly differentiate between infant age groups, with post-conceptionally younger infants receiving relatively lower pain ratings than post-conceptionally older infants

4) Judges high on empathic concern would rate infant pain significantly higher than judges low on empathic concern.

METHOD

Ethical approval was obtained from the University of British Columbia Committee for Research and Other Studies Involving Human Subjects. All subjects provided informed consent (see appendix A), with the exception of the infants, whose parents provided consent as part of the earlier Scott et al. (1997) study.

Participants

Student-observers

Twenty one undergraduate students were recruited from an advertisement posted in the UBC Department of Psychology. Subjects were offered their choice of course credits or a small monetary payment (20 dollars) for their participation (see appendix B for recruitment posters). The demographic makeup of the judge sample is shown in appendix C. One participant withdrew from the study after completing task one for undisclosed reasons. Of the remaining 20 subjects, 12 were paid and 8 received subject credits. This number of subjects was considered adequate, given the high levels of reliability achieved.

Neonates

The infants whose videotaped responses were used as the stimuli for this study had been participants in the Scott et al., (1997) study. All infants for whom there was complete videotaped data available were included in this study. In the Scott et al. study, every patient in the Special Care Nursery at British Columbia's Children's Hospital less

than 40 weeks post-conceptional age who received morphine was eligible. The morphine was used only when clinically necessary. Neonates who met the selection criteria were stratified by post-conceptional age, resulting in four groups; Group 1 24-27 weeks, Group II 28-31 weeks, Group III 32-35 weeks, and Group IV 36-39 weeks. Only those infants who received morphine by infusion at the same dosage rate for greater than 60 hours but less than 14 days completed the entire protocol. In total, complete data were collected on 50 infants. The painful procedure observed was heel lancing to collect blood for clinical assessment.

Videotapes

Each observer was presented with two tasks, completed in two separate sessions:

Part One

The tapes for the first task contained 150 ten-second segments portraying each of the 50 neonates in each of three conditions, 1) neutral (at rest), 2) during heel stick for blood testing while being administered morphine and 3) during heel stick while not on morphine. Segments were separated by four seconds of blank blue screen and six seconds of a screen identifying the number of the upcoming clip. The 150 segments were randomly ordered and broken into three blocks of fifty segments each on a separate tape. The presentation order of the three blocks to each observer was counterbalanced to control for order effects.

Part two

This part of the experiment attempted to more closely approximate the task of caretaker judgement, by informing the judges of the presence or absence of morphine in the infant's system and by always presenting the pain/no morphine clip immediately preceding the pain morphine clip for each baby. The tapes for the second task consisted of 100 ten-second segments portraying each of the 50 neonates in each of two conditions, 1) heel stick without morphine followed by 2) heel stick with morphine. The two segments were presented in the same order for each infant, but the order of infants was randomized. Segments were separated by seven seconds of a screen identifying the number of the upcoming clip. The segments were recorded as two blocks of fifty on two separate tapes and the order of block presentation was counterbalanced to control for order effects.

Measures

Neonatal Facial Coding System

NFCS coding was completed as part of the Scott et al study. Each infant was taped a total of four times: at 60 $\pm$ 4 hours of morphine infusion during quiet time (baseline morphine), during heel lance while on morphine (pain/morphine), at four days after discontinuation of morphine infusion during quiet time (baseline/no morphine) and during heel lance after discontinuation of morphine. Segments were scored with the

NFCS by a primary coder who had passed a proficiency test on the NFCS and a randomly selected 25% of the segments were coded by second coder to determine reliability. Reliability as determined by percent agreement for items coded as present by at least one codet was 85%. In order to ensure that the coders were blind to the condition portrayed on the tape, they coded four segments per tape; baseline, swab, lance (mock for the non-pain tapes, where no actual lance occurred) and post lance. For the purposes of coding, each ten second segment was divided into five two second segments and the ten NFCS actions were coded as present or absent during each two second clip. Total scores are a sum of these individual actions over the five seconds each weighted by a factor loading. Descriptions of NFCS variables can be found in Appendix D.

Judge individual differences

Empathic concern

The empathic concern scale is one scale on the Interpersonal Reactivity Index (Davis, 1980), a measure of empathy, the tendency to react to the observed experiences of others. The empathic concern scale has demonstrated reliability and validity (Davis, 1983) and has been used to study empathy and its relationship to aspects of social interaction (e.g. Kochanska, 1997). It consists of 7 items and was administered to all judges at the end of their second task (see Appendix E).

Demographic information

Demographic information was collected from all judges with the questionnaire in Appendix F).

Observation rating scales

Gracely scales

In part one of the study observers were asked to make ratings of the sensory intensity and affective distress of the infant on verbal descriptor scales (Gracely, Dubner, & McGrath, 1979) (see Appendix G) . The adjectival descriptors have been analysed to provide ratio-scaled quantitative values (Gracely, Dubner, & McGrath, 1979). These ratio scales have been demonstrated to be a valid and sensitive tool for the evaluation of pain and pain control methods through self-report. They have also been used successfully and proven to be valid and sensitive in judgement studies (e.g. Prkachin & Mercer, 1989; Hadjistavropoulos, Craig, Grunau, Hadjistavropoulos & Whitfield, 1994; Hadjistavropoulos, McMurtry, & Craig, 1996).

Need for intervention

The need for intervention was rated as a yes or no response to the question "does this infant require intervention to reduce their pain?" in part one of the study (see Appendix G). The ratio scale scores for the verbal descriptors on this scale appear in Appendix H.

Visual analogue scales

For part two of the study, judges were asked to rate the amount of pain they believed the infant to be experiencing on a visual analogue scale (VAS) anchored with "No pain" and "extreme pain" for each segment (see Appendix I). VAS measures are among the most reliable and simple of self-report measures (Abu-Sadd, 1984).

Adequacy ratings

Adequacy judgements consisted of a yes or no response to the question "was the pain reduction adequate" (see appendix H).

Equipment

Infant expressions were recorded on a VHS videotape recorder. Judgement tapes were edited from the source tapes onto a 60-minute VHS videotape and was presented to subjects using a VCR and 19-inch television screen. The observer was seated at a table directly in front of this screen.

Procedure

Participants who provided informed consent were asked to perform two tasks, during one session. Participants were ran one at a time in a separate room with a VCR and television set for stimuli presentation and a table on which to write.

In the first task, the experimenter showed the participant a sample response sheet for the judgement task and allowed for questions. The participant was shown a sample video clip, similar to those to judged, and asked to fill in the sample response. The experimenter then explained that, "each tape contains fifty segments, each one portraying an infant's face either during quiet time, during a heel stick while the infant is on morphine, or a heel stick with no morphine, and these clips are presented in random order. Each clip number corresponds to a number at the top of response sheet." The experimenter explained the difference between affective and sensory aspects of pain experience by way of analogy to sound coming from a radio which can vary in volume or in the type of music (e.g. classical, pop). The sensory intensity of pain was described as analogous to the volume of the radio sound while the unpleasantness of pain was described as analogous to the degree to which the individual liked or disliked the type of music being played.. The experimenter then began the first tape and left the room.

The participant watched the segments in real time in sequential order and had no access to the VCR control. Judgements were to be made during the blue screens. When the end of a tape was reached (i.e. after fifty segments) a message to contact the experimenter flashed on the screen. The experimenter replaced the tape with the second tape, also containing fifty segments and the same procedure was followed. After the second tape was completed, the third tape was started and the same procedure was followed. When all three tapes for part one had been viewed, the participants were requested to take a ten to fifteen minute break. Upon their return, part two began.

The procedure for the second task was identical to that for part one except that the participant was asked to make judgements of a different type and there was a total of 100 clips on two tapes. Participants were informed that they were to "see two clips for each infant, with the no-morphine clip always preceding the morphine clip". After completing the demographic and empathic concern questionnaires, participants were verbally debriefed and provided with the written debriefing form appearing in Appendix J.

Overview of statistical analyses

To establish that judge ratings might appropriately be used as a measure of pain in preterm infants, reliability coefficients were calculated. Cronbach's Alpha (Guilford, 1954), an index of effective reliability, was calculated over the entire sample of judges, for each of the five scales (pain intensity, pain unpleasantness, need for intervention, visual analogue scale and adequacy of pain reduction), in each of the treatment conditions (no pain, pain without morphine and pain with morphine). As recommended by Rosenthal (1987), inter-judge correlations, indices of mean inter-judge reliability were then derived from these alphas using the Spearman-Brown formula (Walker & Lev, 1953).

To divide judges into high and low empathic concern groups, a median split procedure was used. Cronbach's Alpha and average inter-judge correlations were then calculated separately for high and low empathic concern judge groups to determine if it would be appropriate to use the judge groups separately in the remaining analyses.

To address the questions of clinical validity and the effects of judge empathy and infant age on pain ratings, a 4 (infant age group) x 2 (judge empathy) x 3 (pain-morphine condition) between-within MANOVA was conducted with the three judge ratings from task one as dependent variables. Based on the results of the MANOVA, appropriate follow up analyses were conducted. Similarly, a 4 (infant age) x 2 (judge empathy) x 2 (morphine condition) between-within ANOVA was conducted using the task two VAS ratings as the dependent variable. Based on the results of this ANOVA, appropriate follow up analyses were conducted. Given the potential number of follow up analyses, alpha was set at .01 for these tests.

The values for each dependent variable in these analyses were the mean of judges' ratings in each empathic concern group on the relevant scale. In the case of need for intervention (task one), which allowed only for "yes" or "no" responses, "yes" responses were coded as "1" and "no" responses coded as "0". The means for this variable therefore actually constituted a proportion of judges responding yes to the question.

To determine the degree to which morphine analgesia was rated as resulting in adequate pain reduction (task two), the proportion of infants rated as receiving adequate pain reduction was calculated. Similar to the intervention measure, this measure was dichotomous, and so was handled in the same manner as the need for intervention. The proportion reported is actually a mean calculated over all judges in the relevant group. Given significant results for judge empathic concern in the main analyses, separate proportions were calculated for high and low empathic concern judges.

RESULTS

Derivation and Testing of Judge Empathy Groups

Median split

The median judge empathy score was 23. Judges whose empathy scores fell above 23, (M = 22.9; SD = 3.71) were assigned to the high empathy group and those falling below 23(M = 19.9; SD = 2.33), were assigned to the low empathy group. For demographic breakdown of the separate groups refer to Appendix C.

Reliability of ratings

The results of reliability analyses are reported in Table 1. The average effective reliability across all scales was .91, which corresponds to an approximate average mean reliability of r = .30. The corresponding effective reliability for the high and low empathic concern judges, (both = .83) were slightly lower than those calculated across all judges. This is to be expected, given the relationship between effective reliability and the number of judges used (Rosenthal, 1987), but they were sufficiently high to conclude that judges' ratings within each empathy group were measuring the same phenomenon.

Table 1

Reliability for all Judges and for Judges Divided Into High and Low Empathic Concern

Cronbach's alpha (avg. Inter-judge correlations)

Part One		Judge Group	
Condition	All Judges (n=20)	High Empathic Concern Judges (n=10)	Low Empathic Concern Judges (n=10)
Baseline (no pain)			
Gracely Intensity	.90(.30)	.81(.30)	.81(.30)
Gracely Unpleasantness	.89(.29)	.81(.30)	.79(.27)
Need for Intervention	.83(.21)	.73(.22)	.66(.17)
Untreated Pain (pain/no morphine)			
Gracely Intensity	.96(.53)	.93(.56)	.91(.50)
Gracely Unpleasantness	.94(.45)	.93(.56)	.91(.50)
Need for Intervention	.88(.26)	.86(.38)	.87(.40)
Treated Pain (pain/morphine)			
Gracely Intensity	.93(.39)	.87(.40)	.84(.36)
Gracely Unpleasantness	.93(.39)	.85(.37)	.88(.42)
Need for Intervention	.88(.26)	.80(.28)	.76(24)
Part Two			
Visual Analogue Morphine	.90(.30)	.75(.23)	.87(.40)
Visual Analogue No Morphine	.94(.45)	.91(.50)	.82(.32)
Adequacy	.88(.26)	.74(.23)	.81(.30)
Mean over all scales	.91(.30)	.83(.35)	.83(.34)

Analyses of Judged Effectiveness of Morphine Analgesia and the Effects of Infant and Judge Characteristics

For each infant, average scores were calculated for each scale on each condition (no pain, pain/no-morphine and pain/morphine) by computing the average score of all high empathy judges and of low empathy judges on that scale and condition. The resultant data set was subjected to a 4 (infant age) x 3 (condition) x 2 (judge empathic concern), between (infant) within (condition and judge) MANOVA with the rated sensory intensity, unpleasantness, and need for intervention as dependent variables. Six of the seven effects in this analysis were significant based on Wilks' criteria. The exception was the main effect for age. Given that all interactions involving age were statistically significant, it was decided to include that variable in follow up analyses. Separate univariate 4 x 3 x 2 between-within Analyses of Variance were conducted to ascertain the effects of the various independent variables on each rating scale.

