View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by NORA - Norwegian Open Research Archives

Assessment of Pain in Adults with Intellectual and Developmental Disabilities

Meir Lotan

Dissertation for the degree of doctor philosophiae (PhD)

University of Bergen, Norway

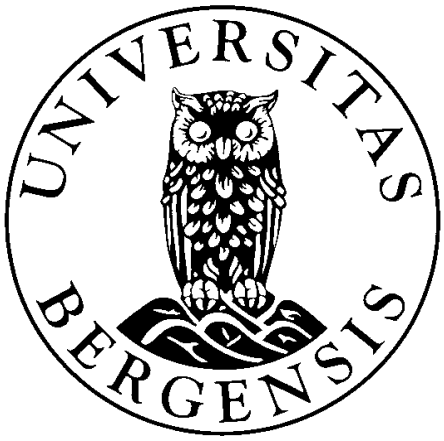
2009

UNIVERSITY OF BERGEN

Assessment of Pain in Adults with Intellectual and Developmental Disabilities

Development of a scale to assess pain behaviors

Meir Lotan



Department of Public Health and Primary Health Care

Section for Physiotherapy Science

University of Bergen

Elwyn residential setting, Jerusalem, Israel

Shalem Foundation, Israel

Table of contents

Page

Table of contents…….. III-VI

Acknowledgements……... VII

List of papers……. VIII

Abbreviations…… IX

Abstract……. XI-XIV

1. **INTRODUCTION**…… 1-28

1.1 INTELLECTUAL AND DEVELOPMENTAL DISABILITY ……. 1

1.1.1 Definitions…….……..... 1

1.1.2 Levels of IDD…….…… 2-3

1.1.3 Health status of individuals with IDD…….... 3-5

1.2 PAIN…….…… 5-9

1.2.1 Introduction …..…… 5

1.2.2 Types of pain…..……. 5-8

1.2.3 Pain in individuals with IDD…...…….. 8-9

1.3 PAIN ASSESSMENT…….. 9-12

1.3.1 Self report……..……. 10

1.3.2 Physiological pain assessment.……. 11

1.3.3 Behavioral pain assessment …….. 11-12

1.4 PAIN ASSESSMENT IN INDIVIDUALS WITH IDD…….. 12-22

1.4.1 The importance of pain assessment in individuals with IDD……. 12-13

1.4.2 The complexity of assessing pain in individuals with IDD…..…. 13-15

1.4.3 Existing pain scales …..……. 15-21

1.4.4 The use of proxy observer…...……. 21-22

1.5 CONSIDERATIONS IN DEVELOPING A PAIN ASSESSMENT SCALE…….....…….. 22-28

1.5.1 Reliability……... 23-24

1.5.2 Validity…… 25- 26

1.5.3 Additional measurement properties …..…..……... 27

1.5.4 Clinical considerations ……..…. 27 - 28

2. **THE CURRENT STUDY**…..……..…… 29-50

2.1 AIMS OF THE STUDY…….……... 29-30

2.2 REVIEW OF PAPERS……..……. 31-37

2.2.1 Paper I……...……. 31-32

2.2.2 Paper II……..……... 33-34

2.2.3 Paper III…….…….. 35-36

2.2.4 Paper IV……..……... 37

2.3 METHODS AND PARTICIPANTS……..……... 38-50

2.3.1 *Identifying promising pain assessment tools*……..…….. 38-39

2.3.1.1 *The Non-Communicating Children's Pain Checklist-Revised (NCCPC-R)*…… 38

2.3.1.2 *The Facial Action Coding System (FACS)*……..…. 38-39

2.3.2. *Participants*…… 39-41

2.3.3. *Data collection*…… 42-44

2.3.4. *Statistical analyses* ……..…….. 45-48

2.3.5. *Ethical issues and approval process*…… 49-50

3. **RESULTS ……..……..** 51-53

4. **DISCUSSION ……** 55-74

4.1 MAIN FINDINGS …….. 55-67

4.2 METHODOLOGICAL CONSIDERATIOS 67-72

4.2.1 *Internal validity*…… 67-68

4.2.2. *External validity*…….. 68-69

4.2.3. *Examining the quality of the NCAPC*……. 69-72

4.2.3.1 *Reproducibility (reliability)* 69-71

4.2.3.2 *Validity* 70-71

4.3 LIMITATIONS OF THE PRESENT INVESTIGATION…….. 73-74

5. **CONCLUSIONS……..** 75-76

5.1. PERSPECTIVES TOWARDS FUTURE RESEARCH…… 76-77

6. **REFERENCE LIST** …… 79-93

7. **PAPERS I- IV**……

I Defrin R, Lotan M, Pick CG. The evaluation of acute pain in individuals with cognitive impairment: A differential effect of the level of impairment. Pain 2006; 124(3):312-20.

II Lotan M, Ljunggren AE, Johnsen TB, Defrin R, Pick C.G, Strand LI. A modified version of the Non-Communicating Children Pain Checklist-Revised (NCCPC-R), adapted to adults with intellectual and developmental disabilities. Sensitivity to pain and internal consistency. Journal of Pain 2009;10(4):398-407.

III Lotan M, Moe-Nilssen R, Ljunggren AE , Strand LI. Reliability of the Non-Communicating Adult Pain Checklist (NCAPC), assessed by different groups of health workers. Research in Developmental Disabilities 2008;30:735-45.

IV Lotan M, Moe-Nilssen R Ljunggren AE, Strand LI. Measurement properties of the Non-Communicating Adult Pain Checklist (NCAPC): a pain scale for adults with Intellectual and Developmental Disabilities, scored in a clinical setting. Research in Developmental Disabilities (Accepted).

8. APPENDICES

APPENDIX I – Non-Communicating Adult's Pain Checklist (NCAPC).

APPENDIX II - Non-Communicating Children Pain Checklist PV (NCCPC-PV).

APPENDIX III – Approval document from the ethical committee at the University of Haifa (Hebrew).

APPENDIX IV - Approval document from the ethical committee at the University of Haifa (Translated).

APPENDIX V - Approval document from the head of the department for care for individuals with IDD, Ministry of social welfare (Hebrew).

APPENDIX VI - Approval document from the head of the department for care for individuals with IDD, Ministry of social welfare (Translated).

APPENDIX VII - Permission letter to use and rename the NCCPC.

APPENDIX VIII - Faces pain scale (Faces VAS).

APPENDIX IX- A comparative summary of all existing pain scales for individuals with IDD.

ACKNOWLEDGEMENTS

I would like to thank all the people who participated in this investigation and helped promote the knowledge and hopefully the management of pain for all individuals with intellectual and developmental disabilities, thereby reducing unnecessary suffering for these people. I would like to thank my devoted group of supervisors that directed me in all the right directions and corrected so many of my mistakes. You taught me so much yet at the same time it was an extreme pleasure to work with you all.

Special thanks goes to my main supervisor, Associate Professor Liv Inger Strand for the constant support and the never ending oceans of patience, for your smart yet kind and gentle guidance. I was privileged to be under the supervision of someone that endlessly continues to amend and correct, until the work is as close to perfection as it gets. From the bottom of my heart I would like to thank Professor Elisabeth Ljunggren that invested ever lasting resources of energy and time in this work, with her knowledge and her all seeing eyes that detected even the smallest of mistakes.

Warm thanks go to Professor Rolf Moe-Nilssen for his amazing ability to unveil the most difficult ideas and abstract concepts into clear and simple understandings.

Thank you Associate Professor Tom Backer Johnsen for your intelligent support.

It has been an honor and a pleasure to have met you all – thank you.

I would like to thank my amazing partner Dr. Cochavit Elefant who is always there for me in poorer in reach, and in sickness and health and forever.

Thanks are due to the Shalem Foundation that went out of their way and partially financed this project despite the fact that it was connected with an overseas university.

Bergen, May, 2009

LIST OF PAPERS

I. Defrin R, Lotan M, Pick CG. The evaluation of acute pain in individuals with cognitive impairment: A differential effect of the level of impairment. Pain 2006; 124(3):312-20.

II. Lotan M, Ljunggren AE, Johnsen TB, Defrin R, Pick C.G, Strand LI. A modified version of the Non-Communicating Children Pain Checklist-Revised (NCCPC-R), adapted to adults with intellectual and developmental disabilities. Sensitivity to pain and internal consistency. Journal of Pain 2009;10(4):398-407.

III. Lotan M, Moe-Nilssen R, Ljunggren AE , Strand LI. Reliability of the Non-Communicating Adult Pain Checklist (NCAPC), assessed by different groups of health workers. Research in Developmental Disabilities 2008;30(4):735-45.

IV. Lotan M, Moe-Nilssen R Ljunggren AE, Strand LI. Measurement properties of the Non-Communicating Adult Pain Checklist (NCAPC): a pain scale for adults with Intellectual and Developmental Disabilities, scored in a clinical setting. Research in Developmental Disabilities (Accepted).

ABBREVIATIONS

AAMR	American Association of Mental retardation (changed to AAIDD)
AAIDD	American Association for Intellectual and Developmental Disability
AUC	Area Under the Curve
CI	Cognitive Impairment
CI	Confidence Interval
DD	Developmental Disability
ESPCP	The Evaluation Scale for Pain in Cerebral Palsy
FACS	Facial Action Coding System
FDR	False Discovery Rate
GFI	Goodness of Fit Index
IASSID	International Association for the Scientific Study of Intellectual Disabilities
ICC	Intra Class Correlation
ID	Intellectual Disabilities
IDD	Intellectual and Developmental Disabilities
INRS	Individualized Numeric Rating Scale
IQ	Intelligence Quotient
MOBID	Mobilization-Observation-Behavior-Intensity-Dementia Pain Scale
NCCPC	Non Communicating Children's Pain Checklist
NCCPC-R	Non Communicating Children's Pain Checklist-Revised
NCAPC	Non Communicating Adults' Pain Checklist
NI	Neurological Impairment
PADS	Pain and Discomfort Scale
PICIC	Pain Indicator for Communicatively Impaired Children
PPP	Pediatric Pain Profile
RCT	Randomized Controlled Trial
RMSEA	Root Mean Square Error of Approximation
ROC	Receiver Operation Curve
SD	Standard Deviation
SDD	Smallest Detectable Difference
SEM	Structural Equation Modeling
SEM	Standard Error of Measurement
SRM	Standardized Response Mean
S_w	Within subject standard deviation

ABSTRACT

Background

Pain measurement is a challenging task in most populations, but individuals with intellectual and developmental disabilities (IDD) display specific barriers to adequate pain evaluation, since they cannot give valid self-reports. Despite enhanced interest in manifestations of pain in adults with IDD in recent years, the characteristics of pain behavior in this group have scarcely been examined.

Aims

The aim of this thesis was to develop a clinical tool for assessing pain behaviors in adults with intellectual and developmental disabilities (IDD), to test the scale in respect to reliability and validity, and to test its clinical applicability in a painful situation with the research population.

Population

A total of 266 adults (mean age: 42.2 years) with different levels of IDD, living in community and residential centers were included in the study.

Method

In **Paper I** the Non-Communicating Children's Pain Checklist (NCCPC) and the Facial Action Coding System (FACS)'s ability to capture pain behavior in adults with IDD was examined. All participants were videotaped before and during an annual influenza vaccination, and scored using the NCCPC and FACS on both occasions. In **Paper II** based on scores from the video uptakes, the sensitivity to pain of each test item of the NCCPC (total of 27 items) was examined by Signed rank test,

and contribution of each item to internal consistency was examined by Cronbach's alpha. Sensitivity to change of the total scale by Standardized Response Mean (SRM) was evaluated. Thirteen items were excluded from the original NCCPC scale while four new items were added, making a modified scale named the Non-Communicating Adults' Pain Checklist (NCAPC). The internal consistency and sensitivity of the scale was reexamined after the reduction of items.

In **Paper III** intra- and interrater reliability of the NCAPC were investigated on video vignettes. Intrarater reliability was evaluated by the first author on a group of 50 randomly selected individuals. Interrater reliability was investigated in two stages. In the initial step different groups of health care workers (caregivers, nurses, case managers, and therapists), each including five raters, viewed a sample of 12 adult participants with IDD (3 at each level of IDD), who were extracted from the population sample. In the second stage 3 participants from each of the groups showing high interrater reliability (caregivers and therapists) evaluated interrater reliability in a randomly selected group of 40 individuals**.**

In **paper IV** the NCAPC's was examined in clinical settings for internal consistency, validity and clinical usability. To achieve these aims 58 adults at all levels of IDD, receiving dental hygiene treatment, were observed for pain behavior, before and during dental hygienist treatment, using the NCAPC. The results were compared with scores of the same participants during an influenza injection.

Results

The results from **paper I** suggested that The NCCPC was superior to the FACS in capturing pain behavior in adults with IDD and was sensitive to pain

behaviors at all levels of IDD, and development of a new scale was continued with NCCPC.

In **paper II** The scale was named the Non-Communicating Adults Pain Checklist-Revised (NCAPC). All items remaining in the modified scale were found to show Sensitivity to pain ($P<0.05$) and high internal consistency ($\alpha=0.773$) was demonstrated. Large sensitivity to pain at all levels of IDD was shown (SRM 1.20-2.07). The NCAPC was found to demonstrate better measurement properties than the NCCPC-R in the target population.

In **paper III** intrarater reliability was found at 0.94. Interrater reliability was very high in caregivers, physical- and occupational therapists, and was found at 0.91 and 0.92, correspondingly.

Results of **paper IV** show that the NCAPC was affective in assessing pain behaviors in a clinical setting. The scale showed satisfactory internal consistency, was able to differentiate between pain and non-pain situations and different pain experiences (influenza injection, dental hygienist treatment).

Discussion

Findings from **Paper I** showed that the NCCPC should be used as a basis to further development of a pain assessment tool for adults with IDD. Other findings suggested that some pain behaviors were commonly observed in adults, but not sufficiently captured by the NCCPC. Therefore there seemed to be a need to adapt the NCCPC to the adult IDD population. In **paper II** it was established that the NCAPC was showing better psychometric properties than NCCPC in adults with IDD. In **paper III** we concluded that the NCAPC have been found to hold high intra- as well as inter-rater reliability values. In **paper IV** The NCAPC was found a valid and

reliable tool, and the authors concluded that it can be used clinically to detect acute pain behaviors of individuals at all levels of IDD at different settings by different health care workers, and during various pain experiences.

Summation

The procedure that was performed in the present thesis has led to the construction of a pain behavior evaluation scale for adults with intellectual disability, showing good measurement properties that allow the use of this tool in clinical settings.

1. INTRODUCTION

1.1. Intellectual and developmental disability

1.1.1. Definitions

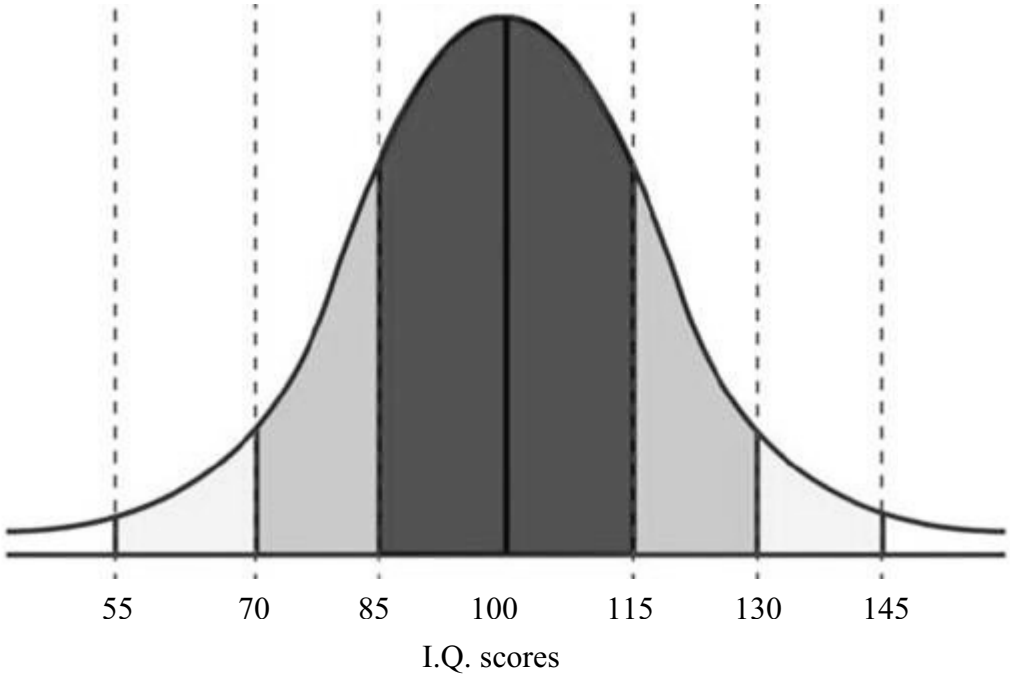
Developmental disabilities (DD) are defined as severe chronic mental or physical disabilities that are manifested before a person reaches 18 years of age (Developmental Disabilities Assistance and Bill of Rights Act, 2000). These disabilities are likely to continue indefinitely and result in substantial functional limitations in three or more of the following areas: self-care, receptive and expressive language, learning, mobility, self-direction, capacity for independent living, and economic self-sufficiency.

Individuals with intellectual and developmental disability (IDD) constitute a subset of persons with DD who have below-average general intellectual functioning, as measured through standardized general aptitude evaluation tools, such as the Wechsler Intelligence Scales or the Stanford-Binet Intelligence Scales. IDD can be identified if it is accompanied by two or more deficits in adaptive behavior used for everyday living (e.g. communication, self-care, home living, social skills, community use, self-direction, health and safety, functional, academics, leisure, and work), as determined by a structured evaluation tool such as the Vineland Adaptive Behavior Scales (Developmental Disabilities Assistance and Bill of Rights Act, 2000; Individuals with Disabilities Education Act, 1992). Individuals with IDD comprise approximately two and a half percent of the general population (Krahn et al., 2006) and are divided into four subgroups according to their level of IDD.

1.1.2. Levels of IDD

In recent years, a significant trend has emerged in regard to the classification of intellectual disability. Traditionally, classification systems have revolved primarily around the range of Intelligence Quotient (IQ) scores (see figure 1) presented by people at two or more standard deviations below the mean of the general population (IQ 100± S.D. 15). The most common such classification scheme involves grouping people into one of four subgroups based on their IQ scores: mild (69 to 55), moderate (54 to 40), severe (39 to 25), and profound (below 25).

Figure 1 – Distribution of intelligence quotations in the general population



In 2002, the classification manual of the American Association for Intellectual and Developmental Disability (AAIDD, recently changed from the American Association for Mental Retardation - AAMR) suggested four levels of support intensities: intermittent, limited, extensive, and pervasive (Luckasson et al., 2002). These support intensities are not correlated in line with the IQ levels of mental retardation (mild, moderate, severe, and profound). Although traditional classification

systems are still in use, there is no consensus as to the best way to classify people within this population or as to whether a change in classification is useful or even necessary. Therefore, in the present research, we divided and related to the participants with IDD in subgroups according to their intellectual abilities.

In addition, there is some controversy regarding the term used to describe this population. Although the term 'mental retardation' is still valuable for diagnostic purposes, it is sometimes used synonymously with several other terms in the literature, such as 'cognitive impairment' (CI), 'neurological impairment' (NI), 'developmental disability' (DD), and 'intellectual and developmental disability' (IDD). The term IDD is preferred for several reasons and is employed by the World Health Organization (WHO, 2002) and the International Association for the Scientific Study of Intellectual Disabilities (IASSID), which is a sister organization of the AAIDD. Furthermore, the term is used in a number of journal names, all with clear links to 'mental retardation.' Thus, the term IDD will be used to define the participants in the present research.

