Assessing the affective state of dementia patients

do very severe Alzheimer patients benefit from visual stimulation? : an explorative study

de Pee, J.

Award date:
2012

Eindhoven, September 2012

Assessing the affective state of dementia patients

Do very severe Alzheimer patients benefit from visual stimulation? An explorative study.

identity number 0575427

in partial fulfilment of the requirements for the degree of

Master of Science
in Human-Technology Interaction

Department of Industrial Engineering and Innovation Sciences
Eindhoven University of Technology

Supervisors:
Yvonne de Kort, Eindhoven University of Technology
Inge Bongers, Tilburg University
Jettie Hoonhout, Jon Mason, & Tom Bergman, Philips Research

Key words: Alzheimer’s disease, Dementia, Explorative study, Observation method, Sensory experience, Wellbeing

Preface

Eight years. Eight years, it took me to get here, to become a Master of Science. And although I didn't imagine it would take that long for me to graduate, I wouldn't have done it differently if I could. In hindsight, even the "wasted year" during my Bachelor was well spent, since it taught me that I should try to please everyone less, and trust on my own intuition more. And that is exactly what I did in this project; I went along with the things that came my way and as a bonus I got the opportunity to work together with a great group of people. It was this great bunch of enthusiastic people, all with different kinds of expertise, skills, and backgrounds, that made this project so much fun for me. Therefore I want to take this opportunity to thank them personally. First, my supervisor Yvonne who whenever I entered her office - even when I had only 2 weeks left, run barely half of my analyses, and still no clue whether I would have any physiological data to report on - welcomed me with never diminishing optimism and gave me encouraging advice, due to which I always left with the feeling that I was doing well and that it was not so impossible to finish my project in time. Secondly, my external supervisor Inge Bongers from Tilburg University who made time for me in her busy schedule to attend my presentations and give her professional opinion on my work. Then of course Jettie, Jon, and Tom from Philips Research who believed in the potential of the Snoezel experience from the start and spent a lot of their free time - whether it was working on the Snoezel box on Friday afternoon or revising my concept thesis on Wednesday evening at ten pm - on my project. Additionally, I want to thank Ad Spaan on who's door I knocked numerous times with camera setting and recording problems and who always came up with a solution, even if that meant that he had to lend me his own laptop for the remainder of my study. Of course I want to thank the relatives of my participants, who gave their consent to let their wife or mother participate in my study and who showed a warming interest in my research. And finally I want to thank the staff of Vitalis Vonderhof who welcomed me like I was one of them and made me feel comfortable enough to be able to conduct my study in an environment that was totally unknown to me prior to this project. In particular Antonet, Lieke and Marianne, who made time in their demanding schedules to bring and collect inhabitants to and from my sessions, fill out the questionnaires and who showed a true interest in my study. But above all Tanja, without whom I would never have been able to conduct this study. From day one she was super enthusiastic, thought along with me about the difficulties and possibilities of my study set-up, involved me in her Snoezel sessions, made me feel like I was part of the activity team, and on top of that she could always find or make time to call a participant's guardian or a fellow occupational therapist and make them enthusiastic about my study. I couldn't have wished for a better partner during this project. And I couldn't have wished for a better project to close my eight years of study.

Jente de Pee
Eindhoven, September 2012

PS The unconditional support from my family during my study as well as during the preceding years is something I've always taken for granted - and hence almost forgot to mention in this preface- but nevertheless something that I am very grateful for.

Summary

The present study was initiated by Philips as part of a larger project on Snoezelen to improve the care for and the wellbeing of people with severe dementia. The term Snoezelen is used for interventions that focus on pleasurable multi-sensory stimulations, and an atmosphere of trust, restoration, and relaxation without any (cognitive) pressure (Lancioni, Cuvo, & O'Reilly, 2002). The initial goal was to validate the prototype developed in the context of this project, the Snoezel box, and create recommendations and guidelines for future iterations in its development. However, a literature study revealed that existing measures did not meet the requirements for validating this kind of products for people with very severe dementia. Hence, the goal of the present study became to develop a tool to measure the current affect, or appreciation, of people with very severe dementia with which interventions to promote their wellbeing could be validated and compared.

In order to achieve this, the following types of measurements were explored and compared to each other: 1. Structured behavioral observations (by means of a newly developed structured behavioral observation scheme suiting this specific target group), and 2. Proxy measures (Interact questionnaire filled out by caregivers and researcher). These measures were applied to a visual stimulation study (N=4) based on the Snoezelen philosophy with a 3 (Snoezel box, TV, Control) x 3 (repeated sessions) design. The study was carried out at care center Vitalis Vonderhof in Eindhoven, and sessions took about 20 minutes (6 minutes of baseline, followed by 14 minutes of intervention).

To validate the two types of applied measures, their sensitivity and convergent validity were assessed.