Sensory intensity

With rated sensory intensity as the dependent variable, Mauchley's test for non-sphericity was found to be significant for tests involving empathic concern, (X^2, 2, N = 50, = 13.16 , $p < .01$) so Greenhouse Geisser adjustments for degrees of freedom were applied to those tests. Employing these adjustments, the effects of condition, $F(2,92) = 51.78$, $p < .001$, empathic concern, $F(1, 46) = 33.25$, $p < .001$, as well as the interaction between empathic concern and condition, $F(2, 92) = 7.82$, $p < .01$, were significant. Age

was not significant, either as a main effect or in interaction with other variables. The significant interaction was further explored with simple effects analyses of condition at each level of empathy. In simple effect analyses, condition was found to be significant for both low empathy judges, F (2,92) = 367.69, $p < .01$ and high empathy judges F (2,76) = 516, $p < .01$. Tukey's post-hoc analyses were conducted to isolate differences due to the condition in each of the two judge empathic concern groups. The results of these analyses are displayed in table 2.

Table 2

Effects of Condition Within Levels of Judge Empathic Concern (Horizontal) and Judge Empathic Concern Within Condition (Vertical) for Intensity Ratings

	Condition				
Judge Empathic Concern	Untreated Pain (pain/no morphine)		Treated Pain (pain/morphine)		Baseline (no pain)
Low	18.76(10.94)	*	7.95(6.31)	*	6.41(6.05)
	*				*
High	22.35(12.21)	*	8.73(5.12)		8.35(5.66)

* indicates statistical significance at the $p < .01$ level

Among the low empathy judges, (see table 2) sensory intensity ratings were significantly higher in the no morphine/pain (untreated pain) condition, (M = 18.76, SD = 10.94) than in the pain/morphine (treated pain) condition, (M = 7.95, SD = 6.31) where

they were significantly higher than in the no pain/no morphine (no pain) condition (M = 6.41, SD = 6.05). High empathy judges rated the sensory intensity significantly higher in the untreated pain condition, (M = 22.35, SD = 12.21) than in either the treated pain (M= 8.73, SD = 5.12) or no pain (M = 8.354, SD = 5.66) conditions, but rated sensory intensity for the latter did not differ significantly.

In order to better isolate the effects of empathic concern, simple effect analyses of judge empathic concern at each level of condition were also conducted. Empathic concern was found to significantly affect judges' sensory intensity ratings of infants in the no pain, F (1,92)= 15.30, $p < .001$, and the pain/no morphine condition F (1,92) = 52.16, $p < .01$ but not in the pain/morphine condition, F (1,92) = 2.46, $p > .05$. Examination of cell means indicated that in each of the no pain and the untreated pain segments high empathy judges ratings of sensory intensity were higher (see Table 2).

Unpleasantness

The univariate 4 x 3 x 2 mixed ANOVA with unpleasantness as the dependent variable revealed significant main effects for condition, F (2, 92)= 36.77, $p < .001$, and empathy, F (1, 46)= 153.56, $p < .001$, significant two-way interactions for condition by age, F (5, 81)= 2.47, $p < .036$, condition by empathy, F (2,92)= 22.2, $p < .001$, and a significant three-way interaction for condition by empathy by age, F (6, 92)= 2.76, $p <$ .017. The three-way interaction was examined with simple-effect analyses.

First, the effects of infant age and condition were analysed at each level of judge empathy. For high empathic concern judges' ratings of pain unpleasantness, significant main effects were found for both infant age, F (3,80)= 21.62, $p < .01$, and condition, F (2, 80) = 193.46, $p < .01$. The interaction between age and condition, F (5, 80)= 16.63, $p < .01$, was also found to be significant. The effects of condition in each infant age group were examined separately for high and low empathic concern judges in further simple-simple effect analyses. Cell means and significant differences between them are displayed in table 3.

Table 3

Mean Ratings of Unpleasantness in each treatment condition for each Infant PCA Group by each judge group, showing significant differences between conditions

Age group	Condition		
	Untreated Pain	Treated pain	Baseline
		Low empathic concern	
1	12.18_a	4.06_b	5.12_c
2	10.63_a	4.36_b	6.04_b
3	14.96_a	8.18_b	3.78_c
4	13.41_a	6.75_b	5.74_b
		High empathic concern	
1	14.33_a	10.02_b	6.54_c
2	14.28_a	7.52_b	7.87_b
3	18.04_a	15.93_a	5.05_c
4	19.18_a	14.47_b	6.87_c

Note. Age gp1 = (24-27wks PCA); age gp2 = (28-31wks PCA); age gp3 = (32-35wks PCA); age gp 4 = (36-40wks PCA). Means in the same row which do not share the same subscript differ significantly at $p < .01$.

For high empathic concern judges, the effect of condition was significant for all infant age groups (group 1, F (1,58)= 29.75, $p < .01$, age group 2 , F (2,92) = 41.14, $p < .01$, age group 3, F (1,65)= 138.01 $p < .01$, and age group 4, F (2,92)= 47.84, $p < .01$). Tukey's post-hoc analyses were conducted to isolate the differences between conditions for each of the four age groups (see Table 3). These analyses revealed significant differences between all three conditions for age group 1 and age group 4. In Tukey's post hoc comparisons, significant differences were found between the pain/morphine and pain no morphine conditions and between the pain/morphine and the no pain conditions but not between the pain/morphine and no pain conditions in age groups two. In age group three, the no pain condition was significantly lower than either the treated pain or untreated pain, but the latter two did not differ significantly.

For low empathic concern judges, simple effect analysis of infant age and condition also revealed significant main effects for both age, F (3, 92) = 7.35, $p < .01$, and condition, F (2, 80) = 140.99 $p < .01$, and for their interaction, F (5,80) = 7.08, $p < .01$. Tukey's post hoc analyses were conducted to clarify the effects of condition in each age group for the low empathic concern judges (see table 4). In age group one, rated unpleasantness was significantly higher in the untreated pain condition, (M = 12.18, SD = 5.71), than the no pain condition, (M = 5.12, SD = 3.36), which in turn had a higher mean unpleasantness rating than the treated pain, (M = 4.06, SD = 1.66), condition. In age

group 2 the untreated pain condition, (M = 10.63, SD = 8.11), also elicited significantly higher unpleasantness ratings than either the no pain, (M = 6.04, SD = 4.00), or treated pain (M = 4.36, SD = 2.82) conditions but the latter did not differ significantly. For babies in age group 3, once more the untreated pain, (M = 14.96, SD = 6.16), was rated significantly more unpleasant than was the treated pain, (M = 8.18, SD = 4.93), which in turn was rated as more unpleasant than the no pain condition (M = 3.78, SD = 3.08). In age group 4, the untreated pain condition, (M 13.41, SD = 6.73), was also rated as significantly more unpleasant than the treated pain condition, (M = 6.75, SD = 4.75), and the no pain condition (M = 5.74, SD = 5.40) which did not differ significantly from each other.

To explore the effects of age more directly, simple interaction effects for age within each condition within each level of judge empathy were conducted. For low empathic concern judges, age group was significant for the untreated pain condition, $F(3,92) = 9.25$, $p < .01$ and the treated pain condition $F(3, 92) = 9.61$, $p < .01$, but not in the no-pain condition. Similarly, for high empathic concern judges the untreated pain condition $F(3, 92) = 12.62$, $p < .01$, and treated pain condition $F(3, 92) = 38.45$, $p < .01$ but only marginally significant for the no pain condition $F(3, 92) = 3.82$, $p < .05$. Tukey's post hoc analyses of age in each condition were examined separately for the high and low empathy judges. These results are displayed in Table 4.

Table 4

Mean Ratings of Unpleasantness for each Infant PCA Group for High and Low Empathic Concern Judges

Condition	Infant Post Conceptional Age Group			
	Low empathic concern judges			
	gp3	gp4	gp2	gp1
Pain/Morphine (treated pain)	8.18(4.93)	6.75(4.750	4.36(2.82)	4.06(1.67)
	gp3	gp4	gp1	gp2
Pain/No Morphine (Untreated Pain)	14.96(6.16)	13.41(6.73)	12.18(5.71)	10.63(8.11)
	High empathic concern judges			
	gp3	gp4	gp1	gp2
Pain/Morphine (treated pain)	15.93(9.14)	14.47(9.51)	10.02(4.17)	7.52(4.31)
	pg4	gp3	gp1	gp2
Pain/No Morphine (untreated pain)	19.18(9.92)	18.04(7.26)	14.33(4.38)	14.28(8.81)

NOTE. Groups are arranged from left to right in descending order of means for each condition gp1=(24-27wks PCA); gp2=(28-31wks PCA); gp3=(32-35wks PCA); gp4=(36-40wks PCA). Break in line represents significance at $p < .01$

For low empathic concern judges in the pain/morphine condition, age group 3, (M = 8.18; SD = 4.93) rated unpleasantness was not significantly higher than age group 4, (M = 6.75, SD = 4.75), but was significantly higher than age group 2, (M = 4.36, SD = 2.87), and age group 1, (M = 4.06; SD = 1.67). Age group 4 did not differ from age group 2 or group 1. In the pain/no morphine condition, age group 3, (M = 14.96, SD =6.16), did not differ from age group 4, (M = 13.41, SD = 6.73), but was significantly higher than age groups 1, (M =12.18, SD =5.71), and 2, (M = 10.63, SD = 8.11). The latter three age groups did not differ significantly.

For high empathic concern judges (see table 4) within the pain/morphine condition, age groups 3, (M = 15.93, SD = 9.14), and age group 4, (M = 14.47, SD = 9.51), did not differ significantly from each other but were both significantly higher in rated unpleasantness than age group 1, (M = 10.02, SD = 4.17), and age group 2, (M = 7.52, SD =4.31), which did not differ significantly from each other. In the no morphine condition, age group 4, (M = 19.18, SD = 9.92), did not differ significantly from age group 3, (M = 18.04, SD =7.26), but both were significantly higher than group 1, (M = 14.33, SD =4.38), and age group 2, (M = 14.28, SD =8.81), which did not significantly differ from each other.

Requirement for intervention

The univariate ANOVA for the decision to intervene revealed no significant interactions but significant main effects for condition, $F(2,77) = 42.18$, $p < .001$, and

judge empathic concern, F (1,46) = 175.36, $p < .001$, were found. High empathic concern judges, (M = 0.37, SD = 0.16), were significantly more likely than low empathic concern judges, (M = 0.22, SD = 0.15), to indicate that the infants required intervention to reduce their pain. Tukey's post-hoc analyses to clarify the effects of condition revealed that judges were significantly more likely to rate infants as requiring intervention to reduce their pain in the untreated pain condition, (M =.50, SD = 0.29), than in either the treated pain condition, (M =.23, SD = 0.19), or the no pain condition, (M = 0.17, SD =.16), but the latter conditions did not significantly differ ($p < .01$).

Visual analogue ratings

Data from task two were analysed with a 4(age) x 2 (condition) x 2 (empathic concern) mixed ANOVA with VAS pain ratings as the dependent variable. No significant interactions were found, but there were significant main effects for condition, F (1,46) = 167.08, $p < .001$ and judge empathic concern, F (1,46) = 17.98, $p < .001$. Examination of the mean VAS ratings indicated a lower average rating of the infants' pain response by judges high on empathic concern, (M = 4.16, SD = 1.47), than by judges low on empathic concern, (M = 4.62, SD = 1.23). Examination of the mean VAS ratings of infant pain response across all judges for condition, revealed that the morphine condition, (M = 2.53, SD = 1.47), was rated significantly lower than the no morphine condition, (M = 6.25, SD = 1.76). The mean change in ratings between the two conditions, across all judges, was found to be 3.72 (SD = 1.89). Since there was no

interaction between judge empathic concern and condition, separate mean change scores were not calculated for the two groups of judges.