1.1.3. Health status of individuals with IDD

Persons with IDD often have multiple and sometimes complicated medical problems (Prater & Zylstra, 2006). In fact, a number of studies have documented substantially higher rates of both chronic and acute medical conditions in people with IDD as compared to the general population (Cooper, 1998). Miniham and Dean (1990) and Miniham et al. (1993) reported that 99% of individuals with IDD in a state institution had at least one chronic medical condition requiring regular follow-up (e.g., cardiac conditions, diabetes, ulcers, chronic otitis media, recurrent pneumonia, and progressive renal failure). Janicki et al. (1999) found that 49% of people with

IDD admitted to hospitals had a visual impairment, 27% had a hearing impairment, and over 50% were obese. Likewise, Beange et al. (1995) found that people with IDD had increased cardiovascular risk factors and higher rates of medical consultations and hospital admissions than the general population, with 4.5 medical disorders per person on average.

The medical problems diagnosed in this population are diverse, ranging from limb contractures and scoliosis (Berven & Bradford, 2002; Thacker et al., 2002) to spasticity (Pfister et al., 2003) and osteoporosis, particularly among non-weight-bearing patients (Henderson et al., 2002; Tyler, Snyder, & Zyzanski, 2000). Persons with IDD often suffer from a host of behavioral and psychiatric problems as well (Prater & Zylstra, 2006). These findings highlight the need for intense and specifically tailored medical coverage for individuals with IDD. However, as a minority group largely lacking empowerment and advocacy, they are constantly challenged by unmet health care needs. Several investigations (Beange et al., 1995; Fisher, 2004; Kerr et al., 1996; Whitfield et al., 1996) have suggested that the health mismanagement of this population has a severe impact on mortality (Bittles et al., 2002; Durvasula & Beange, 2001; Hollins et al., 1998); morbidity (Beange et al., 1995; Janicki et al., 1999); and quality of life (Hensel et al., 2002).

The present research will address the issue of pain in IDD, and we assume that there is a similar disparity in pain assessment and management for this population as there is in other health care issues. Indeed, recent findings support our assumption and show that the IDD population receives reduced levels of post-surgery analgesics as compared to control groups (Gauthier, Finley, & McGrath, 1998). Moreover, mistreatment of individuals with IDD has been found to delay diagnosis and

management of painful medical conditions, causing setbacks in hospitalization and even death (Carter & Jancar, 1984; Mata, 1960; Roy & Simon, 1987).

1.2. Pain

1.2.1. Introduction

Pain is referred to by the International Association for the Study of Pain (IASP) as "an unpleasant sensory or emotional experience associated with actual or potential tissue damage or described in terms of such damage" (Latham & Davis, 1994). Pain can have a negative effect on the individual's functional ability, mobility, emotional status, ability to work, interpersonal relationships, and social activities, leading to increased use of health care services and an accompanying increase in health care costs (Merskey & Bogduk, 1992). This situation is worsened when the individual suffering from pain cannot voice discomfort, as in the very young, the very old, and persons with IDD. Unequal access to pain relief and failure to treat pain is viewed as poor medicine, unethical practice, and an abrogation of fundamental human rights (Brennan, et al., 2007).

1.2.2. Types of pain

There are different types of pain, including chronic pain, pain associated with terminal illness, acute pain, and procedure-related pain (American Academy of Pediatrics, 2001).

Chronic pain

Chronic pain is defined as pain lasting longer than the expected healing time (3-6 months) (Mersky & Bogduk, 1994). Research into chronic pain

management for people with IDD is currently limited. The evidence base consists predominantly of studies assessing chronic pain through behavioral techniques for children (Allen, et al., 2002) and multidisciplinary programs for adults (Guzman et al., 2002).

Pain associated with terminal illness

Terminal illness is defined as an illness that cannot be cured by means of current medical technology and generally leads to death (Turk & Feldman, 1992). At this stage in life, the physician implements medical methods, including pain relief medication, with the intent of prolonging life as well as improving the patient's quality of life and well-being (Bonica, 1979). The complexity of managing the pain associated with terminal illness stems from its interaction with other common symptoms of terminal illness, such as fatigue, weakness, dyspnea, nausea, and constipation, as well as anxiety, fear, and sleep deprivation. This type of pain will not be the issue of the present investigation.

Acute pain

Acute pain is defined as pain that is continually changing and transient. It is accompanied by a high level of emotional and autonomic nervous system arousal and is usually associated with tissue pathology or surgery (American Pain Society, 2003; Melzack & Wall, 2003). Acute pain assessment is complex and requires consideration of individual pain perception, psychological and developmental factors, and potential severity of the specific types of pain experienced (McGrath & Brigham, 1992).

Procedural pain

Procedural pain is defined as pain caused by different medical procedures or examinations (Halimaa, 2003). The complexity of factors involved in acute pain perception and reaction (McGrath & Brigham, 1992) have an even greater impact in the case of procedural pain. Such factors frequently induces anxiety and distress (Carrougher et al., 2006) and involve a degree of anticipation that can compound patients' distress, especially if they have had "bad" past experiences (Von Baeyer et al., 2004). Individuals with such chronic diseases as IDD are likely to undergo more medical tests and treatments than individuals without unique medical conditions. Therefore, it is reasonable to assume that their exposure to procedural pain will also be higher over the course of their lifetime.

Influenza injection, which is the focus of the present investigation, belongs to a group of pain stimuli called needle pain. This type of pain is the most common type of procedural pain and has been found to cause considerable distress in children (Fradet et al., 1990). It was found that more than 50% of children and adolescents who undergo venipuncture for routine blood sampling, experience moderate to severe levels of distress or pain ((Fradet et al., 1990). The evidence shows that fear of a painful procedure causes anticipatory anxiety, which in turn increases the likelihood of experiencing more pain and distress during the actual procedures (Blount et al., 2006).

Yet, the findings on reactions to procedural pain at different ages are unclear. On the one hand, the fear and pain experienced during medical procedures in childhood are found to be predictive of fear and pain during medical procedures and avoidance of medical care during young adulthood (Pate et al., 1996). Other emotional factors, such as elevated anxiety, anger, and low mood, have been found to confound the perception of pain during medical procedures (McGrath, 1994) and may

also render subsequent procedures and pain management more difficult (Frank et al., 1995). On the other hand, younger children are typically found to report greater levels of pain intensity and unpleasantness from needles than older children (Goodenough et al., 1997; Goodenough et al., 1999). This may indicate a trend towards a reduction in the impact of psychological elements with growing age and exposure to routine procedures, thereby minimizing their effect in the adult population. However, no data could be found to shed light on the pain reaction during medical procedures among adults with IDD, and further in-depth research is needed to fully understand this complex phenomenon in this unique population.

1.2.3. Pain in individuals with IDD

The prevalence of pain in the IDD population is unclear, mainly due to communication problems that make the recognition of pain difficult (Reid et al., 2003). People with IDD are vulnerable to the same range of pain-inflicting procedures as the non-IDD adult population, but in addition they are also vulnerable to experiencing pain from falls, leg braces, and ill-fitting wheelchairs (Regnard et al., 2003). A study investigating the frequency, duration, intensity, and location of pain, as well as the interference of pain with activities, in adults with cerebral palsy (CP) and IDD, found that pain was a significant problem for the majority of participants (Schwartz et al., 1999). Of the 93 participants (with an average age of 38), the majority had quadriplegia (84%)[1] and were non-ambulatory (94%). One or more areas of chronic pain (minimum of three months' duration) was reported by 67% of the participants, and 53% experienced moderate to severe pain on an almost daily basis. Lower-extremity pain (66%) and back pain (63%) were the most common complaints.

[1] CP affecting all four limbs

The duration of pain ranged from a mean of 7.5 years for upper-extremity pain to a mean of 20 years for hip or buttock pain. Likewise, McGrath et al. (2000) investigated the pain experience of 64 children with IDD and found that they suffered pain on a regular basis, with 83% suffering constant pain at a level higher than 3 on a 10-point scale.

Thus, the accumulating evidence suggests that individuals with IDD suffer from more pain incidents than the general population and can be considered as a population at risk in regards to pain. Most researchers recommend that additional research is needed to carefully examine how pain can be better managed in people with IDD and multiple disabilities (Schwartz et al., 1999). However, it is clear that better pain management should start with proper pain evaluation and that it is essential for the clinician to use reliable evaluation tools to initiate the pain assessment and intervention processes.

1.3. PAIN ASSESSMENT

Pain is a subjective and multidimensional phenomenon (Abu-Saad, 2000). It can be assessed using a variety of modalities, including self-report, behavioral observation, or physiological measures, depending on the individual client and his or her communication capabilities. The adequacy of pain assessment through different modalities will be presented and discussed in the next paragraph within the context of the research population.

1.3.1. Self-report

Given that pain is a subjective experience, 'self-report' is usually considered to be the criterion standard or the "gold standard" in pain assessment (McCaffery & Beebe, 1989). Therefore it is not surprising that 90% of research projects published in PAIN journal were using self reports' based measures (Craig, 1989). There are many psychometric instruments available that translate the subjective experiences of patients into meaningful data which can be used to assist health care providers with pain diagnosis and treatment. Simple pain assessment tools, such as the Visual Analogue Scale (VAS), the Numeric Rating Scale (NRS), and the Verbal Descriptor Scale (VDS), are uni-dimensional and refer only to intensity of pain. There are also multidimensional tools, such as the McGill Pain Questionnaire (MPQ; Melzack, 1975) and the Multidimensional Pain Inventory (MPI; Kerns et al., 1985).

Yet, self-report scales can only be used in individuals old enough or cognitively competent to provide valid information regarding location, quality, intensity, and tolerability of the painful experience (American Academy of Pediatrics & American Pain Society, 2001; Johnston, 1998). Even individuals with mild or moderate levels of IDD were found unable to submit reliable self-reports regarding pain experience (Abu-Saad, 2000; Devies & Evans, 2001; Fanurik et al., 1998; Hadden & Von Baeyer, 2002). Therefore, such instruments may be inadequate and ineffective for individuals who use non-conventional forms of communication or who lack the cognitive sophistication to convert their internal experiences into a conventional, standardized expressed language.

1.3.2. Physiological pain assessment

There are a number of physiological measures of pain in use, including vagal tone (Gunnar et al., 1995), heart rate (Cohen et al., 1999), blood pressure (Marchette et al., 1991), salivary amylase activity (Yamaguchi et al., 2006), and intracranial pressure (Stevens & Johnson, 1994). Although physiological measures may be viewed as free of response bias and therefore more conducive to objectivity, no single physiological index has been shown to be ideal and specific enough for measuring pain. In fact, many physiological measures vary not only in accordance with the level of pain, but also in accordance with emotional states, temperature in the environment, body movement, and other extraneous factors. Furthermore, some of the measures are invasive and therefore introduce discomfort that might further exacerbate the distress and pain experience.

Finally, physiological instruments can be impractical in terms of the time and costs associated with their use, especially when considering the existing conditions in institutions for individuals with IDD. The American Academy of Pediatrics and the American Pain Society (2001) have stated that physiological measures should generally be conceived as measures that reflect stress reactions during acute pain and are usually only tenuously correlated with self-reports of pain. In light of these limitations, physiological measures were not considered for use in the present investigation.

1.3.3. Behavioral pain assessment

Individuals with cognitive and verbal deficits may be unable to describe their feelings of pain or physical discomfort in a conventional manner (Abu Saad, 2000), thus rendering valid self-report as infeasible. In such cases, observation of behavior

can be used as an acceptable alternative (McGrath, 1998; McGrath et al., 1985; Stevens, 1998). Behavioral indicators, such as facial expressions, crying and body movements, are used to estimate the presence and intensity of pain in nonverbal or preverbal children (Lawrence et al., 1993; Hunt et al., 2004), as well as in elderly persons with cognitive impairment (Husebo et al., 2007; Kaasalainen, 2007). Behavioral pain measures have been successfully used in the past to assess individuals with IDD (Hadjistavropoulos et al., 2001) and were therefore chosen as the preferred method of pain assessment for the present investigation.

1.4. PAIN ASSESSMENT IN INDIVIDUALS WITH IDD

1.4.1. The importance of pain assessment for individuals with IDD

Assessing pain in individuals with IDD is a challenging task and can become extremely difficult at the levels of severe and profound IDD, the ability to verbally communicate pain experience being severely compromised (Lachapelle et al., 1999). Without objective assessment, pain can be misinterpreted or underestimated, which might lead to inadequate management and undermine quality of life (Malviya et al., 2001).

Very few studies on pain in individuals with IDD have been published. Available findings suggest that pain in people with severe intellectual disability is common, yet rarely actively treated (Stallard et al., 2001). Studies in this field also indicate that people with IDD have 2.5 times more health problems than people without IDD (Van Schrojenstein et al., 2000). Individuals with severe or profound levels of IDD are more likely to have additional disabling conditions or multiple complex medical problems coupled with communication difficulties. Such medical problems, whether directly or indirectly linked to the disability, often necessitate

painful procedures, including physical therapy treatments and various medical interventions. Recent data reveals that “sick days” in this population were associated with higher levels of pain and discomfort than “well days” (Carr et al., 2007) and that people with severe cognitive impairments and low communication abilities are likely to experience the most pain over time (Breau et al., 2003).

The current situation puts individuals with IDD at a constant impediment to their quality of life, and therefore there is an urgent need to develop proper pain measures for this population. Yet, there are some objective difficulties in assessing pain in this population.

1.4.2. The complexity of assessing pain in individuals with IDD

Given the constant hindrance of pain to quality of life among individuals with IDD, there is an urgent need to develop a proper pain assessment tool for this population. However, the scientific world has lagged behind when it comes to pain assessment in individuals with IDD, and there are several reasons for this situation. First, many individuals with IDD have neurological problems that may affect their ability to comprehend and effectively communicate pain, thus complicating evaluation of the qualitative and quantitative aspects of their pain experience (Oberlander et al., 1999). Typical cognitive difficulties among this population involve abstract thinking and spatial orientation. Therefore, individuals with IDD may be unable to give valid reports of the features of their pain sensation, such as location, intensity, or quality of their pain. They may not be able to respond to questions about their pain or they may respond in a way that is not meaningful to caregivers (Breau et al., 2004). These circumstances make pain measurement in these patients highly difficult or in some cases impossible (Mafrica et al., 2006). Thus, due to this reduced

ability to verbally communicate pain, the gold standard of pain assessment, namely self-report, cannot be used with this population.

Second, individuals with IDD often have multiple handicaps and form an extremely heterogeneous group in terms of functional and behavioral repertoires. Functional limitations, such as paralysis and inability to move, may also mask expressions of pain (McGrath et al., 1998). To further complicate the issue of unclear communicative signals, challenging behaviors such as aggression, self-injury, and tantrums can be observed in this population (Carr et al., 2007). Such behaviors have been connected with painful medical problems (de Lissovoy, 1962; Hart et al., 1984), but can also mask pain in individuals with IDD (Clements, 1992). This makes it difficult to ascertain whether the behavior is attributable to pain or another source of distress or whether it is simply part of the individual's regular aberrant behavior.

Third, behavioral indicators of pain in the general population, such as facial grimaces, groaning, or altered sleep patterns (Bodfish et al., 2001), may well appear in individuals with IDD at times when they are not in pain (McGrath et al., 1998). It is therefore not surprising that such behaviors are attributed to the intellectual level of the individual rather than to pain (Mason & Scior, 2004), probably resulting in under-diagnosing of pain.

Finally, assessing and managing pain in people with IDD can be complicated by the effects of medication (Turk & Melzack, 2001), as well as the lack of appropriate pain assessment tools. Despite the increased research attention focused on expressive behavior related to pain in individuals with IDD (Carter et al., 2002; Donoven, 2002; Fanurik et al., 1999; Hadden & von Baeyer, 2002; Oberlander & O'Donnell, 2001; Stallard et al., 2001; Stallard et al., 2002a), research on this topic is

still scarce and there are but few pain assessment scales available for use in this specific population.

1.4.3. Existing pain scales

Several scales for pain assessment in individuals with IDD have been developed, the majority over recent years mostly for the pediatric population. One scale was developed for the general population, but has been used for individuals with IDD in the past. The following scales are ordered chronologically and their main features are summarized in appendix 8.7.

1) *The Facial Action Coding System* (*FACS*; Ekman & Friesen, 1978): The FACS is a list of facial actions (action units – AUs) based on movements of specific muscles or groups of muscles in the face. FACS was repeatedly found to be highly reliable by Craig and associates (Craig et al., 1988, 1991, 1992, 2002), as well as by other researchers (LeResche & Dworkin, 1984; Prkachin et al., 1994). This scale has been used for pain evaluation among adults with cognitive impairment (CI) due to dementia (Hadjstavropoulos et al., 1997; Hurley, Volier, & Hanrahan, 1992) and among individuals with IDD. The FACS was found suitable for detecting pain behaviors in individuals with mild to moderate levels of IDD undergoing influenza injection (Lachapelle et al., 1999).

2) *The Evaluation Scale for Pain in Cerebral Palsy* (*ESPCP*; Giusiano et al., 1995): The ESPCP consists of 22 items of pain behaviors derived from physicians' reports of cues considered to be indicative of pain during medical examination. The items included various facial expressions: crying, movements and posture (increase in muscular tone and/or involuntary movements, analgesic postures); protective reactions (movement towards painful areas), and social behaviors (e.g. reduced

interest in surroundings). Although there appears to be a common set of pain behaviors in people with cerebral palsy and severe intellectual disabilities, the importance of the different items in determining pain is dependent on the individual's level of development.

Using the ESPCP, Collignon et al. (1997) developed a 10-item observational scale to evaluate pain and facilitate therapeutic decision-making in children with severe handicaps and adults with cerebral palsy. Collignon and Giusiano (2001) then further developed the tool to better fit an adolescent population with IDD. These researchers investigated pain behaviors in 100 individuals, ranging in age from 2 to 33 years (mean 16 years), with multiple physical disabilities and profound IDD and without speech or any means of communication ability through symbols. Pain could only be detected by observing global behavioral changes, rather than by the presence of a single sign. In addition, each combination of disabilities appeared to evoke a specific set of behaviors. For instance, behaviors associated with the voluntary protection of painful areas were more likely to occur in individuals with a lesser degree of motor impairment. This tool was not further investigated for psychometric properties.

3) *The Non-Communicating Children's Pain Checklist* (*NCCPC*; Breau et al., 2000): The collection of pain items for this scale was initiated by McGrath and associates (McGrath et al., 1998). This group of researchers interviewed twenty parents or caregivers of cognitively impaired children, ranging in age from 6 to 29 years, regarding cues they considered to be indicative of pain in their children. The interviews included instances of short, sharp pain, such as needle pain, as well as longer-lasting pain, such as headache or injury. A list of 31 cues was elicited. While specific behaviors often differed from one child to another, classes of behaviors

(vocal, eating/sleeping, social/personality, facial expressions, body and limbs activity, and physiological reactions) were common to almost all children.