On average, the pain reduction achieved with morphine was judged to be adequate 64% of the time. Given the significant effect of judge empathic concern on judges' VAS ratings of pain, a paired sample t-test was conducted to determine if high and low empathic concern judges were significantly different in their likelihood of rating pain reduction achieved with morphine to be adequate. Low empathic concern judges ($\underline{M}$ = 0.70, $\underline{SD}$ = 0.27) were significantly more likely than high empathic concern judges ($\underline{M}$ = 0.59, $\underline{SD}$ = 0.25) to rate the infant's pain reduction as adequate T^2 (49)= -.465, $p < .001$.

Relationship Between Judge Ratings and Available Facial Cues

This study design allowed only preliminary investigation of this question. To explore the relationship between NFCS coding and the judge ratings, the score on each NFCS variable in each condition was entered into a stepwise multiple regression equation predicting judge ratings of sensory intensity for the corresponding segment and another equation predicting judge ratings of unpleasantness. Separate equations were calculated for high and low empathic concern judge groups. The derived equations were all significant ($\underline{p} < .001$). The mean multiple R for predicting high empathic concern judges' ratings of intensity was .82 (range = .80 to .85) and for unpleasantness it was also .82 (range = .77 to .86). For low empathic concern judges, the mean multiple R for predicting sensory intensity ratings was .84 (range = .79 to .87) and for predicting

unpleasantness the mean multiple R was .82 (range = .74 to .87). The mean multiple R over all judges and scales was .83 indicating that NFCS measured facial actions accounted for a substantial proportion of the variance in judge ratings (69%). In general, a limited number of NFCS variables contributed significantly to the equations, reflecting high inter-correlations between many of the variables.

The relative importance of each NFCS variable is probably best reflected by its correlation with the judge ratings. Table 5 indicates the correlations between each NFCS variable and rated sensory intensity and unpleasantness for each condition in each group. For all conditions and both groups of judges, five NFCS variables consistently showed correlations of .5 or greater with the respective judge ratings of sensory intensity and unpleasantness. Those variables were brow bulge, eye squeeze, open lips, naso-labial fold, and taut tongue. In addition, horizontal mouth stretch also correlated above .5 with low empathy judge ratings in all three conditions. Chin quiver did not occur in either of the no pain or treated pain segments and lip purse did not occur in either of the treated pain or untreated pain segments.

Table 5

Pearson product moment correlation coefficients between NFCS actions and judge ratings on each scale in each condition

	No Pain		Pain/No Morphine		Pain Morphine	
	Intensity	Unpleas	Intensity	Unpleas	Intensity	Unpleas
			Low empathic concern judges			
Naso-labial fold	.79	.74	.77	.73	.80	.78
Brow bulge	.66	.65	.76	.69	.75	.76
Horizontal mouth	.62	.56	.49	.56	.39	.42
Lips open	.62	.54	.63	.65	.51	.51
Taut tongue	.62	.54	.63	.65	.51	.51
Eye squeeze	.61	.61	.80	.76	.79	.79
Vertical mouth	.59	.52	.42	.48	.32	.33
Tongue Protrusion	.74	.08	-.052	-.082	.13	.11
Lip Purse	-.67	-.06	no occurrence	no occurrence	no occurrence	no occurrence
Chin Quiver	no occurrence	no occurrence	.16	.20	no occurrence	no occurrence
			High empathic concern judges			
Brow bulge	.74	.68	.66	.63	.77	.75
Naso-labial fold	.72	.66	.69	.66	.77	.75
Eye squeeze	.68	.63	.73	.69	.78	.76
Lips open	.63	.63	.73	.72	.51	.54
Taut tongue	.63	.63	.73	.72	.5	.54
Vertical mouth	.47	.41	.38	.38	.27	.33
Horizontal mouth	.45	.37	.49	.47	.35	.42
Tongue protrusion	.09	.101	-.17	.22	.16	.08
Lip purse	-.03	-.062	no occurrence	no occurrence	no occurrence	no occurrence
Chin quiver	no occurrence	no occurrence	.45	.15	no occurrence	no occurrence

DISCUSSION

The Efficacy of Morphine Analgesia

The potential harm to preterm infants from untreated repeated exposures to painful stimuli during hospitalization is considerable and the need to develop and evaluate the effectiveness of interventions to minimize that pain is great. While morphine analgesia has been demonstrated to result in statistically significant drops in facial activity presumed to reflect pain, the meaning of these changes in terms of subjective experience and social impact is not clear. These issues can be conceptualized as comprising the question of clinical validity or efficacy. The primary question addressed by this research was the clinical and social validity of morphine analgesia for procedural pain in preterm infants. This question was addressed in three different ways employing subjective judgements as criteria: 1) whether changes in facial expression known to occur with painful procedures and to be reduced with morphine analgesia are detectable and interpreted by observers to represent changes in the infant's pain experience, 2) whether pain stimuli are judged by observers as increasing the need for pain control intervention and whether morphine analgesia reduces that need, and 3) whether morphine analgesia is judged to result in an adequate reduction in the infants' pain experience.

It was expected, given the documented post-conceptional age-related differences in infant responses to painful stimuli (Johnston et al. 1993; Landers, 1990) and the potential

effects of judge empathic concern on judgements, that these factors might moderate the judged efficacy.

Judged reduction in infants' pain experience

The results of this study provide evidence that observers can and do judge infants subjected to painful stimuli without treatment as experiencing more sensory intensity and unpleasantness than the same infants undergoing the same procedure with morphine analgesia or not being subjected to a painful stimuli. These effects held over all infant age groups with both levels of judge empathic concern with the exception of age group 3 ratings of unpleasantness by high empathic concern judges. The difference between treated and untreated pain ratings did not appear to depend upon the judges knowledge of actual treatment condition as it occurred whether or not judges were blind to the treatment condition (task one) or not (task two) although the two tasks involved somewhat different rating scales. The mean sensory intensity rating was reduced from 20.6 which falls between "barely strong" and "slightly intense", to 9.84, which falls between "mild" and "moderate". The mean unpleasantness rating dropped from 14.5, which falls between "slightly intolerable" and "very distressing", to 7.3, which falls between "slightly distressing" and "very unpleasant". The corresponding mean ratings for the "no pain" condition were 7.4, describing a level of sensory intensity falling between "mild" and "moderate", and a level of unpleasantness of 5.8, falling just slightly above "annoying". (see Appendix H for relative magnitudes of particular scores on the Gracely scales)

Notably, the pain levels in the treated pain (pain/morphine) and no pain conditions were quite similar across most age groups and conditions for both levels of judge empathic concern, suggesting that judges did not generally differentiate between the two, however, some significant differences were found. This pattern of results is discussed in more detail under the "Judged Adequacy of morphine analgesia" section to follow.

Also of note is the fact that distress was perceived by the judges even in the no pain condition. There are three potential explanations for this finding. The first is that infants on average displayed some level of background suffering even in the absence of an acute pain stimulus. Alternatively the nature of the study may have created a particular response set. Judges were informed that the study was examining pain and they were aware that the infants are hospitalized. This suggests the second and third potential explanations for non-zero pain ratings in the absence of heel stick. The second possibility would be that judges who were expecting pain display may have been cued to find indicators of pain where there were none or to exaggerate the sensory intensity of those which occur spontaneously during non-pain related displays. For example, an action such as "brow bulge" might occur occasionally in a context unrelated to pain. A judge cued to look for pain might be more likely to read such an action as indicative of pain. A response set might also result in the third possibility, a basement effect on ratings, such that even when no pain was perceived, judges are reluctant to assign a zero pain rating. A basement effect might result either because of general expectations that the infants hospitalized and used in a pain study "must be" experiencing some pain in all conditions

or because of an expectation that heel stick must be painful, combined with uncertainty about whether the clip portrayed represented exposure to the heel stick or not.

The latter two explanations, hinging on the existence of a response set, suggest that characteristics of the judges might influence their tendency to rate "no pain" segments as being painful. If so, it might be expected that high empathic concern judges would be more likely to show such a tendency. High empathy judges would seem most likely to be concerned about missing real signs of pain, potentially resulting in increased sensitivity or in reluctance to assign ratings of zero.

While it is difficult to ascertain if either or both a basement effect or increased sensitivity were occurring in this study, each would have different implications. Increased sensitivity would suggest that ratings in all conditions might be inflated, while a basement effect would suggest that it would be difficult to achieve ratings of total pain elimination of pain in such a study and that the magnitude of treatment effects was actually greater than that measured. Since these processes may be occurring within the judges, further research on how perceptual and cognitive factors predict judgement ratings might help to show if either or both are operating.

Another potential explanation for the perception of distress even in the no pain condition, is that these infants were experiencing some suffering even between painful procedures, possibly as a carryover from previous procedures, or as a result of general sensitivity to more mundane internal or environmental stimuli. This supposition would

seem warranted given the physiological fragility of these infants. The possible existence of consistent low level suffering suggests that continual administration of analgesics during hospitalization may be warranted in this population of infants.

The non-zero ratings of pain in the "no pain" segment also complicate interpretation of the effects of morphine analgesia, particularly when ratings assigned to treated pain (morphine) conditions do not significantly differ from those assigned to the "no-pain" condition. However, the fact that morphine analgesia produced changes in pain ratings from the untreated pain (no pain/ no morphine) condition does provide convincing evidence for its effectiveness.

In general, these results support those of Scott et al. (1997) who found that facial activity was significantly reduced with the administration of morphine analgesia. The current study extends those findings by demonstrating that the changes achieved are perceptible as changes in pain level by naive judges who have no training as to the meaning of those cues. The effects of morphine analgesia were apparent to both high and low empathy judges, suggesting that no special sensitivity to pain communication was necessary to perceive them.

The current findings also make it possible to assign verbal labels to the amount of pain being experienced by preterm infants undergoing procedural by reference to the Gracely descriptor scales. Such verbal labelling might allow easier comparison from this

type of pain to pain of other types, either from different procedures in this population or in other populations, i.e. full term infants, older children or adults.

It should be acknowledged that the relative strength of interaction effects often appeared quite small on examination of post-hoc comparison of means, particularly given the relatively large standard deviations. In the analyses both judge empathic concern and condition were treated as within subject effects and therefore were calculated based on sums across relevant groups of judges. This effectively eliminated the variance attributable to differences between judges within groups, increasing power for these analyses. This may explain the findings of significant interactions when the differences between means relative to respective standard deviations appears relatively small. Given that analyses were conducted at conservative alphas ($p < .01$) and that adjustments made for non-sphericity, always render statistical tests more conservative (Glass & Hopkins, 1984), it must be concluded that these effects are robust.

Judged reduction in need for intervention

Another way to address the clinical significance of morphine analgesia is to determine whether it eliminates the need for further intervention. The results of this study indicated that the perceived need for intervention was significantly reduced with the administration of morphine analgesia, even though judges had no knowledge as to the presence or absence of pain stimuli and/or analgesic. The average rated need for intervention did not return to zero with the administration of morphine, but did fall

significantly below that of the no pain condition, suggesting that, on average, morphine analgesia was judged to be effective. This difference was consistent across infant age groups and judge empathic concern.

Converging with the finding of perceived suffering in the "no pain" condition, there was some perceived need for intervention in that condition as well. It is possible, as it was for the Gracely pain ratings, that a response set, dictated by the nature of the study, led to this finding. The fact that intervention was significantly less likely to be rated as necessary in the treated pain (morphine) condition than in the no pain condition, however, argues against the existence of such an effect on intervention ratings.

This suggests that the level of suffering perceived in the no pain condition is considered worthy of treatment and therefore not trivial. The perceived need for intervention was significantly lower in the morphine condition than in the "no pain" condition, implying that morphine analgesia may reduce ongoing low level suffering.

Rated adequacy of pain reduction

The question of adequacy was approached in two different ways using the available data. The first was to determine whether morphine analgesia reduce the judged pain experience of the infants to a level comparable to judged baseline levels, that is whether it eliminated the pain attributable to the procedure.

Ratings of sensory intensity and unpleasantness did not uniformly return to baseline (no pain stimuli) levels with the morphine analgesia. While the mean ratings for the baseline condition were generally lower than the corresponding mean ratings for the treated pain condition, this difference was not uniformly statistically significant. The significance of effects varied as a function of the pain rating scale used, and in the case of pain unpleasantness it varied as a function of the judges' level of empathic concern and the infants' age. On the sensory intensity measure, although the mean ratings were always lower for the baseline condition, the difference was significant only for low empathic concern judges but not for the high empathic concern judges.

On the unpleasantness rating scale, the relationship was even more complex, with age of the infant predicting the significance of differences between treated pain and no pain conditions. High empathic concern judges significantly differentiated between the no pain and treated pain conditions for all age groups but the second youngest (28-31 wks. PCA.) group of infants. Low empathic concern judges significantly differentiated between no pain and treated pain conditions for the youngest and second oldest group of infants (24-27 wks. and 32-35 wks. PCA) but not for the second youngest and oldest group of infants.

Over both scales all but one significant difference between treated pain and no pain reflected higher ratings in the treated pain condition. In the case of the unpleasantness scale, low empathic concern judges ratings on the youngest infants were significantly

lower for treated pain than for untreated pain. Given the inconsistency in this pattern of results, it is difficult to argue that either sensory intensity or pain unpleasantness levels uniformly did return to baseline levels. Since it is impossible to disprove the null hypothesis such an argument would be questionable in any case.

These results also do not suggest a particular interpretation in terms of the moderating or mediating effects of infant age or level of judge empathic concern since there is no pattern which could be explained on theoretical grounds. There is some evidence, however in the trend of the results, i.e. the tendency for mean ratings to be higher for the treated pain condition across significant and insignificant tests, to surmise that the judges did not, on average, perceive the infants' pain levels to have returned to baseline levels with the administration of morphine analgesia. This suggests that morphine analgesia is not uniformly successful at eliminating judged procedural pain in preterm neonates.

Given that the mean sensory intensity rating for this residual pain represent "mild" to "moderate" sensory intensity and "slightly intolerable" to "very distressing", it is possible that the treatment would be considered inadequate. However, treatment adequacy, like pain level, is a subjective determination.

It is possible that pain reduction might be judged as adequate even when pain was not completely eliminated. Supporting this contention, the judged need for intervention was reduced to below baseline (no pain stimuli) levels by morphine analgesia.

In the most direct measure of adequacy, judges in the second task viewed the majority (64%) of infants to be receiving adequate pain reduction. The mean pain level of the treated pain condition which resulted in pain reduction being rated as adequate over all judges was 1.55 (SD = 2.14), suggesting that judges believed pain reduction could be adequate without total elimination of the infants' pain.

Overall, it would appear that in the majority of cases, infants were judged to have received adequate pain reduction, but the effects were not universally adequate and on average more than a negligible amount of suffering remained. This might be the result of less than perfect dosing for some infants because the appropriate dosage of morphine in this population is still under investigation (Scott et al. 1997).

Judge Empathic Concern

The mean empathic concern score over both sexes (M = 22.9) was similar to that found by Davis, (1980) in his sample of 452 individuals. The range of scores in this sample (14 to 29 out of a possible 35) suggested that this group was sufficiently heterogeneous to allow examination of the effects of empathic concern. Consistent with previous findings (e.g. Davis, 1980) females tended to score higher than males on the empathic concern scale, accounting for the greater proportion of females in the high empathic concern group.

It is important to acknowledge that empathic concern was treated as a within subjects variable. This step was taken since the primary questions of interest involved effects on individual infants, and so effects were calculated based on summing over judges, reducing variability due to judges. This provided greater statistical power for all analyses involving judge empathic concern effects, however, this gain in statistical power limits generalizeability to other populations of judges to some extent (Rosenthal, 1987).

The effects of judge empathic concern were analysed concurrently with those of morphine effectiveness, and therefore the findings already discussed are relevant. The expected relationship between empathic concern and pain ratings held when judges were blind to the infant's pain/treatment condition and were asked to rate the sensory intensity and unpleasantness of the pain on separate scales. In both unpleasantness and sensory intensity ratings, the empathic concern effects were embedded in interactions, but follow up analyses revealed that high empathic concern judges rated the sensory intensity and unpleasantness of pain being experienced by the infant to be greater. The mean sensory intensity rating of 14.5 assigned by high empathic judges over all conditions, falls between the verbal descriptors " "barely strong "(12.6) and "slightly intense" (21.3). The mean pain rating across conditions by low empathic concern judges (11.04) falls between "mild" (5.5) and "moderate" (12.4). For high empathic concern judges the mean unpleasantness rating of 10.48) falls between "slightly distressing" (6.2) and "very unpleasant" (10.7). In the untreated pain condition, where the greatest variation might be expected, high empathic concern judges average rated sensory intensity of 22.35

approached the verbal descriptor close to "strong" while the mean unpleasantness rating of 16.18 corresponded to a verbal descriptor falling between "slightly intolerable and very distressing". The mean ratings by low empathic concern judges (M's = 18.76 and 12.74) to a sensory intensity between "barely strong" and "slightly intense " and an unpleasantness between "very annoying" and "slightly intolerable".

High empathic concern judges also showed a greater likelihood of believing intervention to be warranted over all conditions. These findings suggest that judges of high empathic concern are either more sensitive to pain displays in these infants or they simply interpret the same perceptual input as reflecting more pain and are more liberal in their assignment.

Closer examination of interaction effects suggests that the differences are more likely due to higher perceptual sensitivity. From the interaction effects, high empathic judges' ratings of sensory intensity differentiated between treated pain and baseline condition while the corresponding ratings by low empathic concern judges did not. Likewise, the unpleasantness ratings of high empathic concern judges discriminated between baseline and treated pain conditions for more infant age groups than did those of low empathic concern judges. Thus, it would appear that when judges are blind to the pain and treatment condition, high empathic concern judges are capable of differentiating conditions better. However, high empathic concern judges did not discriminate in their ratings of unpleasantness, between treated and untreated pain for one group of infants.

It is important to note that the treated pain condition and baseline condition differ in both the presence of a painful stimulus and in the presence of analgesic. It is possible that the high empathic concern judges are more sensitive to some aspect of the facial display indicating the presence of the analgesic, outside of its effect on reducing pain, but it is unclear why they would rate it as pain. The relationship between judge empathic concern pain ratings appears to be more complex, however, when the results of task two are examined.

Contrary to expectations, the relationship between judge empathic concern and rated pain levels was reversed when judges made visual analogue ratings of pain level and were informed of the infant's pain treatment condition. In the second task, high empathic judges rated the infants pain as significantly lower than the low empathic concern judges over both morphine and no morphine conditions. Interpretation of this result is complicated by the fact that ratings are summed over both conditions. While the lack of a significant interaction effect argued against statistical analysis of differences between empathic concern groups within each condition, visual examination of the cell means revealed that low empathy judges provided higher ratings in both conditions. Further complicating interpretation, high empathic judges were less likely to rate the pain reduction achieved as adequate. Since a variety of differences exist between the two tasks, only tentative explanations for this discrepancy can be advanced.

Two of the most obvious differences between tasks were the type of scale used and the information about condition provided to the judge. Even if these factors were not confounded, it is difficult to find a compelling explanation for the differential effect of judge empathic concern based upon either or both. It is possible that low empathic concern judges are somehow more susceptible to the information provided in task two that infants are being subjected to a painful stimulus, and that this increases their ratings. Perhaps low empathic concern judges were less sensitive to facial expressions of pain and therefore relied more upon external cues in making their pain ratings. The strength of this argument, however, depends upon the assumption that low empathic concern judges are susceptible to the information on the existence of painful stimulus but less susceptible to the information about morphine analgesia since their pain ratings were higher even in the treated pain condition.

It is also difficult to explain how differences in rating scales might have led to a differential effect of judge empathic concern, since the scales used in the two tasks differ in several ways. The Gracely scales split the phenomenon to be rated into two components (sensory intensity and an affective unpleasantness), while the VAS scale used in task two asked for a more global rating of the "pain response". The Gracely scale instructions which read "circle the words which best describe the sensory intensity and unpleasantness of the pain" may require a higher level of inference requiring complex interpretation of the infant's subjective state. Asking the judge to indicate the infant's level of "pain response" might be interpreted as more of a measurement of sign vehicle

approach (Ekman, 1982). The latter may emphasize the degree of pain the infant is expressing while the latter may focus more on the infant's experience. Interpretation of the extent to which expressions reflect experience is a subjective determination.

Another difference between the task one and task two scales is the number of verbal cues on each scale. The Gracely scales are constructed of fairly specific verbal items, while the VAS scales have only two relatively non-specific anchors, "no pain" and "strong pain".

Any or all of these difference may have contributed to the fact that high empathic concern judges provided higher ratings in task one and lower ratings in task two than did low empathic concern judges. A final potential explanation for the differential effects of judge empathic concern in the two tasks is order effects. The fact that task two was always preceded by task one suggests the possibility of differential sequence effects on the two groups. It is conceivable that the two groups experienced different carry-over effects from task one leading to augmentation of ratings in the low empathic concern group and/or minimiztion of ratings in the high empathic concern group. This possibility could be controlled in future studies where the effects of judge empathic concern were of primary interest by randomizng the presentation of blind versus informed judgement tasks.

Importantly, the average change between VAS pain ratings in the treated and untreated pain conditions did not differ significantly between high ($M = 3.75$, $\underline{SD} = 2.07$)

and low empathic concern judges (M = 3.69, SD = 1.83), indicating that the tendency to lower VAS ratings among high empathic concern judges is stable across the two conditions and that low empathic concern judges were not more sensitive to changes attributable to morphine analgesia.

Infant Post-Conceptional Age

It was expected that younger infants' pain levels would be rated significantly lower in all conditions, but the results were not fully consistent with this hypothesis. The only effect of infant age was found in the interaction of infant age with condition and judge empathic concern on unpleasantness ratings. In the only results consistent with predictions, simple effects analysis revealed that groups three and four were rated as experiencing significantly more unpleasantness by high empathic concern judges than the two younger age groups only in the treated and untreated pain conditions. Among the low empathic concern judges, the mean unpleasantness ratings for infants in age groups 1 and 2 were significantly lower than for age group 3 in both the morphine and no morphine conditions and group one was also lower than age group 4 in the morphine condition, partially supporting the hypothesis.

The lack of a consistent effect for infant age group appears to contradict other research findings suggesting that preterm infants show an attenuated response to painful stimuli (Craig, Whifield, Grunau, Linton, & Hadjistavropoulos, 1993; Johnston, Stevens, Craig, & Grunau, 1993) and that judges rate the pain experience of preterm infants

undergoing heel stick as lower than that of full term infants subjected to the same procedure (Hadjistavropoulos, Craig, Grunau, Hadjistavropoulos & Whitfield, in press; Shapiro, 1991). Perhaps more surprisingly, these results apparently contrast with the effects of age detected with NFCS coding in the Scott et al (1997) study which employed largely the same sample. In that study, the researchers found a main effect of age, such that the two oldest age groups showed significantly greater facial response than the two lowest age groups in terms of their facial responses. Since this effect did not interact with condition, it was argued by the authors to reflect a general tendency toward attenuated response with lower gestational age. The fact that no consistent effect was found using subjective judgements as a measure of nearly the same stimuli, suggests a lack of sensitivity to age effects in observers, which apparently contradicts the Scott et al. (1997) study.

However, the segments analysed in that study were not all employed in this study. In the present study only three of the sixteen separate clips analysed by Scott et al. were presented to judges. Re-analyses of the Scott et al coding data for only the three segments actually included in the current study (see Appendix K) reveal a significant interaction between infant age group and condition. Follow-up analyses revealed a pattern of result very similar to that found in the current study analyzing the high empathy judge ratings of unpleasantness. This suggests a considerable sensitivity among the high empathic concern judges and a convergence between subjective ratings and the more objective NFCS coding.

Judges' Use of Cues

The convergence of high empathic concern judges ratings of unpleasantness and NFCS rated facial activity suggests that judges are using the cues captured in NFCS to a considerable degree in making their ratings. As expected, the judge ratings on each scale in each condition were highly attributable to changes in specific Neonatal Facial Action Coding System variables. The variables which were most predictive of judge ratings were those which tend to be most associated with responses to painful stimuli (e.g. Grunau & Craig, 1993). This is not particularly surprising, given that judges were largely limited to facial expressions in making their judgements, but it does provide evidence that a significant proportion of variance in judges' ratings in all conditions can be attributed to relatively specific facial actions captured by the NFCS.

The mean percentage of variance in judge ratings accounted for by NFCS variables in this study (68%) is comparable to, and actually exceeds the percentage of variance in adult ratings of other adults' pain accounted for by the more comprehensive FACS system variables found in a study by Patrick, Craig and Prkachin (1986).