The NCCPC was developed from this initial study (Breau et al., 2000). It was comprised of 30 items and was to be tested in a home setting. Parents and caregivers assessed whether the pain cues were 'present' or 'absent' in four situations: acute pain, long-term pain, a non-painful but distressing situation, and a non-painful, calm situation. On the average, more than four times as many pain cues were present in painful situations than in calm (no-pain) situations. The total number of present cues did not differ between painful and distressed states, but scores for the 'eating/sleeping' and 'body/limb' subscales were higher during acute pain than during distress.

A second version of the NCCPC checklist, the NCCPC-PV (PV = Postoperative Version), was evaluated in a postoperative setting (Breau et al., 2002). In this study, items related to eating and sleeping were omitted and each of the remaining items was scored on a four-point ordinal scale according to frequency of occurrence. Twenty-four children, ranging in age from 3 to 19 years, were each observed by one of their caregivers and one of the researchers for 10 minutes both before and after surgery. When available, nurses also provided their assessments. Each observer completed the NCCPC-PV independently in addition to giving a global rating of the intensity of the child's pain using a Visual Analogue Scale. The NCCPC-PV was found to show very high internal consistency (Cronbach's alpha=0.91) and good interrater reliability (ICC 0.78 to 0.82). A moderate correlation (from 0.39 to 0.53) was observed preoperatively between scores on the NCCPC-PV and global assessments of the child's pain through the VAS. A score of 11 on the NCCPC-PV

provided 0.88% sensitivity and 0.81% specificity for classifying children who were rated at a moderate-severe level of pain on a verbal rating scale (VRS).

A third revised version of this scale, the NCCPC-R (R= revised), used ordinal ratings according to frequency of occurrence as above, but this time included the items related to eating and sleeping. This version was evaluated in home settings (Breau et al., 2002b). Using the NCCPC-R, 55 caregivers of 71 children with severe cognitive impairments, ranging in age from 3 to 18 years, conducted observations of their children during a time of pain and a time without pain. The NCCPC-R was found to have high internal consistency (Cronbach's alpha=0.93), as well as a moderate correlation with the pain intensity ratings provided by caregivers (Pearson's r=0.46). Sensitivity (0.84) and specificity (0.77) for pain were optimized at a cut-off point of 7 out of a possible total score of 90.

4) *The Pain Indicator for Communicatively Impaired Children* (*PICIC*; Stallard et al., 2002b): The PICIC uses six core items to assess the expression of chronic pain in non-communicative children with significant IDD. A significant relationship was demonstrated between five of the six core items and the presence and severity of pain (Stallard et al., 2002b). However, further research is needed before the PICIC can be established as a tool holding proper psychometric values.

5) *The Pediatric Pain Profile* (*PPP*; Hunt et al., 2004): The PPP is a 20-item behavior rating scale designed to assess pain in children with severe neurological and cognitive disability. The validity and reliability of the scale was assessed in 140 children, ranging in age from 1 to 18 years, who were unable to communicate through speech or augmentative communication. Parents used the PPP to retrospectively rate their child's behavior when 'at their best' and when in pain. Children were found to

display significantly higher scores when in pain than in a non-pain situation, and their scores increased in line with global evaluations of pain.

In order to assess interrater reliability, two raters concurrently observed and individually rated each child's behavior. Interrater reliability by ICC values was found to range between 0.74 and 0.89. In order to assess the construct validity and responsiveness of the scale, the behavior of 41 children was rated before and four hours after the administration of an analgesic. The PPP scores were found to be significantly higher before than after analgesic administration ($p< 0.001$). As part of this process, the behavior of 30 children was rated before and five days after surgery. Internal consistency ranged from 0.75 to 0.89 (Cronbach's alpha), and sensitivity (1.00) and specificity (0.91) were optimized at a cut-off point of 14 on a 60-point scale. Although there was no significant difference between the mean preoperative and postoperative scores, the highest PPP score occurred in the first 24 hours after surgery in 14 (47%) children. Yet, the authors claim that the PPP should be considered as reliable and valid and suggest that it has potential for both clinical and research purposes.

Despite such claims, it seems that more rigorous psychometric properties need to be established for the PPP and that further research is required in order to evaluate the acceptability, feasibility, and usefulness of the PPP as a tool in clinical settings for children with severe to profound neurological and cognitive disabilities. Further validation as an evaluative tool is also required. Finally, it has yet to be determined whether the PPP is also useful for pain assessment in adults with similar degrees of disability (Hunt et al., 2004).

6) *The Pain and Discomfort Scale* (*PADS*; Bodfish et al., 2001): This scale is based on previous research on facial expressions and body movements as indicators of

acute pain and discomfort in children (Breau et al., 2002b). The PADS was developed to assess pain in individuals without the cognitive capacity to convert internal experiences into expressed language. This scale was also designed to aid health care professionals in recognizing, diagnosing, and more effectively treating pain in patients with severe and profound communication difficulties.

Bodfish et al. (2001) conducted three validation studies on PADS. In the first study, 22 adults with severe and profound IDD were assessed with the PADS before and during acute medical procedures known to produce pain and discomfort (i.e. a gastronomy-tube insertion or a toenail removal). The total scores increased significantly during the medical procedures (p<0 .01) as compared to the baseline, and the PADS was interpreted by the authors as being sensitive to pain and discomfort in this population (Bodfish et al., 2001). In the second study, the scores in a group of patients with painful chronic medical conditions and physical disabilities were significantly higher (p<0 .01) than in patients with severe and profound levels of IDD alone (Bodfish et al., 2001). In the last study, eight adults with a profound level of IDD as well as other medical conditions were assessed with the PADS both before and after pain treatment. In all cases, there was a significant reduction in the score from baseline to treatment, which the authors interpreted as indicative of treatment effects and reduced pain (Bodfish et al., 2001).

The work of Bodfish et al. (2001) was later used to detect pain and discomfort during a dental scaling procedure. Twenty-eight subjects with cognitive and communication deficits were assessed at multiple baselines as well as during and after the procedure. Reliability was found to be between 93.6%–99.7%. The results indicated that scores on the PADS were significantly higher during the procedure than during all other non-pain situations quantified by the PADS. An optimal cut-off point

for sensitivity and specificity (Groth-Marnat, 1997) has not yet been demonstrated for the PADS (Phan et al., 2005). However, the accumulating evidence suggests that the PADS is a sensitive measure of pain in adults with IDD (Bodfish et al., 2001; Phan et al., 2005).

1.4.4. The use of proxy observer

Although self-report of pain and symptoms is always preferable, that option is not available for many individuals with IDD. While it is possible for some individuals with neurological and cognitive impairment to use certain forms of self-report (Fanurik et al., 1998; Ferrell et al., 1995; Parmelee, 1996), those unable to report their pain must remain dependent on their caregivers' observational skills. Due to the cognitive and communication difficulties presented by individuals with IDD, a designated external evaluator is usually assigned to perform the assessment. This person should have close knowledge of the observed individual and is termed a proxy observer. When communication is difficult, assumptions by health care professionals about the meaning of any individual's painful experience should be made with caution. Since expression of pain reflects a complex mixture of physical and emotional states, coping style, and family and cultural expectations, it can be misinterpreted by health care professionals (American Academy of Pediatrics and American Pain Society, 2001).

Past reports on the use of proxy observers in regards to abstract concepts, such as quality of life, have been contradictory. Some studies found that relatives and clinicians have a tendency to underestimate the well-being of the person with IDD (Britto et al., 2004; Ennett et al., 1991), while others showed that caregiver ratings were significantly higher than the ratings of the person with IDD (Schwartz &

Rabinovitz, 2003). Still others found substantial positive correlations between the two ratings and no significant difference between self-reports of individuals with IDD and caregiver reports (Stancliffe, 1999).

As for reports related to illness in individuals with IDD, it seems that parents and health care professionals display good agreement on the presence of symptoms for such conditions as asthma and respiratory illness (Gorelick et al., 2002; Samet et al., 1993; Vandvik et al., 1988). When pain reports are in question, recent studies have shown that caregivers' reports provided the best alternative for self-reports in research aimed at understanding and alleviating pain of children with IDD (Breau et al., 2004). When a group of individuals with IDD with good communication abilities was investigated the use of caregivers as proxy observers of pain experience has been found to correspond with the reports made by the individuals experiencing pain themselves, both in regards to acute pain (Schneider et al., 1992) and longer-lasting pain (Miller, 1996). Therefore, most of the existing evidence suggests that the translation of health-related non-verbal communication by proxy observers familiar with the patient does accurately reflect and facilitate the monitoring of health problems and pain in this population.

1.5. CONSIDERATIONS IN DEVELOPING A PAIN ASSESSMENT SCALE

The creation of an assessment tool to examine a construct such as pain is a complex, time-consuming process. The decision to develop a new tool must come after careful consideration of existing tools and their strengths and limitations. Tool development requires significant expertise in both the concept and the research process required to develop and test a tool that will evaluate, discriminate, and/or predict the concept (Duhn, & Medves, 2004).

The strength of a pain scale depends primarily on its reliability and validity (Beyer, & Wells, 1989), and requires an understanding of the various components of each and the methods by which they should be tested. The complexity of establishing reliability and validity of a scale means that multiple studies are necessary before it is deemed adequate for either research or clinical use (Beyer, & Wells, 1989). Further, the assessment of reliability and validity should be conducted prior to using the scale in either of these settings (Beyer, & Wells, 1989).

1.5.1. Reliability

Reliability is “the degree to which test scores are free from error of measurement" (American Educational Research Association, 1999, p.19). There are several components of reliability - instrument reliability, intrarater reliability, interrater reliability and intrasubject reliability (Domholdt, 2000). Instrument reliability depends on the instrument's type. The present investigation concerns an observational measure including a multitude of items, therefore internal consistency should be examined. Internal consistency assesses the extent to which the components of a tool are measuring the same construct (Burns, & Grove, 1997; LoBiondo-Wood, & Haber, 1998).

Intrarater reliability is defined as "the consistency with which one rater assigns scores to a single set of responses on two occasions (Walts, et al., 1984, p.141). Intrarater reliability examines thus the stability of the instrument across time (Burns, & Grove, 1997; LoBiondo-Wood, & Haber, 1998). Interrater reliability is defined as "the consistency of performance among different raters or judges in assessing scores to the same… response…[It] is determined when two or more raters judge the performance of one group of subjects at the same point in time" (Walts, et al., 1984, p.140).

Therefore, interrater reliability is a measure of equivalence and assesses the agreement among users of the tool (Burns, & Grove, 1997; LoBiondo-Wood, & Haber, 1998). Beyer and Wells (1989) suggest that interrater reliability is important when observation is the method of data collection. Intrasubject reliability is the stability of a measured phenomenon within the observed subject across different points in time. Measuring intrasubject reliability is done through test-retest measurements, yet this measure usually also includes tester errors and instrument errors, on top of true subject variability (Domholdt, 2000), which are hard to separate from one another.

Reliability can be quantified by relative and absolute reliability. Relative reliability "examines the relationship between two or more sets of repeated measures" (Domholdt, 2000, p.257). Relative reliability is based on the idea that if a measurement is reliable, individual measurement within a group will maintain their position within the group on repeated measurements. Relative reliability is measured with some form of a correlation coefficient which indicates the association between repeated measures. It is commonly suggested that reliability coefficients greater than 0.8 are considered "very reliable", while coefficients below 0.70 are considered "poor reliability"(Currier 1990, in Domholdt, 2000, p.258).

Absolute reliability "examines the variability of the scores from measurement to measurement (Domholdt, 2000, p.257). A statistic used to measure absolute reliability when variability is unrelated to the size of the score is the Standard Error of Measurement (SEM). Otherwise appropriate transformation of the scale should be considered. Given that the error is proportional to the score, coefficient of variance (CV) is an appropriate measure (Bland, 1996). Measures of reliability are defined by some researchers as reproducibility (Terwee at al., 2007).

1.5.2. Validity

Validity is the "appropriateness, meaningfulness, and usefulness of the specific inference made from test scores" (American Educational Research Association, 1999, p.9), or in different words, an instrument is valid if it measures the construct it intends to measure (Burns, & Grove, 1997; LoBiondo-Wood, & Haber, 1998). The various aspects of validity commonly examined include construct, content, and criterion validity.

Construct validity is the "validity of the abstract constructs that underlie measures" (Domholdt, 2000, p.259). Construct validity provides the strongest evidence for validity (Suraseranivongse, et al., 2001) and is determined by the extent to which the measurement reflects the actual construct (LoBiondo-Wood, & Haber, 1998). Construct validity also implies whether the tool can detect predictable changes in the construct, e.g. whether pain scores change with administration of analgesia (Beyer, & Wells, 1989; LoBiondo-Wood, & Haber, 1998). Construct validity may also be established by demonstrating high correlation with a previously validated instrument (LoBiondo-Wood, & Haber, 1998; Suraseranivongse, et al., 2001), and it is proposed that predefined hypotheses concerning the associations should be stated (Terwee et al 2007).

Content validity is the "extent to which a measure is a complete representation of the concept of interest" (Domholdt, 2000, p.260). Content validity reflects the extent to which the instrument is representative or inclusive of all features of the construct. Content validity can be established by reviewing the literature and collecting experts' opinions regarding the contents (Beyer, & Wells, 1989; LoBiondo-Wood, & Haber, 1998; Suraseranivongse, et al, 2001). Terwee at al., (2007) who have developed quality criteria for questionnaires, proposed that authors of a new measure

should provide a clear description of the aim of the measure, the target population, criteria for item selection and item reduction, and interpretability of the items. Content validity is important for the development of an instrument, yet further than the development phase of the instrument, more rigorous evidence of validity is required (Burns, & Grove, 1997).

Criterion validity is “the extent to which one measure is systematically related to other measures or outcomes" (Domholdt, 2000, p.261). Criterion validity indicates the extent that a subject’s performance on the instrument is related to actual behavior (LoBiondo-Wood, & Haber, 1998). Concurrent validity is a component of criterion validity (LoBiondo-Wood, & Haber, 1998) which describes the level of correlation between two instruments considered to measure the same or a similar construct when applied at the same time; one considered an accepted measure or "gold standard", the other new or undergoing construction. High correlation indicates a high level of agreement between the two measures, and thereby evidence is provided that the new measure is a valid measure of the construct (LoBiondo-Wood, & Haber, 1998). Terwee et al. (2007) propose that the correlation should be at least 0.70. In pain assessment the “gold standard” is usually a self-report measure. Since the target population for the present research intervention is individuals with communication difficulties, the correlation with self-report measures can not be obtained.

Responsiveness could be defined as the ability of a scale to detect even small clinically important changes over time (Guyatt, et al., 1989), thereby distinguishing patients who had changed from those who did change. According to Terwee at al., (2007) responsiveness is an aspect of validity, and can be adequately evaluated by the Receiver Operation Curve (ROC). Responsiveness is considered adequate if Area Under the Curve (AUC) is at least 0.70 (Deyo and Centor, 1986).

1.5.3. Additional measurement properties

Floor and ceiling effects are present if more than 15% of respondents achieve the lowest or highest possible score, respectively (McHorney and Tarlov, 1995). Floor and ceiling effects harm the assessment of reliability, since variability cannot be demonstrated in participants with extreme values. Floor and ceiling effects may also affect responsiveness, since the scale cannot capture improvement and/or deterioration (Terwee at al., 2007).

Interpretability is defined “as the degree to which one can assign qualitative meaning to quantitative scores” (Lohr et al., 1996). When a scale is being developed the researcher is expected to suggest clinically meaningful information such as: means and SD's of target population and of subgroups, relevant cut-off points for discrimination and change, norm values, etc' (Terwee at al., 2007).

We believe that the criteria presented above should be applied for scales that evaluate health issues such as pain. The existence of predetermined criteria is helpful when constructing a new scale and enables the researcher a possibility to check the final result and to evaluate the quality of the scale. The above criteria will be used to evaluate the quality of the present scale after its completion.

1.5.4. Clinical considerations

A pain assessment tool might be highly reliable and valid, yet too difficult or cumbersome to use in a clinical setting. Therefore, when selecting a pain assessment tool in the clinic, the health care provider should not only review the reliability and validity of available tools, but should also consider their clinical utility in relation to the target population and practice setting. Critical factors include the length of the scale as well as its ability to differentiate between different pain situations.

Determining whether a specific tool is appropriate for research purposes and/or clinical use also requires careful consideration. For instance, a video-based evaluation tool would necessitate the proper instrumentation and adequate time to take the video and view it later. On the other hand, the evaluation of brief instances of pain, such as injection-induced procedural pain, would be very difficult to evaluate under real-time conditions without the use of some form of telemetry device.

When constructing a pain assessment tool, it is important to consider the future integration of the assessment tool into practice. This can be achieved by conducting an effective testing phase for the new scale under clinical conditions (Duhn & Medves, 2004). It is also recommended that the population in which the tool was tested be reviewed for measurement properties, as further pre-testing may be required if a different population is targeted.

2. THE CURRENT STUDY

2.1. AIMS OF THE STUDY

Individuals with IDD represent a population at risk for suffering pain. Yet, due to the fact that they comprise a minority group with unique behavioral and communication difficulties, few existing pain scales were developed for use in this population, mostly for children. As a result, pain management for individuals with IDD is sorely lacking and development of appropriate pain assessment tools specifically adapted for this population is required.

The overall aim of the study is to develop a pain assessment tool that can be used for people with IDD, irrespective of IQ level, and that will maintain strong psychometric properties across different pain experiences and settings.

The aims of the separate papers were:

Paper I

To investigate two different pain behavioral scales (the NCCPC-R and the FACS) and to examine whether acute pain behavior in adults with IDD is different from that of adults without IDD, as well as whether such pain behavior is affected by the level of IDD, and is consistent with verbal reports of pain.

Paper II

To explore the adequacy of using the NCCPC-R items to assess pain behavior in adults at different ages and different levels of IDD, and if necessary to adapt the scale to better fit the target population.

Paper III

To evaluate relative and absolute intra- and interrater reliability of the NCAPC total scores, based on the observation of video recordings of adults with IDD who

received an influenza vaccination, also exploring reliability among different groups of health care personnel working with this population and thus considered potential users of the scale.

Paper IV

To evaluate measurement properties and thus applicability of the NCAPC in a clinical setting. This was done by examining the ability of the scale to distinguish between pain and non-pain situations and different levels of pain (discriminate validity) and by examining sensitivity to pain in subgroups of adults with different levels of IDD.

2.2. REVIEW OF PAPERS

2.2.1. The evaluation of acute pain in individuals with cognitive impairment: A differential effect of the level of impairment

Defrin R, Lotan M, Pick CG. Pain, 2006, 124(3):312-20.

Background: Despite enhanced interest in the manifestations of pain in adults with IDD, the characteristics of pain behavior in this group have seldom been examined.

Objective: To investigate whether the level of IDD affects acute pain behavior and how it is manifested.

Method: The behaviors of 159 individuals (mean age 42 years ± 12), including 121 with IDD (divided into four groups according to the level of IDD: mild, moderate, severe, and profound) and 38 with normal cognition (comparison group), were rated by two raters, using the FACS and the NCCPC-R, both before and during an influenza vaccination.

Results: Individuals with severe or profound IDD exhibited more elevated FACS and NCCPC-R values at baseline as compared with all other groups ($p<0.01$). Both the FACS and the NCCPC-R scores of individuals with mild to moderate IDD and the controls increased significantly during vaccination ($p<0.001$). In contrast, individuals with severe or profound IDD exhibited high rates of "freezing reaction" (stillness) during vaccination, manifested mainly in the face and therefore resulting in an elevation of only the NCCPC-R scores, but not of the FACS scores.