The strength of this relationship suggests that examination of the particular pattern of facial cues employed by judges in making their ratings would be of interest. Examining the pattern of cue utilization would help to further refine both the interpretation of objective coding of facial activity and the integration of coding systems into actual practice. Although cue utilization could not be tested statistically from the design of this

study since that would require experimental manipulation of cues, the general patterns revealed by correlations between NFCS variables and each judge rating suggests that judges use a similar pattern of cues in making their ratings regardless of their level of empathic concern. Future research more carefully manipulating the levels of various cues provided to judges might be able to better address the use of particular cues by judges. Specifically an approach based on the social judgement theory paradigm (e.g. Brehmer & Joyce, 1988) which has developed out of Brunswick's (1952) lens model of perception and judgement might profitably be applied to the study of judges' use of facial action cues.

Implications for Pain Treatment and Assessment

The findings of this study argue that morphine analgesia can perceptibly reduce the suffering of preterm infants subjected to procedural pain at a magnitude which is clinically important. Morphine analgesia substantially reduced the average judge ratings of both the sensory intensity and affective unpleasantness of procedural pain. It resulted in a significant decrease in the rated need for further intervention and in the majority of cases it resulted in adequate pain reduction as determined by non-professional judges. Thus these results argue for the clinical and social validity of the already demonstrated differences in facial response to procedural pain resulting with morphine analgesia (Scott et al, 1997). While both NFCS scores and subjective judgements provide a relative index of change, subjective judgements assign meaning to those changes. Some important

limitations of this study must be acknowledged in this context. While this provision of meaning to measurable changes represents one important aspect of clinical significance and efficacy, it does not fully encompass those concepts. The study, as conducted and analyzed, did assign subjective meaning to the changes achieved but did so on the basis of mean changes across groups of infants. The clinical significance of treatment depends upon meaningful change being achieved at the individual level (Kazdin, 1992) rather than at the level of group means and so is ultimately a question to be addressed on a case by case basis. This study does suggest that the use of morphine analgesia is warranted on a general basis, but does not speak to the potential for effectiveness in a particular case.

The use of naive judges rather than health care providers who have the most exposure to preterm infants and who at least during hospitalization play the greatest role in their pain treatment is advantageous in some respects but suggests that further research to extend these results to health care workers might be important. Given the evolutionary benefits of humans having developed a sensitivity to pain in others, there is every reason to believe that even those with little or no exposure to infants would be able to accurately decode their pain experiences. The sample of judges employed in this study came from a population who represent those most likely to already be or soon to become parents. Therefore, they represent a group who may be particularly attuned to perceiving the communicative expressions of infants. For the same reason, they represent the population who are most likely to assume caring for preterm infants after release from hospital and

therefore represent a potentially powerful force in providing a context for the infant's pain experiences (Craig. Lilley, & Gilbert, 1996).

The judges used in this study because they were non-professionals, may also have been less likely to have a vested interest in demonstrating or disproving the effectiveness of morphine analgesia. Because of their experience and training, nurses and other health care providers would be seemingly more likely to hold preconceptions about the need for and effectiveness of pain control in this population. The variability in these opinions is well demonstrated (e.g. Baucher et al. 1992; Edwards et al., 1993; Fernandez & Turk 1992). Therefore the relative naiveté of non-professional judges may have reduced biases due to expectations. There is evidence that expectations about the pain experience of targets can bias judgements, and future research should pursue the effects of beliefs about pain and its treatment in neonates on judgements of pain in that population.

The fact that professionals do assess and treat pain within the hospital context does argue for the importance of extending research to judgements made by them. They represent a different but not necessarily more valid index of clinical significance of changes attributable to morphine. It will also be important to compare judgements of professionals and non professionals. The limited research which has addressed the question suggests that professional health care providers may underestimate pain relative to non-providers (Prkachin et al. 1995).

It will also be important to determine how differences in empathic concern or other individual differences influence professionals' judgements of pain level, decisions to intervene, and determinations of treatment adequacy. Although this study demonstrated that subjective judgements can serve as a sensitive and valid measure of preterm infant pain, it also demonstrated that individual differences between judges may result in differences in their judgements. Further study of these individual differences and their effects could help to fine tune everyday pain assessment of preterm infants, by helping those who must conduct those assessments to understand the particular proclivities they bring to the process. It could also help to provide more reliable and accurate interpretation and communication of subjective judgements of pain. The recording and/or communication of subjective judgements is one step removed from the experience judged. In interpreting the meaning of those judgements in terms of the experience of the patient, information about the factors which might be influencing a particular judge's ratings would be helpful in ascertaining the actual experience of the patient. To know that a particular type of judge tends to rate pain higher than another type of judge should play a role in, for example, how a health care provider interprets and integrates those reports of the infants' pain level.

The differential findings across scales speak to the importance of using multiple measures in pain assessment. Given the results of this study, it would appear that ratings of unpleasantness on a verbal descriptor scale most closely approximate what is captured in objective coding of facial expressions. However, affective unpleasantness represents

only one dimension of pain, and therefore may not capture the infant's experience completely. The apparent differential effects of judge empathic concern across the Gracely scales and VAS scales also suggests that the type of scale used needs to be considered in combination with characteristics of the judge when trying to understand the relationship to actual experience. Because of the multi-dimensionality of pain and of the difficulty of capturing a subjective experience in an instrument measuring a single dimension, multi-method assessments are generally recommended for older children (McGrath & Unruh, 1987) and adults (Turk & Melzak, 1992). Given the added impediments to ascertaining the pain experience of preterm infants (Craig & Grunau, 1993) and the potential variability across instruments which was demonstrated in this study, it is likewise advisable, where circumstances permit, to assess infant pain in more than one dimension with more than one instrument.

Clearly the NFCS and subjective judgements provide complementary but not identical information and both can be used profitably in pain assessment. In this study NFCS variables predicted a significant proportion of variance in subjective judgements. The variables which correlated most consistently with subjective ratings included brow lowering, eyes squeezed, deepening of the nasolabial furrow, open lips and mouth, and a taut cupped tongue, those most associated with pain in previous studies (Grunau & Craig, 1987; Grunau, Johnston, & Craig, 1990). Interestingly it would appear that low empathic concern judges attended more to mouth stretch movements as both horizontal and vertical mouth stretch actions correlated more highly with low empathic concern

judge ratings than with high empathic concern judge ratings. This can only be regarded as a trend however, since no statistical test of significance was conducted.

Subjective judgements allow for the assignment of experiential meaning to pain expressions of infants who are incapable of describing their own experience. While subjective judgement is likely to remain the predominant form of pain measurement and assessment in both clinical practice and everyday care of infants, the increased sensitivity of objective systems such as the NFCS provide for more fine grained analysis often required in research. The NFCS might profitably be used to study how judges select and combine cues in making their judgements and how individual differences might influence these processes.

References

Abbu-Saad, H. (1984). Assessing children's responses to pain. Pain, 19, 163-171.

Anand, K. J. S, McGrath, P. I. (1993). Pain in neonates. Amsterdam: Elsevier

Anand, K. J. S. (1986). Hormonal and Metabolic functions of neonates and infants undergoing surgery. Current Opinions in Cardiology, 1, 681-689.

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

Anand, K. J. S. , & Carr (1989). The neuroanatomy, neurophysiology, and neurochemistry of Pain, stress, and analgesia in newborns and children. Pediatric Clinics of North America, 36, 797-821.

Anand, K. J. S. , & Hickey, P. R. (1987). Pain and its effects in the human neonate and fetus. The New England Journal of Medicine, 317, 1321-1329.

Anand, K. J. S, & Hickey, P. R. (1992). Halothane-morphine compared with high does sufantanil for anesthesia and postoperative analgesia in neonatal cardiac surgery. The New England Journal of Medicine, 326, 1-9.

Anand K. J. S. , Sippell, & Aynsley-Green (1987). Randomised trial of fentanyl anaesthesia in preterm babies undergoing surgery; effects on the stress response. Lancet, 31, 243-247.

Barker, D. P. , & Rutter, N. (1994). Exposure to invasive procedures in neonatal intensive care unit admissions. Archives of Diseases in Children, 72, F47-F48.

Barr, R. G. (1989). Pain in children. In: P. D. Wall & R. Melzack (Eds.), Textbook of Pain, 2nd Ed. (pp. 568-689). London: Churchill Livingstone.

Baucher, H. , May, A. , & Coates, E. (1992). Use of analgesic agents for invasive medical procedures in pediatric and neonatal intensive care units. The Journal of Pediatrics, 121, 647-649.

Beaver, P. K. (1987). Infants' responses to touch and pain: Can nurses make a difference? Neonatal Network, December, 13-17.

Berg, W. K. , & Berg, K. M. (1987). Psychophysiological development in infancy: startle and attention. In J. D. Osofsky (Ed.), Handbook of Infant Development (pp. 238-317). Toronto: John Wiley.

Beyer, J. E. & Wells, N. (1989). The assessment of pain in children. Pediatric Clinics of North America, 36, 837-854.

Bondy, A. S. (1984). Infancy. In S. Gabel & M. T. Erickson (Eds.). Child Development and Developmental Disabilities. Boston: Little Brown (pp. 3-19).

Booth, J. C., McGrath, P.A., Brigham, M. C., Frewen, T. C. & Whittall, S. (1989). Pain in infants: distress responses to repeated noxious stimuli. Presented at the First European Conference on Pain in Children, Maastricht, The Netherlands.

Boucher J. D. (1969). Displays of fear, sadness and pain. Perceptual and Motor Skills, 28, 239.

Brehmer, B. & Joyce, C. R. B. (1988). Introduction. In B. Brehmer & C. R. B. Joyce (Eds.) Human Judgement the SJT View. Amsterdam: North-Holland.

Brunswik, E. (1952). The conceptual framework of psychology. Chicago: University of Chicago Press.

Cohen, F. (1980). Post-surgical pain relief: patients' status and nurses' medication choices. Pain, 9, 265.

Craig, K. D. (1980). Ontogenetic and cultural influences on the expression of pain in man. In H. W. Kosterlitz & L. Y. Terenius (Eds.), Pain and Society. Weinheim: Verlag-Chemie.

Craig, K. D. (1982). A social learning perspective on pain experience. In M. Rosenbaum, C. M Franks et al. (Eds.), Perspectives on Behaviour therapy in the Eighties (pp. 311). New York: Springer.

Craig, K. D. (1987). The facial display of pain in infants and children. In G. A. Finley, & McGrath, P. J. (Eds.), Measurement of Pain in Infants and Children, Progress in research and Management, Volume 10. Seattle: IASP Press.

Craig, K. D. & Grunau, R. (1993). Neonatal pain perception and behavioural measurement. In K. J. S. Anand & P. J. McGrath (Eds.), Pain in Neonates (pp. 67-104). Amsterdam: Elsevier..

Craig, K. D. , Grunau, R. V. E. , & Aquan-Assee (1988). Judgement of pain in newborns: Facial activity and cry as determinants. Canadian Journal of Behavioural Science/ Rev. Can. Sci. Comp., 20, 442-451.

Craig, K. D., Grunau, R. E., Hadjistavropoulos, H. D. & Whitfield, M. F. (in press). Judging pain in infants: Behavioural, contextual, and developmental determinants. Pain.

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. , Lilley, C. M. , Gilbert, C. A. (1996). Social barriers to optimal pain management in infants and children. The Clinical Journal of Pain, 12, 232-242.

Craig, K. D. , McMahon, R. J. , Morison, J. D. & Zaskow, C. (1984). Developmental changes in infant pain expression during immunization injections. Social Science and Medicine, 12, 1331-1337.

Craig, K. D. & Prkachin (1983). Non-verbal measures of pain. In R. Melzack (Ed.), Pain measurement and Assessment. New York: Raven Press.

Craig, K. D. , Prkachin, K. M. & Crunau, R. V. E. (1992). The Facial Expression of pain (pp. 257-276). In D. C. Turk & R. Melzak (Eds.). Handbook of Pain Assessment. (pp 257 - 274).

Craig, K. D. , Whitfield, M. F. , Grunau, R. V. E. , Linton, J. , & Hadjistavropoulos, D. (1993). Pain in the preterm neonate: behavioural and physiological indices. Pain, 52, 287-299.

Dale, J. C. (1986). A multidimensional study of infants' responses to painful stimuli. Pediatric Nursing, 12, 27-31.

Dale, J. C. (1989). A multidimensional study of infants' behaviors associated with assumed painful stimuli; phase II. Pediatric Health Care, 3, 34-38.

Davis, M. (1980). A multidimensional approach to individual differences in empathy. JSAS Catalog of Selected Documents in Psychology, 10, 85.

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

Dixon, S. , Snyder, J. , Holve, R. , & Bromberger, P. (1984). Behavioural effects of circumcision with and without anesthesia. Journal of Behavioural Pediatrics, 5, 246-250.

Durand, M. , Sangha, B. , Cabal, L. A., Hoppenbrouwers, T. & Hodgman, J. E. (1989). Cardiopulminary and intracranial pressure changes related to endotracheal suctioning in preterm infants. Critical Care Medicine, 17, 506-510.

Ekman, P. (1982). Methods for measuring facial activity. In K. R. Scherer& P. Ekman (Eds.), Handbook of Methods in Nonverbal Behavior research (pp. 45-90). Cambridge: Cambridge University Press.

Ekman, P., & Freisen, W.V. (1978). Facial action coding system, investigator's guide. Consulting Psychologists Press. Palo Alto.

Ekman, P. , & O'Sullivan, M. (1991). Who can catch a liar? American Psychologist, 9, 913-920.

Eland, J. M. , & Anderson, J. E. (1977) The experience of pain in children. In A. K. Jacox (Ed.), Pain: a source book for nurses and other health care professionals (pp. 453-473). Boston: Little Brown.

Fernandez, C. V. , & Rees, E. (1994). Pain management in Canadian level 3 neonatal intensive care units. Canadian Medical Association Journal, 150, 499-504.

Fernandez, E. & Turk, D. C. (1992). Sensory and affective components of pain: Separation and synthesis. Psychological Bulletin, 112, 205-217.

Field, T. , & Goldson, E. (1984). Pacifying effects of non-nutritive sucking on term and preterm neonates during heelstick procedures. Pediatrics, 74, 1012-1015.

Fitzgerald, M. (1991). The developmental neurobiology of pain. In M. R. Bond, J. E. Charlton & C. J. Woolf (Eds.), Proceedings VIth World Congress of Pain (pp. 253-261). Elsevier: Amsterdam, 253-261

Fitzgerald, Millard, M. , & MacIntosh, N. (1989). Cutaneous hypersensitivity following peripheral tissue damage in newborn infants and its reversal with topical anaesthesia. Pain, 39, 31-36.

Fitzgerald, M. , Shaw, A. , & MacIntosh, N. (1988). The postnatal development of the cutaneous flexor reflex: a comparative study in the premature infants and newborn rat pups. Developmental Medicine & Child Neurology 30, 520-526.

Fletcher, A. B. (1987). Pain in the neonate. New England Journal of Medicine, 317, 1347-1348.

Franck, L. S. (1987). A new method to quantitatively describe pain behavior in infants. Nursing Research, 35, 28-31.

Franck, L. S. (1986). A new method to quantitatively describe pain behaviour in infants. Nursing research, 35, 28-31.

Fridlund, A. J. (1994). Human Facial Expression: An Evolutionary View. San Diego: Academic Press.

Frodi, A. , & Lamb, M. (1980). Child abusers' responses to infant smiles and cries. Child Development, 51, 238-241.

Fordyce, W. E. (1986). Learning processes in pain. In R. A. Sternbach (Ed.), The Psychology of Pain (pp. 10-23). New York: Raven Pres.

Fuller, B. F. (1990). Potential acoustic measures of infant pain and arousal. In D. C. Tyler & E. J. Krane (Eds.), Advances in Pain Therapy and Research, Pediatric Pain, Vol. 15. New York: Raven Press.

Fuller, B. F. (1991). Acoustic Discrimination of three types of infant cries. Nursing Research, 40, 156-160.

Ganchow J. R., Steiner J. E., Kleiner, M. , & Edelstein E. L. A. (1976). A multidisciplinary approach to the expression of pain in psychotic depression. Perceptual and Motor Skills, 47, 379-398.

Glass G. V. & Hopkins, K. D. (1984). Statistical Methods in Education and Psychology. Englewood Cliffs, N.J.: Prentice-Hall.

Gracely, R. H. , Dubner, R. , & Mcgrath, P. A. (1979). Analgesia: Fentanyl reduces the intensity but not the unpleasantness of painful tooth pulp sensations. Science, 20, 1261-1263.

Gracely, R. H. , Mcgrath, P. A. , & Dubner, R. (1978a). Ratio scales of sensory and affective verbal pain descriptors. Pain, 5, 5-18

Gracely, R. H. , Mcgrath, P. A. , & Dubner, R. (1978b). Validity and sensitivity of ratio scales of sensory and affective verbal pain descriptors: Manipulation of affect by diazepam. Pain, 5, 19-29.

Greeley, W. J., Boyd, J. L. & Kern, F. H. (1993). Pharmacokinetics of analgesic drugs. In K. J. S. Annand & P. J. McGrath (Eds.), Pain in Neonates. (pp. 107-154). Amsterdam: Elsevier.

Grunau, R. V. E. , & Craig, K. D. (1990). Facial activity as a measure of neonatal pain expression (pp. 147-155). In D. C. Tyler & Elliot (Eds.), Advances in Pain Research Therapy. New York: Raven Press.

Grunau, R. V. E. , & Craig, K. D. (1987). Pain expression in neonates: Facial action and cry. Pain, 28, 395-410.

Grunau, R. V. E., Johnston, C. C. , & Craig, K. D. (1990). Neonatal facial and cry responses to invasive and non-invasive procedures. Pain, 42, 295-30.

Grunau, R. V. E. , Whitfield, M. F. , & Petrie, L. (1994). Pain sensitivity and temperament in extremely low-birth-weight premature toddlers and preterm and full-term controls. Pain, 58, 341-346.

Grunau, R. V. E. , Whitfield, M. F. , Petrie, J. H. , & Fryer, L. (1994). Early pain experience, child and family factors, as precursors of somatization: a prospective study of extremely premature and fullterm children. Pain, 56, 353-359.

Grunau, R. V. E., Whifield, M. F. , Petrie, J. H. , & Fryer, L. (1991). Pain sensitivity in toddlers of -birthweight <100 grams compared with heavier preterm and full birthweight toddlers. Pediatric Research, 29, 256A (Abstract).

Gunnar, M. R. , Conners, J. , Isene, J. , & Wall, L. (1984). Adrenocortical activity and behavioral distress in human newborns. Developmental Psychobiology, 21, 297-310.

Guilford, J. P. (1954). Psychometric Methods. Second Edition. New York: McGraw-Hill.

Hadjistavropoulos, H. D., Craig, K. D., Grunau, R. V. E. & Johnston, C.C. (1994). Judging pain in newborns: Facial and cry determinants. Journal of Pediatric Psychology, 19, 485-491.

Hadjistavropoulos, H. D., Craig, K. D., Grunau, R. V. E. & Whitfield, M. F. (in press). Judging pain in infants: Behavioural, contextual, and developmental determinants.

Hadjistavropoulos, T. , McMurtry, B. , & Craig, K. D. (in press). Beautiful faces in pain: Biases and accuracy in the perception of pain. Psychology and Health, 11, 411-420.

Hadjistavropoulos, H. D. , Ross, M. M. , & Von Baeyer, C. (1990). Are physicians' ratings affected by patients' physical attractiveness? Social Science and Medicine, 31, 69-72.

Harpin, V. A. , & Rutter, N. (1983). Making heel pricks less painful. Archives of Disorders of Childhood, 58, 226-228.

Hartley, R. , Green, M. , Quinn, M. , & Levene, M. I. (1993). Pharmacokinetics of morphine infusion in premature neonates. Archives of disease in childhood, 69, 55-58.

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

Hester & Foster (1991). Critical dimensions of nurses' decisions about children's pain. Journal of Pain and Symptom Management, 6, 148.

Holland, R. , Anderson, B. , Watson, T. , & McCall, E. (1994). Introduction of continuous regional techniques for postoperative paediatrics patients: one years' experience from two hospitals. New Zealand Medical Journal, 107, 80-82.

Izzard (1971). The face of Emotion. New York: Appleton.

Izzard, C. E. (1979). The Maximally Discriminative Facial Movement Coding System (MAX). Newark, N. J. : University of Delaware Instructional Resources Center.

Izzard, C. E. , Hembree, E. A. , Dougherty, L. M. , & Cross, C. L. (1983). Changes in facial expressions of 2-to-19 month old infants following acute pain. Developmental -Psychology, 19, 418-426.

Izzard, C. E. , Hembree, E. A. , & Huebner, R. R. (1987). Infant's emotion expressions to acute pain: developmental change and stability of individual differences. Developmental Psychology, 23, 105-113.

Izzard, C. E. , Huebner, R. R. , Risser, D. , McGines, G. C. , & Dougherty, L. M. (1980). The Young infant's ability to produce discrete emotion expressions. Developmental Psychology, 16, 132-140.

Jacobson, N. S. & Truax, P. (1991). Clinical significance: A Statistical approach to defining meaningful change in psychotherapy research. Journal of Consulting and Clinical Psychology, 59, 12-19.

Jacobson, N. S. & Truax, P. (1991) Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. Journal of Consulting and Clinical Psychology, 59, 12-19.

Janko, M. & Trontelj, J. V. (1983). Flexion withdrawal reflex as recorded from single human biceps femoris motorneurons. Pain, 15, 167-176.

Johnston, C.C. (1993). Development of psychological responses to pain in infants and toddlers. In N. L. Schechter, C. B. Berde and M. Yaster (Eds.), Pain in infants, children, and adolescents (pp. 65-74) . Baltimore, MD.: Williams & Wilkins .

Johnston, C. C. , & Strada (1986). Acute pain response in infants: A multidimensional description. Pain, 24, 373-382.

Johnston, C. C., Stevens, B., 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.

Jones, (1989). Identifying signs that nurses interpret as indicating pain in newborns. Pediatric Nursing, 15, 76-79.

Kazdin, (1992) Research Design in Clinical Psychology. Needham Heights, Ma

Kochanska, G. (1997). Mutually Responsive orientation between mothers and their young children: Implications for early socialization. Child Development, 68, 94-112.

Koren, G. , Butt, W. , Chinyanga, H. , Soldin, S., Tan, Y. K. (1985). Postoperative morphine infusion in newborn infants: assessment of disposition characteristics and safety. Journal of Pediatrics, 107, 963-967.

Landers, J. (1990). Clinical judgements in pain management. Pain, 42, 15-22.

Leresche, L. & Dworkin, S. F. (1984). Facial expression accompanying pain. Social Science and Medicine, 19, 1325-1330.

LeResche, L. & Dworkin, S. F. (1988). Facial exprespressions of pain and emotions in chronic TMD patients. Pain, 35, 71-78.

Lester, B. M. (1984). A biosocial model of infant crying. In L Lipsitt (Ed.), Advances in infancy research (pp. 167-212). New York: Ablex.

Loeser, J. D. (1980). Perspectives on pain. In Proceedings of First World Conference on Clinical Pharmacology and Therapeutics, (pp. 313-316). London: Macmillan.

Lynn, A. M. & Slattery, J. T. (1987). Morphine pharmacokinetics in early infancy. Anesthesiology, 66, 136-139.

Marchette, Main, Redick, Bagg, & Eatherland (1991). Pain reduction interventions during neonatal circumcision. Nursing Research, 40, 241-244.

Marshal (1989). Neonatal pain associated with care giving procedures. Pediatric Clinics of North America, 36, 885-903.

Marshall, R. E. , Stratton, W. C. , Moore, J. A. , & Boxerman, S. B. (1980). Circumcision. I. Effects upon newborn behaviour. Infant Behavioural Development, 3, 1-14.

Marshall, R. E. , Porter, F. L. , Rogers, A. G. , Moore, J. A. , Anderson, B. , & Boxerman, S. B. (1982). Circumcision. II. Effects upon mother-infant interaction. Early Human Development, 7, 367-374.

Marvin J. A. & Pomietto, M. (1991). Pain assessment in infants (0-12) months using Neonatal Facial Action Coding System. Journal of Pain Symptom Management, 6, 193 - 200.

McCaffery & Ferrell (1992). Does the gender gap affect your pain-control decisions?. Nursing, 42, 49-52.

McCaffery & Ferrell (1992). Does Life-style affect your pain-control decisions?. Nursing, 42, 58-61.

McCaffery & Ferrell (1992) Patient age, does it affect your pain-control decisions?. Nursing, 42, 44-48.

McCrory, L. (1990). The use of continuous sensory stimulation: proprioreceptive-tactile and auditory stimulation to reduce newborns' response to pain, Masters Thesis University of Florida, 1990.

McGrath, P. A. (1987). An assessment of children's pain: a review of behavioral, physiological and direct scaling techniques. Pain, 31, 147-176.