Conclusions: The results suggest that the level of IDD affects baseline behavior as well as pain behavior. Therefore, it is necessary to choose an appropriate behavioral tool to measure pain in these individuals accordingly. For example, tools based on facial reactions alone, such as the FACS, might provide the false impression that individuals with severe or profound IDD are less reactive to pain. The NCCPC-R was

thus more appropriate to serve as a base for further development of a scale for pain measurement in adults with IDD.

2.2.2. A modified version of the Non-Communicating Children's Pain Checklist-Revised (NCCPC-R), adapted to adults with intellectual and developmental disabilities. Sensitivity to pain and internal consistency

Lotan M, Ljunggren AE, Johnsen TB, Defrin R, Pick CG, Strand LI. The Journal of Pain 2009;10(4):398-407.

Background: The characteristics of pain behavior in adults with IDD have seldom been examined. A previous study found the NCCPC-R (a scale designed for the pediatric population) to be a sensitive scale for pain behaviors in adults with IDD.

Objective: To further develop the NCCPC-R and to provide a sensitive pain behavior scale for adults with IDD.

Method: A total of 228 adults (mean age: 38.7 years) with different levels of IDD were videotaped before and during an influenza vaccination and were scored using the NCCPC-R. Observed pain behaviors not captured by this measure were also registered. Each of the 27 items was examined for sensitivity to change by the Signed Rank test and for internal consistency by Cronbach's alpha. Sensitivity to change of the total scale was examined by Standardized Response Mean (SRM) in the whole sample, as well as in sub-samples at different levels of IDD.

Results: Thirteen items were excluded from the original 27-item NCCPC-R scale, while four new items were added, resulting in a modified scale of 18 items. This scale, named the Non-Communicating Adults Pain Checklist (NCAPC), was re-scored and examined for measurement properties in a random sample (N=89). Sensitivity to pain of all items ($p<0.05$) and satisfactory internal consistency ($\alpha= 0.77$) of the total scale were demonstrated. High sensitivity to pain at all levels of IDD was shown (SRM ranging between 1.20 and 2.07). Overall, better measurement properties were demonstrated for the NCAPC than for the NCCPC-R in the target population.

Conclusion: The initial measurement properties of a new measure, the NCAPC, for evaluating pain behavior in adults with IDD were demonstrated. This measure seems to be a promising tool for capturing pain expressions in the population of adults at all levels of IDD and as such may contribute to better pain management for this group of patients. However, it should be further examined in regard to reliability and validity issues

2.2.3. Reliability of the Non-Communicating Adult Pain Checklist (NCAPC), assessed by different groups of health workers

Lotan M, Moe-Nilssen R, Ljunggren AE, Strand LI. Research in Developmental Disabilities (RIDD) 2009;30:735–45.

Background: Developing tools to evaluate pain in adults with intellectual and developmental disability (IDD) is a challenge. The NCAPC, which was recently developed from the NCCPC-R and examined in a group of adults with IDD (N=228), was found to hold satisfactory internal consistency and sensitivity to pain.

Objective: To explore the intrarater and interrater reliability of the NCAPC.

Method: Data collection was done by videotaping the participants before and during an influenza vaccination. Intrarater reliability was examined by the first author on a group of 50 randomly selected adults with IDD from the total sample of 228 individuals (mean age 42.5 years, range 19-72). Interrater reliability was investigated in two stages. In the initial step, different groups of health care workers (caregivers, nurses, case managers, and therapists), each including five raters, viewed a sample of 12 adult participants with IDD (three at each level of IDD) that were extracted from the total sample. In the second stage 3 participants from each of the rater groups showing high interrater reliability (caregivers and therapist) evaluated interrater reliability in a randomly selected group of 40 individuals.

Results: The mean ICC(1,1) for intrarater reliability was found at 0.94. The interrater reliability of all raters according to ICC values within the groups varied from low to very high [ICC(1,1)=0.40–0.88]. The interrater reliability was very high among caregivers and physical and occupational therapists and these groups were considered as potential users of the measure. In the second stage, three participants from each of the groups showing high interrater reliability (i.e. caregivers and therapists) were examined for their interrater reliability with 40 randomly selected adults with IDD.

The overall interrater reliability for the caregivers and therapists, according to ICC values, were 0.92 and 0.91, respectively.

Conclusion: The NCAPC was found to hold high reliability values when scoring was performed by primary caregivers and physical and occupational therapists, based on video uptakes.

2.2.4. Measurement properties of the Non-Communicating Adult Pain Checklist (NCAPC): a pain scale for adults with Intellectual and Developmental Disabilities, scored in a clinical setting.

Lotan M, Moe-Nilssen R. Ljunggren AE, Strand LI. Research in developmental disabilities (RIDD). Accepted.

Background: The NCAPC was developed to capture pain behavior in adults with IDD. Measurement properties concerning internal consistency, sensitivity to pain, and reliability were found to be satisfactory in a previous study, using scores from observations of video uptakes.

Objective: To examine the NCAPC's discriminative ability and sensitivity to pain based on scores from a clinical situation.

Methods: Fifty-eight adults at all levels of IDD were observed for pain behavior, both before and during a dental hygiene treatment and an influenza vaccination, using the NCAPC.

Results: The NCAPC differentiated between pain and non-pain situations, as well as between different pain incidents, and was found to be sensitive to pain at all levels of IDD.

Conclusions The results add to previous findings of measurement properties of the NCAPC, and support that it can be scored directly in a clinical setting.

2.3. METHODS AND PARTICIPANTS

2.3.1. Identifying promising pain assessment tools

Available pain behavior scales that had been tested prior to the initiation of the current investigation were evaluated (Appendix IX) in order to identify the most promising pain assessment tools. The two most promising scales were The Non-Communicating Children's Pain Checklist-Revised (NCCPC-R) and the Facial Action Coding System (FACS).

2.3.1.1 The Non-Communicating Children's Pain Checklist-Revised (NCCPC-R) is a pain measurement tool specifically designed for children with cognitive impairments. The scale was previously described at length in the introduction (pages 17-18). The present investigation employed the revised version (Breau et al., 2002a), which includes a list of 30 items divided into seven categories (subscales): vocal, eating and sleeping, social and emotional reactions, facial expressions, body and limb activity, and physiological signs. The observer provides a score for each item on a scale of 0–3 according to the frequency of its occurrence during a specific time period, as follows: 0 = not at all, 1 = just a little, 2 = fairly often, 3 = very often. The NCCPC-R total score is computed for each participant by summing up the scores of the items appearing on the scale.

2.3.1.2. The Facial Action Coding System (FACS) (Ekman & Friesen, 1978) is a list of facial actions (action units – AUs) based on the movement of specific facial muscles. The scale, as well as its use in individuals with IDD, was described in the introduction (pages 15-16). In the present study, we used the modified FACS,

including 14 AUs that were previously found to be specifically characteristic of pain (Craig et al., 1992, 1994, 2002; Prkachin, 1992; Prkachin & Mercer, 1989): Brow lowerer (AU4), cheek raiser (AU6), lid tightener (AU7), nose wrinkler (AU9), upper lip raiser (AU10), oblique lip puller (AU12), lip stretcher (AU20). lip presser (AU24). mouth opener (AU25). jaw dropper (AU26), mouth stretcher (AU27), eyelid dropper (AU41), eye closer (AU43), and blinker (AU45). The intensity of most AUs was coded on a 6-point intensity scale, ranging from 0 = no action, through 1 = minimal action/trace to 5 = maximum action (Prkachin, 1992). The final score for each participant was the sum total of the intensity scores of the AUs. The raters were trained in FACS ratings with the manual of Ekman and Friesen (1978).

2.3.2. Participants

In **paper I**, the total sample consisted of 159 adults receiving a mandatory influenza vaccination, including 121 with varying levels of IDD: 22 mild (IQ 70–55), 43 moderate (IQ 54–40), 32 severe (IQ 39–25), and 24 profound (IQ< 25). Individuals with mild IDD were living in community settings, while all other individuals with IDD were living in a residential center, closely representing the diversity of the IDD population in Israel. A comparison group was composed of 38 individuals with a normal level of cognition, but living in a residential center due to physical disabilities (associated with cerebral palsy and other neurological conditions). This specific group was included in order to control for the potential effects of physical disabilities and institutionalization on behavioral gestures of pain. To be included, all participants had to be healthy adults receiving an influenza vaccination and to have resided for at least three months in their present habitat.

Individuals were excluded if they: 1) had recently moved into the residential setting and might not be well known to the staff or familiar with the vaccination procedure, 2) had suffered physical or emotional distress within the last three months and therefore behaved in an abnormal way due to trauma or depression, or 3) did not appear clearly on the video recording during the vaccination shot, preventing their behavior to be clearly judged.

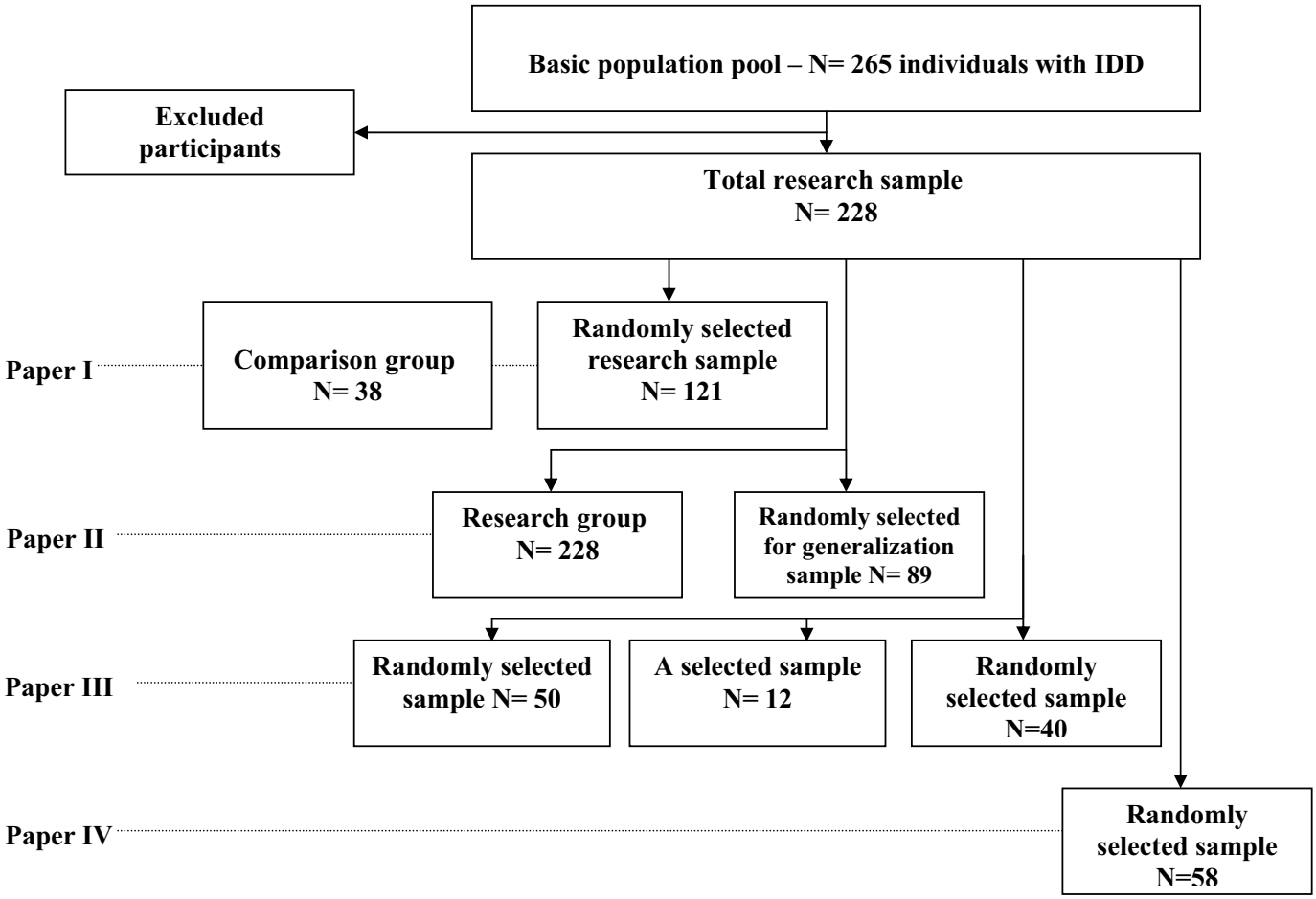
In **paper II**, all participants considered for inclusion in the study were adults with IDD who received a mandatory annual influenza vaccination (n=265)[2] and were living either in a residential center or in community settings, closely representing the diversity of the IDD population in Israel. After excluding 37 individuals for various reasons, 228 participants were evaluated for their pain behaviors from video recordings. After constructing the new pain behavior scale, it was tested on a randomly selected group of 89 participants to enable generalization of findings.

In **paper III**, three samples were drawn from a total research group of 228 adults with IDD who were diagnosed with different levels of IDD. Sample 1 (N = 50) was extracted for the intrarater reliability study by randomly drawing numbers of the participants. Sample 2 was extracted for the initial interrater reliability evaluation (N=12) by including three participants from each level of IDD (mild, moderate, severe, and profound). Pain behavior during the influenza vaccination shot was initially evaluated using the NCAPC. This procedure allowed the inclusion of participants with low, intermediate, and high levels of pain behaviors in each group. Sample 3, the main sample for examining interrater reliability (N=40), was extracted by randomly drawing numbers of the participants.

[2] N=265 are all participants in this project with IDD. All other subgroups in papers I, III, and IV were extracted from this group of clients.

In **paper IV**, a random sample of 58 participants was selected from the research population with different levels of IDD: profound (N=10); severe (N=12); moderate (N=26); and mild (N=10). Of this group, 39 were mobile and the others were ambulatory only by means of a wheelchair. All received dental hygiene treatments at the dental clinic within the residential setting. Figure 2 summarizes the different research samples.

Figure 2 - Flowchart presenting participants in the different research projects



2.3.3. Data collection

In **papers I, II** and **III** video recording were used, and data collected according to the following procedure: All participants were videotaped before (baseline) and during influenza vaccination that took place in the institutions and residential houses (familiar environments). The behavioral responses of the participants during these two time frames (baseline = T^0, vaccination = T^1) were analyzed retrospectively with the aid of the frame-by-frame analysis and the slow motion video. During T^0, the participants sat on a chair or on a wheelchair and were not engaged in any specific activity. Rating was conducted for a random 5-s segment. During T^1, the participants were injected with the vaccination. Rating T^1 commenced the second the nurse touched the skin with the syringe, immediately after swabbing and lasted 5 s (Craig et al., 1984; LaChapelle et al., 1999; Breau et al., 2001). One camera photographed only the face area of each participant, while the other camera photographed the whole body including the face. Audio was also recorded by both cameras.

In **Paper I**, two trained raters (authors RD and ML) viewed the videotapes and rated the pain behavior, using two different scales: The NCCPC-R and the FACS. Therefore a total of four time frames were analyzed by the raters for each participant: T^0 and T^1 with the NCCPC-R (using the whole body shots) and T^0 and T^1 with the FACS (using the close-up shots). Each time frame was rated separately. The different time frames of all participants were presented to the raters in a random order. In addition, the two raters conducted their analysis separately to prevent any influences between them. It should be pointed out that one rater (author ML) was familiar with most of the participants as one of their therapists, whereas the other rater (author RD)

was unfamiliar either with the participants or with the level of their IDD. The final coding was based upon consensus between the raters.

In **paper II** all videos of all participants were observed by the first author (ML) and scored with the NCCPC-R scale. Every item's sensitivity to change was examined, as well as internal consistency of the items in the scale and the sensitivity of the scale was checked for sensitivity. The end outcome was the deletion of 14 items and reconstruction of four new items, resulting in a new scale. This scale was rechecked for sensitivity of items, internal consistency and sensitivity of the scale on a random sample of 89 participants to enable generalization of the findings beyond the research sample.

In **Paper III**, the reliability of test scores based on the video recordings was examined. All videos were observed in random order, and the pain behaviors of the participants were graded using the NCAPC form. In order to examine intrarater reliability, the videos of 50 participants (Sample 1) were assessed by the first author (ML) twice with a three-month interval between the two assessments so as to prevent recall of the first scores. In order to assess interrater reliability as well as applicability of the scale among different health care workers, a two-stage procedure was conducted. In the first stage, videos of 12 participants (Sample 2) were assessed by four groups of different professions of health care workers, each including five raters: caregivers, nurses, case managers, and therapists. Only the caregivers and therapists demonstrated high interrater reliability within the groups and were considered as potential users of the NCAPC. Therefore, in the second stage, three participants in each of the groups of caregivers and therapists were asked to evaluate a larger sample, including 40 adults with IDD (Sample 3).

In **paper IV**, the usability of the NSAPC was examined in a clinical situation. Participants who were invited to a dental clinic for a routine scaling procedure were evaluated for their pain reactions. Observations took place on randomly selected days, and participants were included in the present study and evaluated for their pain experience if they had been previously evaluated with the NCAPC during their annual influenza vaccination. The procedure of observing clients in the dental clinic continued until there were at least 10 individuals at each level of IDD (profound, severe, moderate, and mild). All treatments were performed during the morning work hours at the dental hygienist's clinic.

Prior to the dental hygiene treatment, all patients invited to the dental clinic on a given day were observed by the first author (ML) using the NCAPC. Every patient was observed while sitting in the waiting room before the treatment (with other individuals in the room) for two consecutive five-minute periods. The observation took place while the observer was standing outside the waiting room, watching the patients through a window. This was done to prevent the participants from focusing on the observer and thereby changing their behavior. On arrival to the waiting room, the escorting caregivers were instructed to position the participants in a manner allowing a clear view of their face and body. The assessor filled out two forms, one for each five-minute period.

During the dental hygiene treatment, every patient was observed for two consecutive five-minute periods. The observer was standing on the left-hand side of the treatment bench, where the clearest view of the participant's body and face was obtained. The scores obtained in the dental clinic were compared with those collected from video uptakes of the same individuals during their annual influenza vaccination approximately four years before.

2.3.4. Statistical analyses

In **paper I,** the data were analyzed with SAS software. Internal consistency of the FACS and the NCCPC-R was assessed with Cronbach's alpha test. The agreement between the two raters for the FACS was computed with the Ekman and Friesen Conservative FACS reliability formula (Grunau & Craig, 1987) for a random sample of 24. In addition, agreement between raters was also assessed with a mixed-effects model with interactions, in which the coder was the fixed effect and the scores were the random effect. Fixed-effects models were used to assess the effect of several independent variables on the FACS and the NCCPC-R scores. The variables were: time (baseline, vaccination), group (individuals with IDD, controls), level of IDD (mild, moderate, severe, and profound), gender, Age, duration of institutionalization, mobility profile, and diagnoses. Post-hoc tests were corrected for multiple comparisons using the False Discovery Rate (FDR) procedure. The correlation between the two time frames (baseline, vaccination) for the FACS and NCCPC-R, as well as the correlation between the two tools (FACS, NCCPC-R) for each time frame, were examined by Pearson's correlations. Verbal reports based on the faces scale (only for individuals with mild to moderate IDD) were compared with the scores obtained by the FACS and the NCCPC-R. All p-values presented are two-tailed.