McGrath, P. A. & Craig, K. D. (1989). Developmental and psychological factors in children's pain. Acute pain in children, 36, 823-836.

McGrath, P. J. & Unruh, A.M. (1987). Pain in children and adolescents. Elsevier, Amsterdam.

McIntosh, N., VanVeen, L., & Bramayer, H., (1993). The pain of heel prick and its measurement in preterm infants. Pain, 52, 71-74.

McLaughlin, Hull, Edwards, Cramer & Dewdney (1993). Neonatal pain: A comprehensive survey of attitudes and practices. Journal of Pain and Symptom Management, 8, 7-16.

Mcgraw, M. (1943). The neuromuscular maturation of the human infant. New York: Columbia University Press.

Melzack, R. (1975). The McGill Pain Questionnaire: Major properties and scoring methods. Pain, 1, 277-299.

Melzack, R. & Wall, P. D. (1965). Pain mechanisms: A new theory. Science, 150, 971-979.

Melzack,R. & Wall, P. D. (1983). The challenge of pain. New York. Basic Books.

Mersky & Bogduk (1994). Task force on taxonomy of the international association for the study of pain. Classification of chronic pain: Descriptions of chronic pain syndromes and definitions of pain terms 2nd ed.

Murray, A. D. (1979). Infant crying as an elicitor of parental behaviour: An examination of two models. Psychological Bulletin, 86, 191-215.

Murrell, D. , Gibson, P. R. , & Cohen, R. C. (1993). Continuous epidural analgesia in newborn infants undergoing major surgery. Journal of Pediatric Surgery, 28, 548-553.

Obara, H., Sugiyama, D. Maekawa, N. et al. (1984). Plasma cortisol levels in paediatric anaesthesia. Canadian Anaesthesia Society Journal, 31, 24-27.

Olkkola, K. T. , Hamunen, K. , & Maunukesla, E. (1995). Clinical pharmacokinetics and pharmacodynamics of opioid analgesics in infants and children. Clinical Pharmacokinetics, 28, 385-404.

Osofsky, J. D. (1976). Neonatal characteristics and mother-infant interactions in two observational situations. Child Development, 47, 1138-1147.

Oster, H. & Rosenstein, D. (1993). Baby FACS: Analyzing Facial Movement in Infants. Unpublished Manuscript.

Owens, M. E. (1986). Assessment of infant pain in clinical settings. Journal of Pain and Symptom Management, 1, 29-31.

Owens, M. E. (1984). Pain in infancy: Conceptual and methodological issues. Pain , 20, 213-230

Patrick, Craig, K. D. , & Prkachin, K. M. (1986). Observer Judgements of acute pain: facial action determinants. Journal of personality and Social Psychology, 50, 1291-1298.

Pasero, C. L. (1995). Assessing pain neonates. American Journal of Nursing, see e-mail

Pidgeon, H. M. , McGrath, P. J. , Lawrence, J. , & MacMurray, B. (1989). Nurses' perceptions of pain in the neonatal intensive care unit. Journal of Pain and Symptom Management, 4, 179-183.

Pigeon, H. M. et al. (1989). Nurses' perceptions of pain in the neonatal intensive care unit. Journal of Pain and Symptom Management, 4, 179-183.

Porter, F. L. (1993) Pain assessment in children: Infants. In N. Schecter, C. Berde, & M. Yaster (Eds.), Pain in Infants, Children and Adolescents (pp 87-96) . Baltimore: Williams & Wilkens.

Porter, F. L. , Miller, R. H. , & Marshall R. E. (1986). Neonatal pain cries: effect of circumcision on acoustic features and perceived urgency. Child Development, 57, 790-802.

Porter, F. L., Miller, R. H. & Marshall, R. E. (1986). Neonatal pain cries: Effect of circumcision on acoustic features and perceived urgency. Child Development, 57, 790-808.

Porter, F. L. et al. , (1991). A controlled clinical trial of local anaesthesia for lumbar punctures in newborn infants. Pediatrics, 88, 663-669.

Porter, F. L. Porges, S. W. , & Marshall, R. E. (1988). Newborn pain cries and vagal tone: parallel changes in response to circumcision. Child Development, 59, 495-505.

Prkachin, K. M. (1992). Measurement in the face of pain. American Pain Society Journal, 1, 167-169.

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

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 & Craig, K. D. (1983). Judging nonverbal expressions of pain. Canadian Journal of behavioural 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., Berzins, S. & Mercer, S. R. (1994). Encoding and decoding of pain expressions: a judgement study. Pain, 58, 253-259.

Prkachin, K. M. , Solomon, Hwang & Mercer S. R. (1995). Does experience affect judgements of pain behaviour?: Evidence for relatives of pain patients and health care providers. Manuscript submitted for publication

Reisman, J. E. , (1987). Touch, motion, and proprioception. In P. Sapatek & L. Cohen (Eds.), Handbook of Infant Perception (pp. 365-303). New York: Academic Press.

Richards, M. P. M. , Bernal, J. F. & Brackbill, Y. (1976). Early behavioral differences: gender or circumcision? Developmental Psychobiology, 9, 89-95.

Rosenthal R. (1987a). Conducting judgment studies. In K. R. Scherer & P. Ekman (Eds.), Handbook of methods in nonverbal behavior research. (pp. 287- 361). London: Cambridge University Press.

Rosenthal R. (1987b). Judgement studies: Design, analysis, and meta-analysis. Cambridge University Press, Cambridge.

Rosenthal R. , Hall J. A. , DiMatteo M. R. , Rogers, R. L. , & Archer, D. (1979). Sensitivity to nonverbal communication: The PONS Test. Baltimore: Johns Hopkins Press.

Rosenthal, R., Hall, J. A., Dimatteo, M. R., Rogers, P. L., & Archer, D. (1979). Sensitivity to noverbal communication: The PONS test. Baltimore; Johns Hopkins University Press.

Schechter (1989). The undertreatment of pain in children: An overview. Pediatric Clinics of North America, 36, 781-793.

Schecter, Allen, & Hanson (1986). Status of pediatric pain control: A comparison of hospital analgesic usage in children and adults. Pediatrics, 77, 11-15.

Scott, C. S., Riggs, K. W., Ling, E. Fitzgerald, C., Hill, M., Solimano, A. Grunau, R. V. E., & Craig, K. D. (1997). Morphine pharmacokinetics and pain assessment in premature newborns. Unpublished Manuscript.

Shannon & Berde (1989). Pharmacologic management of pain in children and adolescents. Pain Clinics of North America, 36, 855-871.

Shapiro, C. R. (1991). Nurses' judgements of pain intensity in term and preterm newborns. Journal of pain and symptom management, 6, 148.

Scherer, K. R. , & Ekman, P. (1982). Methodological issues in studying nonverbal behavior. In R. Klaus, Scherer & P. Ekman (Eds.), Handbook of Methods in Behavior research. (pp. 1-43). Cambridge: Cambridge University Press.

Spear (1992). Anesthesia for premature and term infants Perioperative implications. The Journal of Pediatrics, 120, 165-176.

Stevens, B. J. & Johnston, C. C. (1994). Physiological responses of premature infants to a painful stimulus. Nursing Research, 43, 226-231.

Stevens, B. J., Johnston, C.C., & Horton, L. (1993). Multidimensional pain assessment in preterm infants: a pilot study. Journal of Obstetrical, Gynecological and Neonatal Nursing, 22, 531-541.

Stevens, B. J., Johnston, C. Petryshen, P. & Taddio, A. (in press). Premature infant pain profile: Development and initial validation. The Clinical Journal of Pain,.

Taddio, A. , Goldbach, M. , Ipp, M. , Stevens, B. & Koren, G. (1995). Effect of neonatal circumcision on pain responses during vaccination in boys. The Lancet, 345, 291-292.

Truog, R. & Anand, K. J. S. (1989). Management of pain in the postoperative neonate. Clinics in Perinatology, 16, 61-78.

Turk, D. C. & Melzak, R. (1992) The measurement of pain and the assessment of people experiencing pain. In D. C. Turk & R. Melzak (Eds.) Handbook of pain assessment. (pp 3 -14).

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

Walco, Cassidy, & Schechter (1994). Pain Hurt & harm: The ethics of pain control in infants and children. The New England Journal Of Medicine, 331, 541-544

Wasz-Hockert, Lind, Vuorenkoski, Partanen & Valanne (1968). The infant pain cry: A spectrographic and auditory analysis. Spastics Int. London.

Walker, H. M., & Lev, J. (1953). Statistical inference. New York: Holt.

Willer, J.C. (1985). Studies on pain: effects of morphine on a spinal nociception flexion reflex and related pain sensation in man. Brain Research, 187, 212-215.

Zeskind, P. S., & Marshal, T. R. (1988). The relation between variations in pitch and maternal perceptions of infant crying. Child Development, 59, 193-196.

APPENDIX A

Consent Forms
Course credit participants

Dear Participants;

We are conducting a study entitled 'Judging Pain in the Preterm Infant'. this project is the second part of a study which looks at the effectiveness of analgesic medication with preterm infants. The first part of the study revealed that some pain relief was obtained, however, we do not know if the pain reduction achieved was clinically significant. That is, we do not know if the pain reduction was large enough to be seen by observers.

The aims of this study are to determine: (1) how much of a decrease in pain expression is observable to untrained judges and (2) how large a change in pain expression represents a meaningful amount of pain reduction. In order to get this information , we will be asking you to perform two separate judgement tasks. In both cases, you will be asked to view video clips of the infants' responses to a routine blood sampling procedure. You will then be asked to rate the amount of pain the infants are experiencing. The judgement task will require about 2 hours of your time. All of the information you provide will be kept confidential and anonymous. Your questionnaire package will be identified by subject number and your responses will be seen only by the experimenters who are running the study.

If you have any questions or if you would like further information on the study, please feel free to ask the experimenter or contact Dr. Ken Craig, UBC Psychology Department (822-3948).

If you do participate in this study, you will receive 2 credit points. We hope that you will be able to participate in this study, but you are free to withdraw from the study now or at any time and doing so will not jeopardize your class standing in any way.

I acknowledge receipt of this form, and:

I consent ☐

I do not consent ☐

Signature _____

Date _____

Consent Forms (cont.)
paid participants

Dear Participants;

We are conducting a study entitled 'Judging Pain in the Preterm Infant'. this project is the second part of a study which looks at the effectiveness of analgesic medication with preterm infants. The first part of the study revealed that some pain relief was obtained, however, we do not know if the pain reduction achieved was clinically significant. That is, we do not know if the pain reduction was large enough to be seen by observers.

The aims of this study are to determine: (1) how much of a decrease in pain expression is observable to untrained judges and (2) how large a change in pain expression represents a meaningful amount of pain reduction. In order to get this information , we will be asking you to perform two separate judgement tasks. In both cases, you will be asked to view video clips of the infants' responses to a routine blood sampling procedure. You will then be asked to rate the amount of pain the infants are experiencing. The judgement task will require about 2 hours of your time. All of the information you provide will be kept confidential and anonymous. Your questionnaire package will be identified by subject number and your responses will be seen only by the experimenters who are running the study.

If you have any questions or if you would like further information on the study, please feel free to ask the experimenter or contact Dr. Ken Craig, UBC Psychology Department (822-3948).

If you do participate in this study, you will receive 20 dollars. We hope that you will be able to participate in this study, but you are free to withdraw from the study now or at any time and doing so will not jeopardize your class standing in any way.

I acknowledge receipt of this form, and:

I consent ☐

I do not consent ☐

Signature _____

APPENDIX B
Recruitment posters
subject credit participants

2 Subject Credits
for
Judging Pain in the Pre-term Infant

We are conducting a study as part of a larger project looking at the effectiveness of analgesic medication with preterm infants. Your task will be to watch videotapes of the infants faces and record your judgements of their pain on simple questionnaires.

The task should take about 2 hours over two sessions (1hour per session) and you will receive 1 subject credit for each session. Sessions may be booked consecutively, but we will ask you to take a break of at least a 15 minutes between.

Please leave your first name and phone number below or call 822-5280 (Dr. Craig's lab) and leave a message for Michelle indicating how you might be contacted.

First Name	Phone Number Day	Evening

Experiment 015-96
Dr. Ken Craig & Bruce McMurtry ph. 822-5280

Recruitment Posters (cont.)
paid participants

$$$
for
Judging Pain in the Pre-term Infant

We are conducting a study as part of a larger project looking at the effectiveness of analgesic medication with preterm infants. Your task will be to watch videotapes of the infants faces and record your judgements of their pain on simple questionnaires.