In **paper II**, data were analyzed with SPSS version 14 and SAS software. Since the test scores were not normally distributed, Wilcoxon matched Signed Rank was used to examine the change in scores of each item between the time periods before and during the injection. Internal consistency by Cronbach's alpha was examined using the "during" scores for all 27 items of the NCCPC-R, as well as for the final 18 items of the NCAPC. It is recommended that the alpha value should be within the range of 0.7-0.9 and that the correlation between the separate items and the total score

should be above 0.2 (Streiner & Norman, 2001). The impact of each item on the total alpha value was examined using the option "if item deleted" in the reliability analysis.

Sensitivity to change of the total sum score was examined and compared by calculating the Standardized Response Mean (SRM), dividing the mean total change scores from before to during the injection by the standard deviation of the total change scores (Finch et al., 2002). SRM was calculated for the NCCPC-R, as well as for the NCAPC, and compared in each subgroup of individuals with different levels of IDD.

In **paper III**, the data were analyzed with SPSS 14.0 for Windows (SPSS, Inc., Chicago, IL). Relative and absolute reliability were quantified. Relative reliability was examined using Intra Class Correlation (ICC) coefficients, reflecting the relation of variability caused by measurement error to total variability in the data (Rankin & Stokes, 1998). In order to evaluate the stability of scores by a rater (intrarater reliability) and between different raters (interrater reliability), ICC(1,1) was used. The ICC(1,1) model is based on a one-way analysis of variance in which all variation between occasions is regarded as measurement error (Shrout & Fleiss, 1979). In a slightly different model, the ICC(3,1), the effect of any systematic shift in the data is not considered as part of the error of measurement (Moe-Nilssen, 1998; Shrout & Fleiss, 1979).

In order to evaluate the usability of the NCAPC by the different health care personnel, reliability analysis for all raters within each group, as well as pair-wise analysis within each group, was performed. Although there is no consensus on how to judge ranges of correlation coefficients, Currier (1990) has suggested that ICCs in the range of 0.60-0.79 might be considered as moderate reliability, while ICCs in the range of 0.80-0.89 would be considered as high reliability.

Absolute reliability indicates the extent to which a score varies on repeated measurements (Domholdt, 2000). In order to demonstrate absolute reliability, within subject standard deviation (Sw) was calculated, expressed in the same units as the measurement tool (Roebroeck et al., 1993). Assuming Sw to be normally distributed, the difference between a patient's measurement and the true value is expected to be less than 1.96 for 95% of observations (Bland & Altman, 1996). The difference between two repeated measurements for the same patient is expected to be less than 2.77 Sw for 95% of pairs of observations. This value is called the Smallest Detectable Difference (SDD). Therefore, only a change in a measure that exceeds the SDD in an individual should be claimed to be a real change, rather than simply measurement error.

In **paper IV**, SPSS 14.0 for windows was used for statistical analysis (SPSS, Inc., Chicago, IL). In order to evaluate the stability of scores by a rater (test-retest), ICC statistics were used, as described above for ICC(1,1) and ICC(3,1). A third model, the ICC(1,2), while based on a one-way analysis of variance, was used to measure reliability through the mean of two scores. The ICC procedure was performed on the scores of the first and second five-minute periods prior to the dental hygiene treatment observations. The procedure was also performed on the scores of the first and second five-minute periods during the dental hygiene treatment observations in the dental clinic in order to examine the stability of pain behaviors across time and to determine the proper length of pain evaluation of the NCAPC in clinical situations.

Since the test scores were not normally distributed, the Wilcoxon matched Signed Rank test was used to examine the hypothesis of no difference between two sets of data in the following comparisons: 1) NCAPC sum scores of the two consecutive

five-minute periods before the dental hygiene treatment in order to examine the stability of baseline (no pain) behavior, 2) NCAPC sum scores of the 10-minute period (two consecutive five-minute periods) in the waiting room prior to the dental hygiene treatment (no pain) and the baseline scores from the video uptakes in the participants' residence prior to the influenza vaccination (no pain) in order to examine the stability of no-pain behavior across locations; 3) NCAPC sum scores of the periods before and during the dental hygiene treatment in order to examine the ability of the NCAPC to distinguish between pain and non-pain situations; 4) NCAPC sum scores of the two periods during the dental hygiene treatment observations and those of the same participants during the influenza vaccination in order to examine the ability of the NCAPC to distinguish between different painful stimuli, with the dental treatment expected to cause more pain behavior than the influenza vaccination.

To examine the association of pain behavior between two painful experiences, Pearson's correlation was performed between the sum scores of the period during the dental hygiene treatment and those of the period during the influenza vaccination in the same participants.

Sensitivity to change of the NCAPC was examined by calculating the Standardized Response Mean (SRM), dividing the change scores from the period before the dental hygiene treatment (no pain) to the period during dental treatment (pain) by the standard deviation of the change scores, as described by Finch et al. (2002). The SRM was calculated for the whole sample, as well as for each of the four different levels of IDD.

2.3.5. Ethical issues and approval process

The ethical guidelines for pain research in humans, published in the IASP web site (IASP, 2009) were taken into consideration when planning and performing the present investigation. As suggested by the guidelines, before starting any study on human subjects, the researcher is to present the experimental protocol to an independent committee on human research. The protocol of the present research was presented and approved by the Haifa University Committee of Ethics in Research (Appendix III-IV).

The guidelines also claim that "potential participants should be informed fully about the goals, procedures and risks of the study before giving their consent. Healthy subjects and patients must be able to decline, or to terminate, participation at any stage without risk or penalty whatsoever. Written consent must be obtained to indicate that the subject understands the nature and purpose of the proposed study, has had the opportunity to ask questions and agrees to participate on a voluntary basis. Where possible, informed consent should be endorsed by an independent signatory" (IASP, 2009). In this regard, written informed consent was obtained from all non–IDD participants (**Paper I**) after explaining them the aims of the study and its protocol.

The duty to protect those who may be incapable of giving fully informed and voluntary consent, including children, elderly, mentally handicapped, prisoners, and the very ill, is also pointed out in the above mentioned IASP guidelines. Such persons, it is said "should not be used for medical research unless they are essential for the goals of the proposed research. In such cases, consent must be obtained also from those who have legal responsibility for their welfare" (IASP, 2009). In the present research informed consent was obtained from the appointed guardians of all participants with IDD, as well as from the head of the section for care for individuals

with IDD (Appendix V-VI), after explaining the aims of the study and its protocol (**Papers I-IV**).

Moreover it should be noted that no pain was ever inflicted by the researcher on any of the participants during the present investigation. To avoid inflicting pain, mandatory painful procedures such as annual influenza vaccination and routine dental cleansing, were carefully selected and observed. Videotaping and observing the participants by the researcher during the mandatory painful situations might potentially be disturbing to the participants. However, effort was made to do this very discretely, and video-uptakes were performed by the primary researcher who is a practicing therapist working with and therefore familiar with the participants, and therefore subjected to privacy protection laws. The video uptakes were only viewed by the researchers.

When the last version of the new scale was complete and psychometric measurements confirmed, permission to change the name of the scale from NCCPC and rename it as NCAPC was approved by the author of the NCCPC, Dr. Lyn Breau (Appendix VII)

3. RESULTS

The most important findings of this research are:

- The FACS scores increased significantly during vaccination relative to baseline only in individuals with mild to moderate IDD, as well as in a comparison group without IDD. Individuals with severe or profound IDD had a significantly lower change in scores as compared to individuals with mild or moderate IDD and the comparison group. This indicates that the FACS is not sensitive to pain in adults with severe or profound IDD (**Paper I**).
- All participants at different levels of IDD exhibited significantly higher NCCPC-R scores during vaccination than at baseline. This indicates that the NCCPC-R was sensitive to pain in all subgroups of IDD (**Paper I**).
- There was no correlation between the pain reports on the visual faces pain scale (Appendix VIII) obtained from individuals with mild or moderate IDD and between their behavioral scores. Most of the individuals with IDD depicted the face representing no pain prior to the vaccination, as well as immediately following the vaccination, indicating low validity of self-report in this population (**Paper I**).
- Some frequent forms of pain behavior that were observed in the research population were not included in the original scale (NCCPC-R). The most prominent pain behavior seen in adults with severe or profound IDD was that of ‘‘freezing’’ (i.e. face and body not moving for several seconds). The observed frequency of this freezing reaction significantly increased as the severity of the IDD increased ($p < 0.01$) (**Paper I**).
- Thirteen items were excluded from the original 27 items of the NCCPC-R, while 14 were retained in the adjusted scale, as they were found to be sensitive to

procedural pain in adults with IDD. Four new or adjusted behavioral items were added in accordance with behaviors observed in the research sample, but not found in the original scale, resulting in a modified scale of 18 items (**Paper II**).

- The sensitivity of the modified scale was examined by the Standardized Response Mean (SRM). The SRM values were large (>0.80) for all subgroups of adults with IDD (mild, moderate, severe, and profound), with a range of 1.20-2.07. In comparison, the SRM of the original NCCPC-R was only moderate for the samples with severe (0.70) and profound (0.72) IDD. Thus, these results indicate a higher sensitivity to pain for the 18-item NCAPC than for the 27-item NCCPC-R, particularly in adults with severe and profound IDD (**Paper II**).
- All 18 NCAPC items demonstrated a statistically significant change ($p<0.05$) in connection with procedural pain, and the new items were found to be common pain behavior indicators. Internal consistency was higher for the 18-item NCAPC ($\alpha = 0.77$) than for the 27-item NCCPC-R ($\alpha = 0.63$) (**Paper II**).
- The intrarater reliability of the NCAPC was found to be very high, with the ICC(1,1) at 0.934. Moreover, no systematic shift in scores was observed, with the ICC(3,1) at 0.937 **(Paper III)**.
- Interrater reliability within the caregivers and the therapists groups was very high [ICC(1,1)=0.92 and 0.91, respectively] **(Paper III)**.
- The mean NCAPC sum scores obtained directly from observation (not from video uptakes) in a non-pain situation for the first and second five-minute observation periods were 4.7 and 4.8, respectively, showing no statistically significant difference ($p<0.89$). The average measure of reliability for the 10-minute period (two consecutive five-minute periods) was high (ICC(1,2)=0.93) **(Paper IV)**.

- The mean NCAPC sum scores obtained during the first and second five-minute periods of dental hygiene treatment were 15.8 and 16.8, respectively. The average measure of reliability for the 10-minute period (two consecutive five-minute periods) was high (ICC(1,2)=0.93). **(Paper IV)**.
- The mean NCAPC sum scores obtained in both non-pain situations (in the waiting room prior to the dental clinic intervention and in the residence prior to the influenza vaccination) were 4.7 and 5.1, respectively. No statistically significant difference ($p<0.45$) was demonstrated between the two non-pain situations **(Paper IV)**.
- The mean NCAPC sum scores monitored across different situations showed statistically significant lower values ($p<0.05$) in no-pain situations (in the residence and the dental clinic waiting room) than during pain situations (influenza vaccination and dental hygiene treatment) **(Paper IV)**.
- The mean NCAPC sum scores obtained during the influenza vaccination and the dental hygiene treatment were 11.5 and 16.3, respectively. The two sets of scores were found to be significantly different ($p<0.001$), as hypothesized. The correlation (Pearson) between scores on the two occasions was high ($r=0.88$), suggesting that the participants reacted in a proportional manner to the two different painful experiences **(Paper IV)**.
- The sensitivity to pain of the NCAPC was examined for the whole scale and for different IDD levels. The SRM values were high for the whole sample (2.18), as well as for all levels of IDD **(Paper IV)**.

4. DISCUSSION

4.1. MAIN FINDINGS

The aim of this thesis was to develop an instrument that can be used to assess pain behaviors in individuals with IDD by health personnel working with this population. As shown in previous studies as well as in the present investigation, the pain experience is generally hard to measure due to its subjectivity and multi-dimensional expressions (Abu Saad, 2001). Pain measurement is particularly challenging in populations with communication difficulties, such as in individuals with IDD (Hadjistavropoulos et al., 2001). Although self-reports of pain are usually considered as the "gold standard" of pain measurement, they have no value in a population lacking clear verbal communication. For these individuals, other means of pain assessment should be investigated. The validation procedure of a pain assessment tool for this population is an extremely challenging task, demonstrating not only the psychometric properties of the scale but also the characteristics of the population. The findings of the series of investigations included in this project are discussed here, followed by the methodological strengths and limitations of the development of the NCAPC.

In **paper I**, we found that the level of IDD significantly affected pain behavior. Individuals with severe or profound IDD exhibited elevated baseline (pre-vaccination) scores on the FACS and the NCCPC-R for facial and body expressions, as compared with individuals with mild or moderate IDD and individuals without IDD. During vaccination, many from the severe or profound IDD group did not demonstrate an increase in the FACS scores, but only in the NCCPC-R scores. Individuals with mild or moderate IDD, on the other hand, exhibited a significant increase in both the FACS and the NCCPC-R scores during vaccination, similar to

that of individuals without IDD. These findings led to the conclusion that not all pain behavioral tools are suitable for measuring pain in adults at all levels of IDD. Specifically, the findings from **paper I** suggest that the NCCPC-R can detect pain behaviors at all levels of IDD, whereas the FACS seems more suitable for detecting pain behaviors in individuals with mild or moderate IDD. These results are consistent with the findings of Lachapell et al., (1999).

The inability of the FACS to identify pain behaviors seemed to emerge from two phenomena that were observed. The first was the high level of spontaneous facial expressions recorded at baseline. The second was the freezing phenomenon observed during vaccination in close to 50% of the participants with severe or profound IDD. This freezing phenomenon has been described by previous researchers (Weiner et al., 1999) and defined as "stillness." The fact that these individuals seem detached and "not bothered" by pain may explain why they are often assumed to be insensitive to pain or to have a high pain threshold (Biersdorff, 1994). This phenomenon may result in the misdiagnosis of pain in this population (Feldt et al., 1998; Malviya et al., 2001), causing under-treatment (Jancar & Speller, 1994; Krauss et al., 2003; Malviya et al., 2001; Smith & Teele, 1980; Weiner et al., 1999) and even death (Carter & Jancar, 1984; Mata, 1960; Roy & Simon, 1987). Since freezing is highly prevalent, it might be helpful to incorporate it as an indicator of pain in pain behavioral checklists. It should be pointed out, however, that while freezing may indeed be a specific indicator for pain, it might also be indicative of fear, surprise, attention, or other phenomena that might be revealed through further study.

The findings in **Paper I** also show that the NCCPC-R scores reflected similar changes in the degree of pain behavior occurring at all levels of IDD, indicating that individuals with severe and profound IDD experience pain similarly to those with

mild to moderate IDD as well as non-IDD controls. The findings in **Paper I** support the notion that when measuring acute pain in individuals with severe and profound IDD, facial expressions should receive less weight than body reactions, which better represent pain expression in this population. The differences between the groups at baseline without any noxious event (no-pain) suggest that the level of IDD influences general behavior in this population, rather than pain behavior. We suggest that this important point should be considered when choosing a behavioral scale. Despite the fact that similar findings were reported by Biersdorff (1994), this phenomenon requires further investigation.

The lack of correlation in **paper I** between self-reports obtained with the visual faces scale, the FACS and NCCPC-R is not surprising. Many individuals with IDD chose the smiling face (indicating no pain) at baseline and immediately after the vaccination. Others chose the face at the middle of the scale, which has a neutral expression. Thus, as has previously been reported (Herr et al., 2004; Scherder & Bouma, 2000), it seems that most individuals with mild and moderate levels of IDD simply cannot use a face scale validly.

Since our observation of pain behavior in **paper I** indicated that the scale should be adjusted to the population of adults with IDD, an in-depth analysis of sensitivity to pain and internal consistency of test items was performed in **paper II**. The investigation resulted in a modified version of the NCCPC-R with 18 items, named the Non-Communicating Adults Pain Checklist (NCAPC).

In **paper II,** six sub-categories were retained within the NCAPC (Appendix I), though some behaviors within these categories were found to be less sensitive to pain than others. As a result, the subcategories of Vocal Expression, Facial Expression, and Protective Reaction were kept with four items in each category, while the

subcategories of Emotional Reaction, Body Language, and Physiological Reaction were reduced to only two items in each category. Our decision to follow this procedure for the Emotional Reaction subcategory is supported by previous research, showing that children with IDD display less vigorous emotional reactions and fewer social responses as compared with non-IDD children under similar circumstances (Gilbert-MacLeod et al., 2000). Moreover, the Body Language category has been found to have very low reliability values (Breau et al., 2002b; Malviya et al., 2006; Voepel-Lewis et al., 2002). In the present study, changes in body movements were often found to be pain-related, but might be expressed in opposite directions. In some cases, the movements were enhanced and in some cases the same movements reduced, probably affecting the reliability of this subcategory. Finally, the Physiological Reaction subcategory was reduced because the attribution of these types of behaviors as pain responses has long been debated for their failure to provide sufficient specificity as pain indicators (Hadjistavropoulos et al., 2001) and for their lack of promise and usability in clinical pain assessment (Breau et al., 2006).

As the newly created assessment tool contains subcategories of relatively few test items, we concluded that a sum score of all items should probably be applied. An initial indication that all test items of the NCAPC are indeed assessing a common construct was provided by the high alpha value of the total score (0.77), which was higher in the NCAPC than in the NCCPC-R when used only in adults with IDD.

In **paper II**, the NCAPC was found to be much more sensitive to pain than the original scale. This applied to the whole sample as well as to each subgroup of IDD. Sensitivity to change of the NCAPC was particularly better for people with severe and profound levels of IDD, who present the most challenging groups for pain evaluation within this population. The large SRM values indicate that the NCAPC can be used to

capture acute pain in the study population, irrespective of the level of IDD. The basis for developing the modified scale from the initial investigation of the NCCPC-R was confirmed when re-scoring the modified scale in a random sample of 89 participants. Thus, we suggest that our findings can be generalized to pain assessment in other groups of adults with IDD exposed to procedural pain.

Among the different types of pain (i.e. acute pain, chronic pain, recurring pain, pain associated with terminal illness, and procedure-related pain, American Academy of Pediatrics, 2001), procedural pain is unique insofar as it tends to involve anticipatory anxiety that may increase the likelihood of experiencing more pain and distress during the actual procedure (Blount, 2006). This additional pain and distress might in turn compound the patient's reaction, especially in the wake of a "bad" experience with the procedure (Von Baeyer, 2004).

Distress from procedural pain is addressed in the literature only in relation to children, and has never been investigated in a population of adults with IDD. Since it is common for individuals with IDD to have chronic medical conditions, it is likely that many of the participants in the present study had experienced painful tests and treatments in the past. However, even though anticipatory anxiety could have been expected in connection with the influenza vaccination, this was not observed in adults with severe or profound IDD. In fact, many seemed to have a rather late pain reaction as if they did not fully understand what was about to happen. Therefore, the topic of procedural pain should be investigated further in this population. Moreover, given that the present instrument was only evaluated during procedural pain, its use in other painful situations should also be further examined

In **paper III**, intrarater and interrater reliability of the NCAPC was examined based on the observation of pain behavior from video recordings of adults with IDD

receiving an influenza vaccination. Relative intrarater reliability was found to be very high, and absolute reliability was used to calculate the smallest detectable difference (SDD), which was 5.2 on the 54-point scale. This represents only about 10% of the full range of the scale, thus enabling the tool to be sensitive to other even more extreme pain behaviors in this population. Further investigations are warranted to examine whether individuals less familiar with this tool would than the first author obtain similar results when testing the same individuals twice.

The design used in **paper III** had two goals. The first goal was to examine the clinical applicability of the NCAPC among four groups of health personnel caring for individuals with IDD in order to evaluate their proficiency as potential users of this tool. The second goal was to evaluate the stability of pain measurements between different groups of health workers across measurements.

Interrater agreement between the groups of health personnel varied from poor to high, and except for the group of caregivers, paired analysis showed remarkable variability between raters within all groups. These findings are consistent with those of another study investigating paired interrater reliability for pain assessment tools (Zwakhalen et al., 2006), thus reflecting the difficulties inherent in assessing pain behavior. The nurses group was found to have the lowest level of agreement. Since this group represents higher mean age and higher professional seniority than the caregivers' group, we speculate that proximity to and daily contact with the patients, rather than professional experience, are the key factors in identifying pain behaviors in this population. In light of these findings, caregivers seem to be promising candidates for using the NCAPC.

The nurses group showed low agreement, both as a group and in paired ICC analysis. Nurses occasionally see individuals with IDD in painful situations as part of

their job, but do not spend long periods of time with individual patients on a daily basis. Therefore we recommend that nurses not be the first choice of personnel for identifying pain behaviors of individuals with IDD by means of the NCAPC unless proper training has been provided beforehand. It is speculated that better results might have been achieved by the nurses if they had felt the topic of investigation to be of greater relevance to them.

Agreement between the case managers was found to be poor, and when paired ICC analysis was performed, the variability within this group was considerable. One of the raters reduced the agreement for the whole group, yet even without this rater the agreement level was only low to moderate. We speculate that the poor agreement in this group was due to their young age and low seniority on the job, as well as their lack of familiarity with the daily behaviors of individuals with IDD. The present results suggest that case managers should also not be the first choice of personnel to use the NCAPC unless prior intensive training has been provided.

As for the therapists group, pair-wise analysis showed that the two music therapists tended to achieve very low ICC values when paired with physical or occupational therapists, but demonstrated high agreement with each other. These results are in line with the difference between the ICC (1,1) and ICC (3,1) values in the therapists group. Likewise, the two physical therapists and the occupational therapist demonstrated excellent agreement among themselves. Since therapists are trained to use observation skills in their work (Brunnekreef et al., 2005; ChihChen, 2006; Cooke et al., 2005; Edwards, 2001; Missiuna et al., 2003), they are apt to use observational scales, such as the NCAPC.

Yet, it is speculated that due to differences in the nature of their professional training, the music therapists observed different aspects of pain expression than the

physical and occupational therapists, thus explaining the different levels of pain scores and hence the poor levels of relative reliability between these two groups of therapists. Physical and occupational therapists are specifically trained to observe the physical aspects of functioning, and the use of such behavioral scales, such as the NCAPC, is consistent with their professional training. It should also be mentioned that besides the caregivers, the therapists are the only group that actually spend lengthy periods of time (i.e. therapeutic sessions) with individuals with IDD and actually might see them in painful situations. This is especially the case for physical and occupational therapists. Therefore, therapists should be considered as potential users of the NCAPC. The present findings suggest that in future investigations, raters should be grouped according to their profession.

As mentioned above, proximity is probably the main aspect explaining the fact that the caregivers and the physical and occupational therapists were found to show sufficient reliability in using the NCAPC for detecting pain behaviors of individuals with IDD. The fact that the agreement between the caregivers was slightly superior to that achieved by the therapists provides support for the assumption that observational skills are improved as more time is spent with the patients. As these are the two groups of health personnel spending the longest periods of time with this population, they are the closest to being proxy observers.

The concept of 'proxy observer' (also termed proxy rater, proxy report, proxy respondent) refers to the need to use significant others in the assessment of individuals with intellectual and/or communicative difficulties (Devies & Evans, 2001; Stancliffe, 1999; Richardson 1997; Turnbull 1999; White-Koning et al., 2005). Besides using proxy observers among individuals with IDD who have insufficient communication skills, the practice is also common in other populations lacking clear

communication abilities, such as individuals with dementia (Husebo et al., 2007) and children (Reinfjell et al., 2007). Despite methodological and ethical problems, the use of proxy observers is sometimes the only way to collect data in a non-communicative population.

Another aspect of care, common to both therapists and caregivers which might explain the high reliability found within the two groups is the physical handling of the patients. Indeed, some pain evaluation tools used in other populations are based on the assumption that physical manipulation can be used to detect pain (Husebo et al., 2007).

The fact that most groups participating in this research demonstrated poor to moderate reliability when using the NCAPC might suggest that a longer, more detailed tutorial period than the one-hour training given to raters in the present investigation should be applied in order to expand the possible users' population. Previous studies of cancer patients examining the value of an educational program for improving pain management found a significant improvement in the staff's knowledge and management of pain (Bauwens et al., 2001; Hoffmann et al., 2004; Morita et al., 2007). The Facial Action Coding System (FACS) requires that coders spend approximately 100 hours learning how to apply the scale by using self-instructional materials (Ekman et al., 1993). A shorter (2.5 hours) educational workshop was implemented by Husebo et al. (2007) to instruct staff on how to use the Mobilization-Observation-Behavior-Intensity-Dementia Pain Scale (MOBID). Since the NCAPC is a larger pain assessment tool (18 items) than the MOBID (5 items), it is estimated that a slightly longer duration program should be required to enable potential users to reach acceptable levels of reliability. The impact of such a program, if implemented in the future, should be investigated.

In **paper IV**, we examined the measurement properties of the NCAPC based on the direct observation of pain behaviors in a clinical setting. Fifty-nine adults with IDD were scored before and during a dental hygiene intervention, and the findings were compared with previous pain measurements based on scores from video recordings of the same individuals receiving an influenza vaccination. The validity of the NCAPC was demonstrated by its ability to discriminate between painful and non-painful situations, as well as between different painful experiences. High sensitivity to pain-provoking procedures was demonstrated for the participants at all levels of IDD. In both of the non-painful situations (the participant's living quarters and the dental clinic waiting room), the NCAPC scores were similarly low, despite the fact that the observations were done four years apart and in different settings. In addition, the residential observations were done through the use of video uptakes, whereas the on-the-spot scoring of participants' behavior in the dental clinic waiting room reflected a more realistic clinical situation. Despite these differences, the NCAPC demonstrated stability over the no-pain situations. The study also showed that the behavioral scores of the no-pain situations were stable over time, as scores from the first and second five-minute observation periods in the waiting room were similar and highly associated.

The NCAPC scores in non-painful situations were found to be higher than 0 in all participants, ranging between 1 and 10 on the 54-point scale. We therefore recommend that when the NCAPC is clinically implemented, a baseline score for every individual in a non-painful situation should be used for comparison purposes with behavior that is suspected as caused by pain. As shown in our study, the observation period for completing an NCAPC form in a non-painful situation does not need to exceed the duration of five minutes.

When scoring pain behaviors among individuals with IDD during two consecutive five-minute periods of dental hygiene treatment, the scores were found to be highly correlated, and a mean difference of only one point on the 0-54 scale was found between the initial and the second periods. The high correlation between the scores suggests that the participants reacted in a similar way in the first and second parts of the dental hygiene intervention. A minimal systematic shift in data was demonstrated between the first and the second observations, suggesting that the pain behavior of individuals with IDD is rather stable when observations are made in the same pain-provoking situation. The average measure of reliability was found to be higher for a 10-minute period than for a five-minute period, suggesting that more reliable results are obtained when the duration of observation is longer. The clinical implication of this finding is that when using the NCAPC, the observation for each client should last for at least 10 minutes in order to establish a sufficient unbiased stable measurement. The same duration was also suggested by the original developers of the NCCPC-R (Breau et al., 2002)

In **paper IV**, the NCAPC demonstrated the ability to distinguish between painful and non-painful situations, as well as between different levels of painful situations and settings. These findings indicate that the NCAPC is a sensitive pain assessment tool for individuals with IDD when used in a clinical situation. Similar results were found by a group of researchers evaluating pain and discomfort via the Pain and Discomfort Scale (PADS; Phan et al., 2005) in 28 individuals with IDD who had very close demographic characteristics to the present research sample. The PADS is a scale that was also originated from the NCCPC-R (Breau et. al., 2002), and the participants were undergoing the same routine dental intervention as in the present research. A future comparison of the two scales (The NCAPC and the PADS) used in

the same painful situation could lead to the construction of a new and improved scale and is therefore recommended.

Our initial expectations that the dental clinic treatment would be significantly more painful than the influenza vaccination were found correct in our study by the statistically significant different, yet correlated, scores found in the two painful situations. Our findings are supported by previous findings (Oosterink et al., 2008; Phan et al., 2005), suggesting that the dental treatment generates a stronger pain experience than an injection.

The high correlations found in **paper IV** between the pain reactions, despite the different settings and time between observations and pain stimuli, suggest that each participant reacted in a proportional manner to the two different painful stimuli. Therefore, the findings suggest that pain reactions in individuals with IDD could be considered reproducible, though this assumption should be reexamined in larger samples.

The high SRM values found in **paper IV** indicate that the NCAPC is very sensitive to pain in adults with IDD. The sensitivity of the tool was found to be highest for the group with a profound level of IDD, representing those having the most difficulty in verbally communicating their pain experience. These results correspond with our previous findings showing that the NCAPC has a high sensitivity to pain (SRM=1.20-2.07) in a large group of participants (N=228) receiving an influenza vaccination.

Differences in the observation situation used should also be taken into consideration in regard to the evaluation of pain behaviors. In the initial stages of developing the NCAPC **(papers I-III)**, video recordings were used to observe and evaluate the pain behaviors, while in **paper IV** scoring was done in a clinical

situation. The use of video recordings enables unlimited observational repetitions so that each test item can be scored separately, resulting in relatively accurate observations. In contrast, the dental hygiene treatment lasted for 20 minutes and required the simultaneous ongoing evaluation of all 18 pain behaviors on the scale related to different parts of the participant's body. This makes "live" scoring a much more challenging task than the observation of video recordings, where it is possible to focus on a few aspects of pain behavior in each repetition of the tape. Thus, when judged from the perspective of clinical usability of the scale, the reduction of items from 27 to 18 in the NCAPC seems to be justified.

4.2. METHODOLOGICAL CONSIDERATIOS

The present research addresses mainly methodological issues concerning the construction of a new pain scale, and some methodological issues that have not yet been discussed, are hereby added.

4.2.1. Internal validity

Internal validity is a term usually referred to in research designs involving treatment or Randomized Controlled Trials (RCT's), yet it could be redefined from Polit and Bech (2008) to meet the needs of the present investigation: the extent to which we can assume that results are not affected by confounding variables. Some issues relating to internal validity are raised.

Data collection: It has been found that the mere act of collecting data from people changes them (Polit and Bech, 2008). In this investigation video uptakes were done and later examined for pain behavior. It is possible that placing a camera in the room could have influenced the behavior of the participants. To minimize the change in

behavior the camera was used by the main researcher, who is known to the participants. Also while taking the video uptakes interaction with the participants was kept to a minimum as to least provoke reactions. When observing the videos the first 10 seconds of each participant were excluded as to avoid mistaking a reaction to the camera as spontaneous reaction. Another point that was considered to prevent change of participants behavior was to maintain the observed painful experiences (influenza vaccination, dental cleansing) in accordance to their natural occurrence. The issue of reaction to being filmed was partially resolved in **Paper IV** were the researcher was out of sight of the participants during all observations, therefore giving a confirmation to the clinical applicability of the scale. Nevertheless, one possibility that would have completely reduced the possibility of any influence of the camera was to implement the whole protocol with hidden cameras. This solution however would have given rise to a set of other practical, methodological as well as ethical issues.

Data analysis: In this study intrarater reliability was only examined by the primary author (ICC at 0.94). As he was the most involved in the development of the scale and familiar with the different items, there is a possibility that a less experience examiner might come up with less impressive results. One might refute this notion by mentioning the fact that the caregivers and therapists examining interrater reliability came up with very good results as well (ICC at 0.91 and 0.92). Despite the fact that having the main researcher examining his results is probably a common protocol, this assumption could be examined in future research

4.2.2. External validity

External validity deals with the possibility of generalizing the results of the present investigation to other conditions (Domholdt, 2005) such as similar

populations (children and adolescents with IDD), settings (hospital), and types of pain (chronic, long-lasting), and the use of others aside the main researcher. When evaluating the present research project we find that characteristics of the sample selected for the research intervention resemble those of the population of adults with IDD in Israel regarding residence, types of disability and the occurrence of gender, age and level of IDD (Merrick, Kandel, 2003). Moreover, after construction of the NCAPC using the total research sample (**Paper II**), all the tests leading to the construction of the NCAPC (sensitivity of items, internal consistency, sensitivity of the scale) were reexamined in a random sample of 89 participants. Therefore we can assume that the NCAPC is valid for assessing procedural pain in the population of adults at all levels of IDD in Israel, both through direct observation and through video recording. External validation of the NCAPC regarding different age groups (children, adolescents), different types of pain (chronic), and ability to evaluate the effect of pain management, are pending future research.

4.2.3. Examining the quality of the NCAPC

Terwee, et al., (2007) suggest that measurement properties of health status measure should be evaluated according to specific quality criteria, which could be marked as positive (+), indeterminate (0), poor (-) or no information available (?) (Table 1).

4.2.3.1 Reproducibility (reliability)

Satisfactory absolute and relative measurement error is required to get a full mark on this issue (Terwee, et al. 2007). ICC values were higher than 0.7, which is considered satisfactory, and standard error of measurement (SEM) and the Smallest

Detectable Change (SDC) were reported. Although we did not examine Minimal Important Change (MIC) in relation to SDC, we consider the outcome of this measurement property satisfactory, giving a positive mark (+).

4.2.3.2 Validity

Content validity - In order for this criterion to be fulfilled as proposed by Terwee et al. (2007) there must be a clear description of the measurement aim, the concept being measured, the item selection and target population (i.e. adults with IDD). Since all these requirements seem to be met regarding the NCAPC, a positive score (+) is suggested.

Criterion validity – Self-reports of pain are considered "gold standard" of pain. Since self-reports of pain in people with IDD lack validity, this aspect of validity is difficult to examine, and was not addressed in the present research. We suggest no information on criterion validity and would like to comment that criterion validity might be impossible to examine in this population. Therefore this criterion can be marked as (0).

Construct validity – Construct validity was not examined regarding the NCAPC by formulating specific hypothesis regarding the relationship with other measures, as suggested by Terwee et al. (2007). However, the construct for the NCAPC was based on thorough work by previous researchers, when establishing the mother scale, the NCCPC, intended for children (Breau, et al., 2000; Breau, et al., 2002A; Breau, et al., 2002B; Breau, et al., 2003). As for now there is not enough evidence to fully support the construct validity of the NCAPC. Therefore this criterion

can be marked as (?), meaning no information available, or rather not sufficient information available.

Internal consistency – was examined during test development and of the final NCAPC scale, using the Cronbach's alpha test. A bootstrapping procedure (1000 samples) was also used in order to achieve robust results of internal consistency. Alpha was found within recommended boundaries of 0.7-0.9. Factor analysis of the measure is also recommended, and has in fact been performed, but not yet published within the scope of the four articles included within the present thesis. Therefore this criterion can be marked as (0), meaning intermediate

Responsiveness should be examined according to Terwee et al. (2007) by the receiver operating characteristics (ROC) curve analysis, which is designed to examine the ability of a measure to discriminate between patients who have vs. have not improved. Only sensitivity to change was addressed to examine the ability of the NCAPC to capture change in individuals at all levels of IDD. Therefore this criterion can be marked as (?), meaning no information available.

Floor and ceiling effects are considered to be present if more than 15% of respondents achieved the lowest or highest possible score (McHorney, & Tarlov, 1995). No subject reached a full score (54) in the present investigation and only 1 person out of the total sample (N=228; 0.4%) showed a floor effect. Therefore this criterion can be marked as positive (+).

Interpretability is defined as the degree to which one can assign qualitative meaning to quantitative scores (Lohr, et al., 1996). Terwee, et al., (2007) suggest that a positive rating should be given to a tool if mean scores and SD are presented at least

in four subgroups, and this was presented in **Paper II**. Other information regarding interpretability such as the SDD and norm values for the population as well as subgroups within the population was presented. On the other hand, minimal important change (MIC) to enable interpretation of change scores over time and sample size calculations has not yet been examined, as suggested by Terwee et al., (2007). Therefore this criterion can be marked 0, meaning intermediate, as not enough information is available.

Table 1 – Estimating the quality of the NCAPC in adults with IDD as suggested by Terwee, et al., (2007).

Criteria	Content validity	Internal consistency	Criterion validity	Construct validity	Reproducibility		Responsiveness,	Floor and ceiling effects	Interpretability
					Agreement	Reliability			
Rating	+	**0**	**0**	**?**	+	+	**?**	+	**0**
Index: (+) positive; (0) intermediate; (-) poor; (?) no information available									

The evaluation of measurement properties of NCAPC (Table 1), based on the criteria of Terwee et al. (2007), indicates that many of the basic criteria (content validity, internal consistency, and the floor and ceiling effects) can be considered satisfactory, while others (construct validity, responsiveness, reproducibility and interpretability) are still lacking and pending future investigation.

4.3. LIMITATIONS OF THE PRESENT INVESTIGATION

As in all research investigations the present research project was showing limitations. The basic limitation might be the use of procedural pain (both influenza injection and dental hygienist intervention). Since procedural pain has specific characteristics (American Academy of Pediatrics, 2001) such as the readiness of the participant towards the intended stimulus, generalization to other types of pain (e.g. chronic, long-lasting, post surgical pain) is pending future investigations. Moreover, the mild, acute phasic pain that characterizes influenza vaccinations is very specific and leads to pain behaviors that might be different from more severe pain experiences. Therefore the applicability of the scale to other types of pain and to more severe levels of pain should be investigated in the future.

Another limitation lies in the presence of confounding elements usually accompanying pain behaviors such as anxiety and discomfort. The measurements described in the present study represent the accumulated reaction of the participants during two different pain experiences. We presume that behaviors gathered and collected by the NCAPC, also include reactions such as discomfort and anxiety. The distinction between these subtle nuances of the total pain reaction is difficult to establish in individuals with IDD (Phan, et al., 2005), as well as in individuals with full mental and verbal capacities (Sokol, Sokol, & Sokol, 1985). In **paper IV** the scale was tested under clinical conditions, and therefore different limitations specific to such situations arose. For instance, some of the participants were given 5 ml of relaxing medication (valium) prior to dental hygiene treatment to reduce resistance to the procedure. This was done to specific individuals according to their past behaviors during such treatment. It is probable that the medication had a distorting effect on the results, reducing the natural reaction of individuals with severe and profound levels of

IDD who tend to react strongly to the dental clinic procedures. Despite this medication, the NCAPC was found sensitive to pain. In **Paper IV** some restrictions prevented the researcher from viewing the full scope of pain behaviors by the participants. These restrictions were holding of the head of some of the participants by the accompanying caregiver, during dental hygiene treatments, restraining the body of some of the participants, hands of the dental hygienist, and of the caregiver prevented a clear view of the face and mouth of the participants. All these events around the participants certainly reduced natural head, body and limb movements and might have prevented observation of items such as protective movements, which have been found to be extremely important when evaluating pain behaviors in previous investigations (Lotan, Ljunggren , et al., 2009). Yet, since these hindrances are part of true clinical situations, it seems worthwhile that the NCAPC was tested under such conditions.