The task should take about 2 hours over two sessions (1hour per session) and you will receive 20 dollars for participating. Sessions may be booked consecutively, but we will ask you to take a break of at least a 15 minutes between.

Please leave your first name and phone number below or call 822-5280 (Dr. Craig's lab) and leave a message for Michelle indicating how you might be contacted.

First Name	Phone Number Day	Evening
__________	__________	__________
__________	__________	__________
__________	__________	__________
__________	__________	__________
__________	__________	__________
__________	__________	__________
__________	__________	__________
__________	__________	__________

APENDIX C

Breakdown of Judge Demographics

	Judge Empathic Concern		
	High (N=10)	Low (N=10)	All Judges N=(20)
Mean Empathic Concern Score (SD)	26.0 (1.63)	19.9 (2.33)	22.9 (3.71)
Age Mean (range)	22.1 (18-32)	20.7 (18-36)	21.4 (18-36)
Sex:			
Male	N = 1	N = 6	N=7
Female	N = 9	N = 4	N=13
Ethnic Identity (self described):			
Asian	N=4	N=5	N=9
Canadian	N=2	N=3	N=5
South/Central American	N=1	N=0	N=1
First Nations	N=1	N=0	N=1
Semitic-Canadian	N=0	N=1	N=1
Not answered	N=2	N=1	N=3
Experience Caring for an Infant:			
Have Children	N=0	N=0	N=0
Baby Sitting	N=4	N=5	N=9
Professional Caregiver	N=0	N=0	N=0
None	N=6	N=6	N=12
Faculty:			
Arts	N=4	N=2	N=6
Science	N=2	N=6	N=8
Fine Arts	N=1	N=0	N=1
Commerce	N=1	N=1	N=2
Unclassified	N=0	N=1	N=1
Graduate Studies	N=1	N=0	N=1
Marital Status:			
Married	N=0	N=0	N=0
Single	N=0	N=0	N=0
Divorced/Separated	N=0	N=0	N=0
Never Married	N=10	N=10	N=20

APPENDIX D

Components of the Neonatal Facial Coding, System (Grunau & Craig (1990).

Action	Description
Brow bulge	Bulging, creasing and vertical furrows above and between brows occurring as a result of the lowering and drawing together of the eyebrows
Eye squeeze	Identified by the squeezing or bulging of the eyelids. Bulging of the fatty pads about the infant's eyes pronounced
Naso-labial	Primarily manifested by the pulling upwards and furrow fold deepening of the naso-labial fold (a line or wrinkle which begins adjacent to the nostril wings and runs down and outwards beyond the lip comers)
Open lips	Any separation of the lips
Stretch mouth, vertical	Characterised by a tautness at the lip comers coupled with a pronounced downward pull on the jaw. Often stretch mouth is seen when an already wide open mouth is opened a fraction further by an extra pull at the jaw
Stretch mouth, horizontal	This appears as a distinct horizontal pull at the comers of the mouth
Lip purse	The lips appear as if an 'oo' sound is being pronounced
Taut tongue	Characterised by a raised, cupped tongue with sharp tense edges. The first occurrence of taut tongue is usually easy to see, often occurring with a wide open mouth. After this first occurrence, the mouth may close slightly. Taut tongue Is still scoreable on the basis of the still visible tongue edges
Chin quiver	An obvious high frequency up-down motion of the lower jaw
Tongue protrusion	Tongue visible between the lips extending beyond the mouth

APPENDIX E

Empathic Concern Scale
(Extracted from the Interpersonal Reactivity Index; Davis, 1980)

1. I often have tender, concerned feelings for people less fortunate than me.

0	1	2	3	4
does not describe me very well				describes me very well

*2. Sometimes I don't feel very sorry for other people when they are having problems.

0	1	2	3	4
does not describe me very well				describes me very well

3. When I see someone being taken advantage of, I feel kind of protective towards them.

0	1	2	3	4
does not describe me very well				describes me very well

*4. Other people's misfortunes do not usually disturb me a great deal.

0	1	2	3	4
does not describe me very well				describes me very well

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

0	1	2	3	4
does not describe me very well				describes me very well

6. I am often quite touched by things that I see happen.

0	1	2	3	4
does not describe me very well				describes me very well

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

0	1	2	3	4
does not describe me very well				describes me very well

Note * indicates reverse scored items.

APPENDIX F

Demographic Information Form

1. Your Gender: female male

2. Your Current Age: _____ (years)

3. Your Current Marital Status (circle one number):

1. Married
2. Divorced/ Separated
3. Remarried
4. Widowed
5. Never married
6. Other _____

4. Your Education (circle the one you last completed):

1. Graduate School/ Professional training
2. University graduate (4 year college)
3. Partial university (at least 1 year)
4. Trade School/Community College
5. High School graduate
6. Some high school (min. 10th grade)
7. Junior high school graduate
8. Less than 7th grade

5.a) Your Occupation (please describe):

_____

b) If you are a student, what is your faculty_____

6. Do you have children? yes no

7. Have you ever been responsible for the care of an infant? yes no
If yes, in what capacity?

_____

8. What languages can you speak fluently? _____

9. Which of the languages did you learn first? _____

10. Which language is most natural to you now? (i.e., in which language do you think?)

_____

11.a) In what country were you born?
____Canada _____Other (please specify: _____) ____don't know
b) If you were born outside North America, at what age did you first come to North America? age: ____
c) In what country was your mother born?
____Canada _____Other (please specify: _____) ____don't know

d) In what country was your father born?
____Canada _____Other (please specify: _____) ____don't know

12. How do you describe your own cultural or ethnic identity? (i.e. Canadian, Japanese, Cantonese, American, South-Asian, First Nations, French-Canadian, Indo-Canadian, Ukrainian-Canadian, etc.)

_____

13. What was the ethnic origin of the friends and peers you had as a child up to age 6?
    a) almost exclusively from my own cultural group
    b) mostly from my own cultural group
    c) about equally from my group and Anglo (English-speaking) or other groups
    d) mostly Anglos or from cultural groups other than my own
    e) almost exclusively Anglos or from cultural groups other than my own

14. With whom do you now associate in the community?
    a) almost exclusively from my own cultural group
    b) mostly from my own cultural group
    c) about equally from my group and Anglo (English-speaking) or other groups
    d) mostly Anglos or from cultural groups other than my own
    e) almost exclusively Anglos or from cultural groups other than my own

15. Do you participate in special occasions, holidays, traditions, etc. that are specific to your culture of ethnic origin?
    a) nearly all of them
    b) most of them
    c) some of them
    d) a few of them
    e) none at all

16. How much do you identify with each culture:
    a) Your culture of ethnic origin:
        i) very much
        ii) mostly
        iii) partially or somewhat
        iv) a little
        v) not at all
    b) Anglo-Canadian culture
        i) very much
        ii) mostly
        iii) partially or somewhat
        iv) a little
        v) not at all

APPENDIX G

Pain Rating Scales For Task One
(based on Gracely, Dubner & McGrath,1979)

Pain Judgement scales

After watching each videoclip, I would like you to judge the pain experience of the infant based on their facial expression. On the first scale I would like you to rate the intensity of the pain you think the infant is experiencing. On the second scale, rate the unpleasantness of the pain the baby is experiencing. Circle the words which best describe the intensity and unpleasantness of the pain. Please answer the final question by checking "yes" or "no":

Baby # _____

Intensity	Unpleasantness
A. Extremely intense	A. Very intolerable
B. Very intense	B. Intolerable
C. Intense	C. Very distressing
D. Strong	D. Slightly intolerable
E. Slightly intense	E. Very annoying
F. Barely strong	F. Distressing
G. Moderate	G. Very unpleasant
H. Mild	H. Slightly distressing
I. Very mild	I. Annoying
J. Weak	J. Unpleasant
K. Very weak	K. Slightly annoying
L. Faint	L. Slightly unpleasant
M. No sensation of pain	M. No discomfort

Does this baby require intervention to reduce their pain?

yes____ no ____

APPENDIX H

Relative magnitudes of Gracely pain descriptors

(Gracely, Dubner & McGrath, 1979)

Sensory Intensity		Unpleasantness	
Descriptor	Relative Magnitude	Descriptor	Relative Magnitude
Extremely intense	59.5	Very intolerable	44.8
Very intense	43.5	Intolerable	32.3
Intense	34.6	Very distressing	18.3
Strong	22.9	Slightly intolerable	13.6
Slightly intense	21.3	Very annoying	12.1
Barely strong	12.6	Distressing	11.4
Moderate	12.4	Very Unpleasant	10.7
Mild	5.5	Slightly distressing	6.2
Very Mild	3.9	Annoying	5.7
Weak	2.8	Unpleasant	5.6
Very Weak	2.3	Slightly annoying	3.5
Faint	1.1	Slightly unpleasant	2.8
No sensation of pain	0	No discomfort	0

Appendix I

Response form for task 2

Baby # 1

pain response (no morphine):

no pain — strong pain

pain response (morphine):

no pain — strong pain

was the pain reduction adequate ?

☐ yes ☐ no

Baby # 2

pain response (no morphine):

no pain — strong pain

pain response (morphine):

no pain — strong pain

was the pain reduction adequate ?

☐ yes ☐ no

APPENDIX J

Debriefing Form

Dear Participant:

Thank you for taking the time to participate in this study on judging pain in the pre-term infant. Pre-term infants often must undergo potentially painful procedures as part of their care, but there is a relative lack of knowledge about pain and its treatment in this population. The primary purpose of this study was to determine the degree to which observers judged morphine analgesia to be effective in reducing pre-term infants' reactions to a common medical procedure. In addition we are interested in how various characteristics of observers influence their judgements of pain in others.

The clips you saw included the infants' facial expressions once during quiet time and twice during a heel stick for blood testing, once with morphine analgesia and once without. All procedures were a necessary and normal part of the infants' medical care. The facial expressions of the infant in each clip had previously been objectively coded. Analyses indicated the heel stick procedure produced a significant increase in facial activity from quiet time and that morphine analgesia significantly reduced this pain expression. The part of the study you participated in was attempting to determine if these changes can be detected by observers, and if they are adequate.

The ratings you have just completed will be pooled with those of other untrained observers and analysed to determine if morphine analgesia was judged to significantly reduce the infants' pain expressions. In addition, the judgements will be compared to objective measures of the facial expressions. We will also be comparing the judgements made by untrained judges to those made by nurses. Finally, we will be analysing whether cultural background and/or reactivity to the experiences of others, influences sensitivity to facially expressed pain. All of the responses you provided will remain strictly confidential, identified only with a participant number.

If you are interested in learning more about pain and its treatment in neonates, the resources listed on the bottom of this page are a good place to begin. If you are interested in the results of the present study, or have any more questions about the study or your participation, please contact Dr. Ken Craig's lab at 822-5280. The study will be completed by the end of 1997, and we will have preliminary results by then.

Thanks again for your help with this project.

Bruce McMurtry, BA., UBC 822-5280
Ken Craig, Ph.D., Professor, UBC 822-3948

References

Carter, B. (1994) Child and Infant Pain: Principles of nursing care and management. London: Chapman & Hall.

Skevington, S. M. (1995). Psychology of pain. Rexdale: John Wiley & Sons LTD.

APENDIX K

RE-analyses of NFCS data for segments used in judgement study

Between Within, 3(condition) X 4(age group) ANOVA:

Effect	df	F	Significance of F
Age	3	357.53	$p < .01$
Condition	2	33.92	$p < .01$
Age X Condition	6	2.47	$p < .05$

Simple Effects of Age within each Condition:

Condition	df	F	Significance of F
Baseline (no /pain)	3	.72	$p > .05$
Untreated Pain (pain/no morphine)	3	3.50	$p < .05$
Treated Pain (pain/morphine)	3	5.67	$p < .01$

Post Hoc comparison of differences between age groups in NFCS (PCA weighted) in the treated pain and untreated pain conditions:

Pain/Morphine (treated pain)	Age gp3	Age gp4	Age gp2	Age gp1
	3.32(1.86)	2.82(1.67)	1.36(1.50)	1.28(1.25)

Pain/No Morphine (untreated pain)	Age gp3	Age gp4	Age gp1	Age gp2
	4.71(1.49)	4.46(1.88)	3.53(1.79)	2.66(2.28)

Note. Scores refer to M and (SD) Age gp1= (24-27wks PCA) ; Age gp2= (28-31wks PCA); Age gp3=(32-35wks PCA); Age gp4=(36-40wks PCA). Break in line represents significance at $p < .05$.