Despite the loss of much information with regards to intensity of movement, since the NCAPC is aimed at recording duration rather than intensity of pain behaviors, we believe that sufficient information was obtained to enable valid pain assessment. Moreover, since many of those restricting elements were applied to the majority of clients, the gathering of pain behaviors was, in a way, standardized. Another point to consider is that this procedure was done in a real clinical situation and it represents common clinical limitations to pain measurements during dental hygiene treatments.

5. CONCLUSIONS

For many years, clinicians did not fully grasp the phenomenon of pain behavior assessment and management in individuals with IDD. Without a substantial knowledge base, practitioners were left to rely on their own subjective judgments. These were fraught with inconsistencies and personal biases, which typically led to undertreatment of pain. Now that the old "myths" have been refuted by solid research data (Schechter, 1989; Schechter et al., 1993), it becomes obvious that practitioners should provide care that reflects current advancement of knowledge in the field. Increasingly, guidelines for the assessment and treatment of pain are being published to educate and direct the health care professionals (Agency for Health Care Policy and Research, 1992; American Pain Society, 1992; Berde et al., 1990; Zeltzer, et al., 1990), and these must be implemented in everyday practice.

It is therefore imperative that each individual caregiver feel the personal responsibility to attend to and alleviate symptoms of pain and distress. This would include the careful assessment of pain and ongoing management to assure that discomfort is minimal. With new scientific data available, including the body of knowledge gathered and developed within the current investigation, both scientific and ethical standards demand that the undertreatment of pain be rejected as substandard practice for individuals with IDD at all ages and at all levels of IDD.

The NCAPC is a scale suitable for detecting pain in adults with IDD. This scale has been found to hold high internal consistency, very high intra and interreliability, high sensitivity and ability to distinguish between pain and non-pain situations and different levels of pain. The NCAPC, if properly used, will hopefully contribute to reduce pain suffering in individuals with IDD, thereby improving their quality of life.

5.1. PERSPECTIVES TOWARDS FUTURE RESEARCH

In past years, the clinical and academic milieus have progressed beyond the traditional understanding that the absence of verbalized complaints of pain do not equate to the absence of pain. The scientific world now recognizes the strong influence that physical limitations, life experiences, level of education and cognitive status has on patients' ability to clearly report and accurately describe their pain experiences.

Further research is warranted in order to identify more characteristics of pain behavior in this population, which in turn could be used in diagnostic profiles or clinical diagnosis of pain in individuals with IDD. This could be especially relevant in individuals with profound and severe levels of IDD, where common ways of communication are absent. This group of clients also presents multiple medical conditions and possible sources of pain, which can result in delays in identifying the current cause of pain, causing untreated illness and suffering for the client with IDD.

The NCAPC is a new scale, and as such has not yet been fully examined for measurement properties. In order to ensure the strength and usability of the scale, all aspects of reliability, validity, responsiveness, clinical utility and feasibility should be examined. Such inquiries should be performed in order to ensure that the NCAPC has the ability to accurately assess pain in individuals with IDD. Future studies should be performed to investigating the association between the NCAPC against similar scales (such as the PADS), examine it's ability to differentiate anxiety related behaviors from pain behaviors, testing the ability of the scale to detect different types of pain (such as pain in individuals with chronic illness, post-operation pain, pain from extremely harmful medical procedures), validating the NCAPC in different settings (such as a hospital), examine the sensitivity of the scale to differentiate between pain

reaction with and without medication, and examine the applicability of the scale in other populations (children, adolescents).

An extremely important part would be to introduce the scale to the clinical field. There are many examples for scales that have been constructed yet have not been clinically used after the construction procedure have ended. Therefore a clinical trial should be implemented in order to suggest initial guidelines for the clinical use of the scale. Possible issues involved in introducing the NCAPC for clinical use would be trials with different target populations, suggestion of potential users, establishing a tutorial program for the use of the NCAPC).

6. REFERENCE LIST

Abu-Saad HH. Challenge of pain in the cognitively impaired. Lancet 2000;(356): 1867-8.

Agency for Health Care Policy and Research. Acute pain management: Operative or medical procedures and trauma. Rockville, MD, 1992.

Allen KD, Elliott AJ, Arndorfer RE. Behavioural pain management for paediatric headache in primary care. Child Health Care 2002;31:175–89.

American Academy of Pediatrics Committee on Psychosocial Aspects of Child and Family Health & American Pain Society Task Force on Pain in Infants, Children, and Adolescents the Assessment and Management of Acute Pain in Infants, Children, and Adolescents. Pediatrics 2001;108(3):793-7.

American Educational Research Association, American Psychological Association, National Committee on Measurement in Education. Standards for educational and psychological testing. Washington, DC.: American psychological association, 1999.

American Pain Society. Principles of analgesic use in the treatment of acute pain and cancer pain (3rd ed.). Skokie, IL:Author, 1992.

American Pain Society. Principles of analgesic use in the treatment of acute pain and cancer pain, 5th ed. American Pain Society, Glenview, IL, 2003.

Bauwens S, Distelmans W, Storme G, Kaufman L. Attitudes and knowledge about cancer pain in Flanders. The educational effect of workshops regarding pain and symptom control. Palliat Med 2001;15:181-9.

Beange H, McElduff A, Baker W. Medical disorders of adults with mental retardation: A population study. Am J Ment Retard 1995;99:595– 604.

Berde C, Ablin A, Glazer J, Miser A, Shapiro B, Weisman S, Zeltzer P. (1990). Report of the Subcommittee on Disease-Related Pain in Childhood Cancer. Pediatrics 1990;86:818-25.

Berven S, Bradford DS. Neuromuscular scoliosis: causes of deformity and principles for evaluation and management. Semin Neurol 2002;22:167-78.

Beyer JE, Wells N. The assessment of pain in children. Pediatr Clin North America 1989;36:837–54.

Biersdorff KK. Incidence of significantly altered pain experience among individuals with developmental disabilities. Am J Ment Retard 1994;98:619–31.

Bittles AH, Petterson BA, Sullivan SG, Hussain R, Glasson EJ, Montgomery PD. The influence of intellectual disability on life expectancy. J Gerontol A Biol Sci Med Sci 2002; 57:M470–M2.

Bland JM., Altman DG. Statistics notes: Measurement error. BMJ 1996;313:744.

Blount RL, Pira T, Cohen LL, Cheng PS. Pediatric procedural pain. Behav Modif 2006; 30(1):24-49.

Bodfish JW, Harper VN, Deacon JR, Symons FJ. Identifying and measuring pain in persons with developmental disabilities: a manual for the Pain and Discomfort Scale (PADS). Western Carolina Center Research Reports, 2001.

Bonica JJ. Cancer pain: Importance of the problem. In: Bonica JJ, Ventafrida V. (Eds.). Advances in Pain Research and Therapy Vol. 2(pp.1-12). New York: Raven press, 1979.

Breau LM, McGrath PJ, Craig KD, Santor D, Cassidy KL, Reid GJ. Facial expression of children receiving immunizations: a principal components analysis of the child facial coding system. Clin J Pain 2001;17:178–86.

Breau LM, McGrath PJ, Camfield C, Rosmus C, Finley GA. Preliminary validation of an observational pain checklist for persons with cognitive impairments and inability to communicate verbally. Dev Med Child Neurol 2000;42:609–16.

Breau LM, Finley GA, McGrath PJ, Camfield CS. Validation of the Non-communicating Children's Pain Checklist-Postoperative Version. Anesthesiology 2002A;96:528–35.

Breau LM, McGrath PJ, Camfield CS, Finley GA. Psychometric properties of the non-communicating children's pain checklist-revised. Pain 2002B;99:349–57.

Breau LM, Camfield CS, McGrath PJ, Finley A. The incidence of pain in children with severe cognitive impairments. Arch Pediatr Adolesc Med 2003;157:1219–26.

Breau LM, Camfield CS, McGrath PJ, Finley GA. Risk factors for pain in children with severe cognitive impairments. Dev Med Child Neurol 2004;46:364–71.

Breau LM, McGrath PJ, Zabela M. Assessing pediatric pain and developmental disabilities. In: Oberlander TF, Symons FJ (Eds). Pain in children & adults with developmental disabilities. Baltimore: Paul H. Brookes Publishing Co, 2006:149-72.

Brennan F, Carr DB, Cousins M. Pain management: A fundamental human right. Anesth Analg 2007;105:205-21.

Britto MT, Kotagal UR, Chenier T, Tsevat J, Atherton HD, Wilmott RW. Differences between adolescents' and parents' reports of health-related quality of life in cystic fibrosis. Pediatr Pulmonol 2004;37:165–71.

Brunnekreef JJ, van Uden CJT, van Moorsel S, Kooloos JGM. Reliability of videotaped observational gait analysis in patients with orthopedic impairments. BMC Musculoskelet Disord 2005;6:17.

Burns N, Grove SK. The Concepts of measurement. In: Burns N, Grove SK. (Eds). The practice of nursing research, 3rd ed. Philadelphia: W.B. Saunders Company, 1997: 319–41.

Carr EG, Jamie Æ, Owen-DeSchryver S. Physical illness, pain, and problem behavior in minimally verbal people with developmental disabilities. J Autism Dev Disord 2007;37:413–24.

Carrougher GJ, Ptacek JT, Honari S, Schmidt AE, Tininenko JR, Gibran NS, Patterson DR. Self-reports of anxiety in burn-injured hospitalized adults during routine wound care. J Burn Care Res 2006;27:676–81.

Carter G, Jancar J. Sudden death in the mentally handicapped. Psychol Med 1984;14: 691-5.

CarterB, McArthur E, Cunliffe M. Dealing with uncertainty: parental assessment of pain in their children with profound special needs. J Adv Nurs 2002;38,449–57.

ChihChen SL. Music therapy in Taiwan, 2003-2006. Accessed 2008 May http://www.voices.no/country/monthtaiwan_november2006.php

Clements J. I can't explain…''Challenging behaviour:'' towards a shared conceptual framework. Clin Psychol Forum 1992 January; 29–37.

Cohen L L, Blount RL, Cohen RJ, Schaen ER, Zaff J. Comparative study of distraction versus topical anesthesia for pediatric pain management during immunizations. Health Psychol 1999; 18: 591-8.

Collignon P, Giusiano B, Boutin AM, Combes JC. Utilisation d'une échelle d'hétéroévaluation de la doleur chez le sujet sévérement polyhandicapé. Douleur et Analgesie 1997;1:27–32.

Collignon P, Giusiano B. Validation of a pain evaluation scale for patients with severe cerebral palsy. Eur J Pain 2001;5:433–42.

Cooke DM, McKenna K, Fleming J. Development of a standardized occupational therapy screening tool for visual perception in adults. Scand J Occup Ther 2005;12: 59-71.

Cooper SA. Clinical study of the effects of age on the physical health of adults with mental retardation. Am J Ment Retard 1998;102:582-9.

Craig D. Clinical pain measurement from the perspective of the human laboratory. In: Chapman CR, Loeser JD. (Eds). Advances in Pain Research and Therapy. Volume 12, Issues in Pain Measurement, 1989.

Craig KD. On knowing another's pain. In: Ekman P, Rosenberg E. (Eds). What the face reveals. Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). Oxford: Oxford University Press, 1998:180–7

Craig KD, Grunau RVE, Aquan-Assee J. Judgment of pain in newborns: facial activity and cry as determinants. Can J Behav Sci/Rev Can Sci Comp 1988;20:442–51.

Craig KD, Hadjistavropoulos HD, Grunau RV, Whitfield MF. A comparison of two measures of facial activity during pain in the newborn child. J Pediatr Psychol 1994; 19:305–18.

Craig KD, Hyde SA, Patrick CJ. Genuine, suppressed and faked facial behavior during exacerbation of chronic low back pain. Pain 1991;46:161–71.

Craig KD, Korol CT, Pillai RR. Challenges of judging pain in vulnerable infants. Clin Perinatol 2002; 29:445–57.

Craig KD, McMahon RJ, Morison JD, Zaskow C. Developmental changes in infant pain expression during immunization vaccinations. Soc Sci Med 1984;19:1331–7.

Craig KD, Prkachin KM, Grunau R. The facial expression of pain. In: Turk D, Melzack R. (Eds). Handbook of pain assessment. New York: Guilford Press, 1992: 257–76

Currier DP. Elements of research in physical therapy. Baltimore: Williams & Wilkins, 1990.

Defrin R, Lotan M, Pick CG. The evaluation of acute pain in individuals with cognitive impairment: a differential effect of the level of impairment. Pain 2006;124 (3):312-20.

De Lissovoy V. Head banging in early childhood. Child Dev 1962;33:43–56.

Developmental Disabilities Assistance and Bill of Rights Act of 2000. (114 STAT. 1684, PUBLIC LAW 106–402—OCT. 30, 2000). Accessed 2006 Feb 02. http://frwebgate.access.gpo.gov/cgi-bin/getdoc. cgi?dbname=106_cong_public _laws&docid=f:publ402.106.pdf.

Devies D, Evans L. Assessing pain in people with learning disabilities. Br J Nurs 2001;10(8):513-6.

Domholdt E. Physical therapy research: principles and applications. Philadelphia: WB Saunders Company, 2000:15-20.

Donovan JD. Learning disability nurses' experiences of being with clients who may be in pain. J Adv Nurs 2002;38:458–66.

Duhn LJ, Medves JM. A systematic integrative review of infant pain assessment tools. Adv Neonatal Care 2004;4(3):126-40.

Durvasula S, Beange H. Health inequalities in people with intellectual disability: Strategies for improvement. Health Promot J 2001 Aug;11:27–31.

Deyo RA, Centor RM Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. J Chronic Dis 1986;39(11):897-906.

Hadden KL, von Baeyer CL. Pain in children with cerebral palsy: common triggers and expressive behaviors. Pain 2002;99:281–8.

Edwards MM. The Betty Project: Jump-starting clinical observation skills in an OTA Curriculum. Occupational Therapy in Health Care 2001;15:99-104.

Ekman P, Huang TS, Sejnowski TJ. Hager J. Final report to NSF of the planning workshop on facial expression understanding. San Francisco, CA; Human Interaction Lab, University of California, 1993.

Ekman P, Friesen W. Investigators guide to the facial action coding system. Palo Alto, CA: Consulting Psychologist Press, 1978.

Ennett ST, DeVellis BM, Earp JA, Kredich D, Warren RW, Wilhelm CL. (1991) Disease experience and psychosocial adjustment in children with juvenile rheumatoid arthritis: children's versus mothers' reports. J Pediatr Psychol 1991;16:557–68.

Fanurik D, Koh JL, Harrison RD, Conrad TM, Tomerlin C. Pain assessment in children with cognitive impairment. An exploration of self-report skills. Clin Nurs Res 1998; 7:103–24.

Fanurik D, Koh JL, Schmitz ML, Harrison D, Roberson PK, Killebrew P. Pain assessment and treatment in children with cognitive impairment: a survey of nurses' and physicians beliefs. Clin J Pain 1999; 15**:**304–12.

Feldt KS, Ryden MB, Miles S. Treatment of pain in cognitively impaired compared with cognitively intact older patients with hip fracture. J Am Geriatr Soc 1998;46: 1079–85.

Ferrell BA, Ferrell BR, Rivera L. Pain in cognitively impaired nursing home patients. J Pain Symptom Manage 1995;10:591–8.

Finch E, Brooks D, Stratford PW, Mayo NE. Physical rehabilitation outcome measures: A guide to enhanced clinical decision making. 2nd. Hamilton, BC Decker Inc., 2002.

Fisher K. Health disparities and mental retardation. J Nurs Scholarsh 2004;36:48–53.

Fradet C, McGrath PJ, Kay J, Adams S, Luke B. A prospective survey of reactions to blood tests by children and adolescents. Pain 1990;40:53-60.

Frank NC, Blount RL, Smith AJ, Manimala MR, Martin JK. Parent and staff behavior, previous child medical experience, and maternal anxiety as they relate to child procedural distress and coping. J Pediatr Psychol 1995;20:277-89.

Gauthier JC, Finley GA, McGrath PJ. Children's self-report of postoperative pain intensity and treatment threshold: determining the adequacy of medication. Clin J Pain 1998;14(2):116-20.

Gilbert-MacLeod CA, Craig KD, Rocha EM, Mathias MD. Everyday pain responses in children with and without developmental delays. J Pediatr Psychol 2000;25:301–8.

Giusiano B, Jimeno MT, Collignon P, Chau Y. Utilization of neural network in the elaboration of an evaluation scale for pain in cerebral palsy. Methods Inf Med 1995; 34:498–502.

Goodenough B, Kampel L, Champion GD, Laubreaux L, Nicholas MK, Ziegler JB, McInerney M. An investigation of the placebo effect and age-related factors in the report of needle pain from venipuncture in children. Pain 1997;72:383-91.

Goodenough B, Thomas W, Champion GD, Perrott D, Taplin JE, von Baeyer CL, Ziegler JB. Unraveling age effects and sex differences in needle pain: Ratings of sensory intensity and unpleasantness of venipuncture pain by children and their parents. Pain 1999;80:179-90.

Gorelick MH, Stevens MW, Schultz TR. Comparability of acute asthma severity assessments by parents and respiratory therapists. Arch Pediatr Adolesc Med 2002; 156:1199–202.

Groth-Marnat, G. Handbook of psychological assessment. 3rd edition, Hoboken: NJ, John Wiley & Sons, Inc, 1997:21–2.,

Gunnar MR, Porter FL, Wolf CN, Rigaruso J, Larson MC. Neonatal stress reactivity: Prediction to later emotional temperament. Child Dev 1995;66:1-13.

Grunau RV, Craig KD. Pain expression in neonates: facial action and cry. Pain 1987;28:395–410.

Guyatt GH, Feeny, DH, Patrick DL. Measuring health related quality of life. Ann Intern Med 1993;118:622-9.

Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A. (1989). Responsiveness and validity in health status measurement: a clarification. J Clin Epidemiol 1989;42: 403-8.

Guzmán J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary bio-psycho-social rehabilitation for chronic low back pain. Cochrane Database Syst Rev 2002; 1: CD000963.

Hadjistavropoulos T, Craig KD, Martin N. Towards a research outcome measure of pain in frail elderly in chronic care. Pain Clinician 1997;10:71-9.

Hadjistavropoulos T, Von Baeyer C, Craig KD. Pain assessment in persons with limited ability to communicate. In: Turk DC, Melzack R. (Eds). Hand Book of Pain Assessment. 2nd edition, New York: The Gilford press, 2001:134-52

Halimaa SL. Pain management in nursing procedures on premature babies. J Adv Nurs 2003;42(6):587-97.

Hart H, Bax M, Jenkins S. Health and behavior in preschool children. Child Care Health Dev 1984;10:1–16.

Henderson RC, Lark RK, Gurka MJ, Worley G, Fung EB, Conaway M, Stallings VA, Stevenson RD. Bone density and metabolism in children and adolescents with moderate to severe cerebral palsy. Pediatrics 2002;110(1 pt 1):e5- e5.

Hensel E, Rose J, Kroese BS, Banks-Smith, J. Subjective judgments of quality of life: A comparison study between people with an intellectual disability and those without a disability. J Intellect Disabil Res 2002;46:95–107.

Herr KA, Spratt K, Mobily PR, Richardson G. Pain intensity assessment in older adults: use of experimental pain to compare psychometric properties and usability of selected pain scales with younger adults. Clin J Pain 2004;20:207–19.

Hollins S, Attard MT, von Fraunhofer N, McGuigan S, Sedgwick P. Mortality in people with learning disability: Risks, causes, and death certification findings in London. Dev Med Child Neurol 1998;40:50 –6.

Hoffmann W, Munzinger H, Horstkotte E, Greiser E. A population-based evaluation of an intervention to improve advanced stage cancer pain management. J Pain Symptom Manage 2004;28:342-50.

Hunt A. Goldman A. Seers K, Crichton N, Mastroyannopoulou K, Moffat V, Oulton K, Brady M. Clinical Validation of the Paediatric Pain Profile. Dev Med Child Neuorol 2004; 46(1):9-18.

Hurley BAC, Volier BJ, Harnahan PA. Assessment of discomfort in advanced Alzheimer patients. Res Nurs Health 1992;15:369-77.

Husebo BS, Strand LI, Moe-Nilssen R, Husebo SB, Snow AL, Ljunggren AE. Mobilization-Observation-Behavior-Intensity Dementia Pain Scale (MOBID): Development and validation of a nurse-administered pain assessment tool for use in Dementia. J Pain Symptom Manage 2007;34(1):67-80.

Individuals with Disabilities Education Act. (Public Law 101-476) Mental retardation: definition, classification, and systems of supports. 9th ed. Washington, D.C.: American Association on Mental Retardation, 1992. Accessed 2006, Feb 02. http://thomas.loc.gov/cgi-bin/bdquery/z?d101:SN01824:@@@D&summ2= m&|TOM:/bss/d101query.html|.

International Association for the Study of Pain – IASP, (2009) Ethical guidelines for pain research in humans. Accessed April 22nd 2009. URL: http://www.iasp-pain.org/AM/Template.cfm?Section=General_Resource_Links&Template=/CM/HTMLDisplay.cfm&ContentID=3052

Jancar J, Speller CJ. Fatal intestinal obstruction in the mentally handicapped. J Intellect Disabil Res 1994;38:413–22.

Janicki MP, Dalton AJ, Henderson CM, Davidson PW. Mortality and morbidity among older adults with intellectual disability: Health services considerations. Disabil Rehabil 1999;21:284 –94.

Johnston CC. Psychometric issues in the measurement of pain. In: Finley GA, McGrath J (Eds). Measurement of Pain in Infants and Children. Progress in Pain Research and Management. Vol. 10. Seattle, WA: IASP Press, 1998: 5–20.

Kaasalainen S. Pain assessment in older adults with dementia: using behavioral observation methods in clinical practice. J Gerontol Nurs 2007;33(6):6-10.

Kerns RD. Turk DC, Rudy, TE. The West Haven-Yale Multidimensional Pain Inventory (WHYMPI), Pain 1985 Dec; 23(4):345–56.

Kerr M, Fraser W, Felce D. Primary health care for people with a learning disability: A keynote review. Br J Learn Disabil 1996;24:2– 8.

Kirschner B, Guyatt G. A methodological framework for assessing health indices. Journal of Chronic Disease 1985;38:27-36.

Krahn GL, Hammond L, Turner A. A cascade of disparities: Health and health care access for people with intellectual disabilities. Ment Retard Dev Disabil Res Rev 2006;12:70–82.

Krauss MW, Gulley S, Sciegaj M, Wells N. Access to specialty medical care for children with mental retardation, autism, and other special health care needs. Ment Retard 2003;41(5):329-39.

Lachapelle DL, Hadijistavropoulos R, Craig KD. Pain Measurement in Persons with Intellectual Disability. Clin J Pain 1999;15:13-23.

Latham J, Davis BD. The socio-economic impact of chronic pain. Disabil Rehabil 1994;16(1):39-44.

Lawrence J, Alcock D, McGrath P Kay J, MacMurray SB, Dulberg C. The development of a tool to assess neonatal pain. Neonatal Network 1993;12: 59–66.

LeResche L, Dworkin SF. Facial expression accompanying pain. Soc Sci Med 1984;19:1325–30.

LoBiondo-Wood G, Haber J. Reliability and Validity. In: LoBiondo-Wood G, Haber J, (Eds). Nursing Research: Methods, Clinical Appraisal and Utilization, 4th ed.. St. Louis: Mosby, 1998: 327–50.

Lohr KN, Aaronson NK, Alonso J, Burnam MA, Patrick,DL, Perrin EB, Roberts JS.

Evaluating quality of life and health status instruments: Development of scientific review criteria. Clinical Therapy 1996;18:979-92.

Lotan M, Ljunggren AE, Johnsen TB, Defrin R, Pick CG, Strand LI.. A modified version of the Non-Communicating Children Pain Checklist-Revised (NCCPC-R), adapted to adults with intellectual and developmental disabilities. Sensitivity to pain and internal consistency. Journal of Pain 2009;10(4):398-407.

Lotan M, Moe-Nilssen R, Ljunggren AE, Strand LI. Reliability of the Non-Communicating Adult Pain Checklist (NCAPC), assessed by different groups of health workers. Res Dev Disabil 2009 Jul-Aug;30(4):735-45.

Lotan M, Moe-Nilssen R Ljunggren AE, Strand LI. Measurement properties of the Non-Communicating Adult Pain Checklist (NCAPC): a pain scale for adults with Intellectual and Developmental Disabilities, used in a clinical setting
Journal of Intellectual and Developmental disabilities (JIDD). Submitted.

Luckasson R, Borthwick-Duffy S, Buntinx WHE, Coulter DL, Craig EM, Reeve A, Shalock RL, Snell ME, Spitalnik DM, Spreat S, Tassé MJ. Mental retardation: Definition, classification, and systems of supports. 10th ed. Washington (DC): American Association on Mental Retardation, 2002.

Mafrica F, Schifilliti D, Fodale V. Pain in Down's syndrome. TheScientificWorld Journal 2006;6:140–7.

Malviya S, Voepel-Lewis T, Trait AR, Merkel S, Lauer A, Munro H, Farley F. Pain management in children with and without cognitive impairment following spine fusion surgery. Pediatric Anesthesia 2001;11:453–8.

Malviya S, Voepel-Lewis T, Burke C, Merkel S, Tait AR. The revised FLACC observational pain tool: improved reliability and validity for pain assessment in children with cognitive impairment. Paediatr Anaesth 2006;16:258-65.

Marchette, L, Main R, Redick E, Baggs A, Leatherland J. Pain reduction interventions during neonatal circumcisions. Nurs Res 1991;40:241-4.

Mason J, Scior K. Diagnostic overshadowing amongst clinicians working with people with intellectual disabilities in the UK. J Appl Res Intellect Disabil 2004;17:85–90.

Mata AC. Management of acute abdomen it the mentally retarded. Va Med Mon 1960;96:732-3.

McCaffery M, Beebe A. Pain clinical manual for nursing practice. St. Louis, Mo. The C.V. Mosby Company, 1989.

McGrath PJ. Behavioral measures of pain. In: Finley GA, McGrath PJ. (Eds). Measurement of pain in infants and children. Progress in Pain Research and Management. Vol. 10. Seattle, WA, IASP Press, 1998: 83–102

McGrath PA. Psychological aspects of pain perception. Arch Oral Biol 1994;39:55-62.

McGrath PJ, Breau LM, Camfield C, Finley GA. Caregivers management of pain in cognitively impaired children with severe speech impairments. Poster presented at the 5th international symposium on pediatric pain. London, England. June 18-21, 2000.

McGrath PA, Brigham MC. The assessment of pain in children and adolescents. In: Turk DC, Melzack R. (Eds). Handbook of pain assessment. New York, NY: Guilford Press, 1992: 295–314

McGrath PJ, Johnson G, Goodman JT, Schillinger J, Dunn J, Chapman J. CHEOPS: A behavioral scale for rating postoperative pain in children. In: Fields HL. Dubner R, Cervero F (Eds). Advances in Pain Research and Therapy. Vol. 9. New York, Raven Press, 1985: 395–402.

McGrath PJ, Rosmus C, Canfield C, Campbell MA, Hennigar A. Behaviours caregivers use to determine pain in nonverbal, cognitively impaired individuals. Dev Med Child Neuro 1998;40:340–3.

McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Quality of Life Research 1995;4:293-307.

Melzack R. The McGill Pain Questionnaire: Major properties and scoring methods, Pain 1975;1:119–33.

Melzack R, Wall PD. Handbook of pain management. New York, NY, Churchill Livingstone, 2003.

Merrick J, Kandel I. Medical services for persons with intellectual disability in Israel. Public Health Review 2003;3:45-68.

Merrick J, Cahana C, Lotan M, Kandel I, Carmeli E. Snoezelen or controlled multi-sensory stimulation. Treatment aspects from Israel. TheScientificWorldJournal 2004; 4:307-14.

Mersky H, Bogduk N. (Eds). Classification of chronic pain: descriptions of chronic pain syndromes and definitions of pain terms. Seattle, WA: IASP Press, 1994.

Merskey H, Bogduk N. Taxonomy of chronic pain. Seattle WA: IASP Press, 1992.

Miller D. Comparisons of pain ratings from postoperative children, their mothers, and their nurses. Pediatr Nurs 1996;22:145–9.

Miniham PM, Dean DH. Meeting the needs for health services of persons with mental retardation living in the community. Am J Pub health 1990;80:1043-48.

Miniham PM, Dean DH, Lyons CM. Managing the care of patients with mental retardation: A survey of physicians. Ment Retard 1993;31:239-46.

Missiuna C, Rivard L, Bartlett D. Early identification and risk management of children with developmental coordination disorder. Pediatr Phys Ther 2003;15:32-8.

Moe-Nilssen, R. Test-retest reliability of trunk accelerometry during standing and walking. Arch Phys Med Rehabil 1998; 79,1377-85.

Moe-Nilssen R, Nordin E, Lundin-Olsson L. Criteria for evaluation of measurement properties of clinical balance measures for use in fall prevention studies. J Eval Clin Pract 2008 Apr;14(2):236-40.

Morita T, Murata H, Hirai K, Tamura K, Kataoka J, Ohnishi H, Akizuki N, Kurihara Y, Akechi T, Uchitomi Y. Meaninglessness in terminally ill cancer patients: a validation study and nurse education intervention trial. J Pain Symptom Manage 2007;34:160-70.

Oberlander TF, O'Donnell ME, Montgomery C.J. Pain in children with significant neurological impairment. J Dev Behav Pediatr 1999;20:235–43.

Oberlander TF, O'Donnell ME. Beliefs about pain among professionals working with children with significant neurological impairment. Dev Med Child Neurol 2001;43: 136–40.

Oosterink FMD, de Jongh A, Aartman I.HA. What are people afraid of during dental treatment? Anxiety-provoking capacity of 67 stimuli characteristic of the dental setting. European Journal of Oral Science 2008;116:44–51.

Pate JT, Blount RL, Cohen LL, Smith AJ. Childhood medical experience and temperament as predictors of adult functioning in medical situations. Child Health Care 1996;25:281-96.

Parmelee PA. Pain in cognitively impaired older persons. Clin Geriatr Med 1996;12: 473–87.

Pfister AA, Roberts AG, Taylor HM, Noel-Spaudling S, Damian MM, Charles PD. Spasticity in adults living in a developmental center. Arch Phys Med Rehabil 2003; 84:1808-12.

Phan A, Edwards CL, Robinson EL. The assessment of pain and discomfort in individuals with mental retardation. Res Dev Disabil 2005;26(5):433-9.

Prater CD, Zylstra RG. Medical Care of Adults with Mental Retardation Fam Physician 2006;73(12):2175-83.

Prkachin KM, Berzins S, Mercer SR. Encoding and decoding of pain expressions: a judgement study. Pain 1994;58:253–9.

Prkachin KM. The consistency of facial expressions of pain: a comparison across modalities. Pain 1992; 51:297–306.

Pkachin KM, Mercer S. Pain expression in patients with shoulder pathology: validity, coding properties and relation to sickness impact. Pain 1989;39:257–65.

Polit DF, Beck CT. Nursing research. Generating and assessing evidence for nursing practice. 8th edition. New York, J.B Lippincott company, 2008.

Rankin, G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analysis. Clin Rehabil 1998;12:187-99.

Regnard C, Mathews D, Gibson L, Clarke C. Difficulties in identifying distress and its causes in people with severe communication problems. Int J Palliat Nurs 2003; 9: 173–6.

Reid BC, Chenette R, Macek MD. Prevalence and predictors of untreated caries and oral pain among Special Olympic athletes. Spec Care Dentist 2003;23:139–42.

Reinfjell T, Lofstad GE, Veenstra M, Vikan A, Diseth TH. Health-related quality of life and intellectual functioning in children in remission from acute lymphoblastic leukaemia. Acta Paediatrica 2007;96:1280-5.

Richardson M. Reflection and celebration — Neal (1960-1987): narrative of a young man with profound and multiple disabilities. Journal of Intellectual Disabilities 1997; 1(4):191-5.

Roebroeck M, Harlaar J, Lankhorst GJ. The application of generalizability theory to reliability assessment: an illustration using isometric force measurements. Physical Therapy 1993;6:386-401.

Roy A, Simon GB. Intestinal Obstruction as a Cause of Death in the Mentally Handicapped. J Ment Defic Res1987;31:193-7.

Samet JM, Cushing AH, Lambert WE, Hunt WC, McLaren LC, Young SA, Skipper BJ. Comparability of parent reports of respiratory illnesses with clinical diagnoses in infants. Am Rev Respir Dis 1993;148: 441–6.

Schneider EM, LoBiondo Wood G. Perceptions of procedural pain: parents, nurses, and children. Special Issue: measurement issues with child-subjects. Children's Health Care 1992;21:157–62.

Schechter NL. The under treatment of pain in children: An overview. Pediatr Clin North Am 1989;36:781-94.

Schechter NL, Berde CB, Yaster M. Pain in infants, children, and adolescents: An overview. In: Schechter NL, Berde CB, Yaster M. (Eds.). Pain in infants, children, and adolescents. Baltimore, MD: Williams & Wilkins, 1993: 3-9.

Scherder EJ, Bouman A. Visual analogue scales for pain assessment in Alzheimer's disease. Gerontol 2000;46:47–53.

Schwartz L, Engel JM, Jensen MP. Pain in persons with cerebral palsy. Arch Phys Med Rehabil 1999; 80:1243-6.

Schwartz C, Rabinovitz S. Life satisfaction of people with intellectual disability living in community residences: perceptions of the residents, their parents and staff members. J Intellect Disabil Res 2003;47:75–84.

Shrout PE, Fleiss FL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420–8.

Smith GV, Teele RL. Delayed diagnosis of duodenal obstruction in Down syndrome. A.J.R 1980;134:937-40.

Sokol D, Sokol S, Sokol C. A review of non-intrusive therapies used to deal with anxiety and pain in the dental office. J Am Dent Assoc 1985;110:217–22.

Stallard P, Williams L, Lenton S, Velleman R. Pain in cognitively impaired, non-communicating children. Arch Dis Child 2001;85:460–2.

Stallard P, Williams L, Velleman R, Lenton S, McGrath P.J. Brief report: behaviors identified by caregivers to detect pain in noncommunicating children. J Pediatr Psychol 2002A;27:209–14.

Stallard P, Williams L, Velleman R, Lenton S, McGrath PJ, Taylor G. The development and evaluation of the pain indicator for communicatively impaired children (PICIC). Pain 2002B;98:145–9.

Stancliffe RJ. Proxy respondents and the reliability of the Quality of Life Questionnaire Empowerment factor. J Intellect Disabil Res 1999;43:185–93.

Stevens B. Composite measures of pain. In: Finley GA, McGrath PJ. (Eds). Measurement of Pain in Infants and Children. Progress in Pain Research and Management. Vol. 10. Seattle, WA: IASP Press, 1998: 161–78.

Stevens BJ, Johnson C. Physiological responses of premature infants to a painful stimulus. Nurs Res 1994; 43:226-31.

Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to their Development and Use. New York, NY. Oxford Medical Publications, 2001:65.

Suraseranivongse S, Santawat U, Kraiprasit K Petcharatana S, Prakkamodom S, Muntraporn N. Cross-validation of a composite pain scale for preschool children within 24 hours of surgery. Br J Anaesth 2001;87:400–5.

Terwee CB, Bot SDM, de Bor MR, van der Windt DAWM, Knol DL, Dekker J, Bouter LM, de Vet HCW. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34-42.

Thacker M, Hui JH, Wong HK, Chatterjee A, Lee EH. Spinal fusion and instrumentation for paediatric neuromuscular scoliosis: retrospective review. J Orthop Surg (Hong Kong) 2002;10:144-51.

Turk DC, Feldman CS. Non invasive approaches to pain control in terminal illness: The contribution of psychological variables. In Turk DC, Feldman CS. (Eds). Non invasive approaches to pain management in the terminally ill, New York: Haworth Press, 1992.

Turk DC, Melzack R. (Eds) Handbook of pain assessment,2nd ed. New York, NY: Guildford Press, 2001.

Turnbull J. Intuition in nursing relationships: the result of 'skills' or 'qualities'? Br J Nurs.1999;8(5):302-6.

Tyler CV Jr, Snyder CW, Zyzanski S. Screening for osteoporosis in community-dwelling adults with mental retardation. Ment Retard 2000;38:316-21.

Vandvik IH, Hoyeraal HM, Larsen S. Agreement between parents and physicians regarding clinical evaluation of patients with juvenile rheumatoid arthritis. Scand J Rheumatol 1988;17:459–63.

Van Schrojenstein Lantman-De Valk HMJ, Metsemakers JFM, Haveman MJ, Crebolder HFJM. Health problems in people with intellectual disability in general practice: a comparative study. Fam Pract 2000;17:405–7.

Voepel-Lewis T, Merkel S, Tait AR, Trzcinka A, Malviya S. The reliability and validity of the face, legs, activity, cry, consolability observational tool as a measure of pain in children with cognitive impairment, Anesth Analg 2002;95:1224-9.

Von Baeyer CL, Marche TA, Rocha EM, Salmon K. Children's memory for pain: overview and implications for practice. J Pain 2004;5(5):241-9.

Walts CF, Strikland OL, Lenz E.R. Measurement in Nursing Research. Philadelphia, Pa: F.A. Davis, 1984:141

Weiner D, Peterson B, Keefe F. Chronic pain-associated behaviors in the nursing home: resident versus caregiver perceptions. Pain 1999;80:577–88.

White-Koning M, Arnaud C, Bourdet-Loubère S, Colver A, Grandjean H. Subjective quality of life in children with intellectual impairment – how can it be assessed? Dev Med Child Neurol 2005;47:281–8.

Whitfield M, Langan J, Russell O. Assessing general practitioners' care of adult patients with learning disability: Case-control study. Qual Health Care 1996;5:31–5.

World Health Organization (2002) International Classification of Functioning, Disability and Health [ICF Home Page]. Accessed 2008 July 14. . http://www3.who.int/icf/onlinebrowser/icf.cfm?parentlevel=3&childlevel=4&itemsle

vel=4&ourdimension=b &ourchapter=1&ourblock=1& our2nd=17& our3rd =0& our 4^{th}=0.

Yamaguchi M, Takeda K, Onishi M, Deguchi M, Higashi T. Non-verbal communication method based on a biochemical marker for people with severe motor and intellectual disabilities. J Int Med Res 2006 Jan-Feb;34(1):30-41.

Zeltzer LK, Altman A, Cohen D, LeBaron S. Report of the subcommittee on the management of pain associated with procedures in children with cancer. Pediatrics 1990;86(5):826-31.

Zwakhalen SMG, Hamers JPH., Berger MPF. The psychometric quality and clinical usefulness of three pain assessment tools for elderly people with dementia. Pain 2006;126:210-20.