For both proxy measures, Interact During (14 items) and Interact Short (12 items), only one item ('Attentive/responding to/focused on activity/objects' and 'Talked spontaneously' respectively) showed an effect of Stimulus condition.

Of the three domain scales, Attention, Arousal, and Valence, from the structural observation measure, two (Attention and Arousal) were sensitive enough to capture an effect, or at least a trend towards an effect, of Stimulus condition in the present study, which is a rather promising finding for the future development of this kind of observational measures for this target group. In addition, this indicates that even very severe dementia patients are susceptible to these kinds of stimuli, which is an exciting finding by itself, but also a big motivation to carry on with the research to Snoezelen or other sensory stimulating interventions, and the development of supporting applications like the Snoezel box.

Table of Contents

1. Introduction

The present study was initiated by Philips as part of a larger project on Snoezelen to improve the care for and the wellbeing of people with severe dementia. The initial goal was to validate the prototype developed in the context of this project, the Snoezel box, and create recommendations and guidelines for future iterations in its development. However, a literature study revealed that existing measures did not meet the requirements for validating this kind of products for people with very severe dementia. Hence, the goal of the present study became to develop a tool to measure the current affect, or appreciation, of people with very severe dementia with which interventions to promote their wellbeing could be validated and compared.

In this section, first, in 1.1, a general introduction to the field of study is given, followed by 1.2 in which interventions for promoting the wellbeing of people with dementia are discussed. In 1.3 relevant measurements for the current study and its target group are described after which, in 1.4, the reasons why the available measures do not meet the requirements for assessing the wellbeing of people with severe dementia are summarized. Finally, in 1.5 Rationale, the objectives for the present study are presented.

1. 1 Field of study

The number of people suffering from dementia is growing rapidly. Alzheimer's Disease International (2009) estimated that there would be 36 million people with dementia worldwide in 2010, rising to more than 115 million by 2050. Dementia mainly affects the elderly. In the population younger than 65, about 0.1 percent suffers from dementia. This increases to 5 percent for people above the age of 65, and to 20 percent for those above the age of 80 (Alzheimer's Disease International, 2009).

Fifty to 60 percent of the dementia cases is caused by the most common cause of dementia, Alzheimer's disease (AD), a progressive brain disorder which destroys nerve cells in the brain, particularly in the regions that are responsible for storing and retrieving memories and new information (Alzheimer's Disease International, 2012). Consequently, people with AD go through different stages of cognitive decline during which their abilities to remember, speak, think, and make decisions deteriorate. These stages may overlap and range from no to little impairment (normal functioning) towards the final (7th) stage of very severe cognitive decline at which apathetic and aggressive behaviors are common and people "lose the ability to respond to their environment, to

carry on a conversation and, eventually, to control movement" (Alzheimer's Association, Seven Stages of Alzheimer's).

Several interventions[1)] have been applied to promote the wellbeing of people with dementia, e.g. Snoezelen or multi-sensory stimulation (Baillon et al., 2004; Baker et al., 1997; Baker et al., 2001; Staal et al., 2007; Van Weert, Van Dulmen, Spreeuwenberg, Ribbe, & Bensing, 2005), reminiscence therapy (Baillon et al., 2004), and activity therapy (Baker et al., 1997; Baker et al., 2001; Staal et al., 2007; Van Weert et al., 2005). (More information about these interventions can be found in section 1.2.) It is hard to verify whether these interventions are successful, because people who suffer from the most severe form of AD cannot communicate anymore how they feel about an intervention.

A tool to assess the pleasure and wellbeing experienced by this user group would be useful, because most people with AD spend a considerable number of years in this final AD stage (Alzheimer's Association, 2011) and each person should be able to live in conditions that promote his or her quality of life.

Additionally, in conventional product development processes, next to conducting user and usability tests, the appreciation of users is normally measured by asking them what they thought about the product, if they liked interacting with it, etc. Since we can not ask severe dementia patients what they think about a product or how they felt when using it, a tool to measure their appreciation, or current affect, would be helpful. Such a tool would make it possible to measure the effect of interventions targeted to people with dementia, like Snoezelen, to simplify the process of comparing these, and select the intervention that produces the most positive affect or appreciation for this particular target group.

Hence, the goal of this study was to create a tool to measure the affect of people with very severe Alzheimer's disease dementia by triangulation of three measures: 1. Structured behavioral observation during the intervention, 2. proxy questionnaires about the participant's behavior during and in the ten minutes just before and immediately after the intervention, and 3. Physiological measures heart rate and respiration rate during the intervention.

1) The term intervention is used throughout this report for applications or approaches that are applied for pleasure and should thus not be mistaken for treatments or applications intended to cure people.

1.2 Interventions for people with dementia

Care centers have adopted several approaches and interventions to promote the wellbeing of people with dementia, like activity therapy, Snoezelen or multi-sensory stimulation, and reminiscence therapy. A popular means to promote the wellbeing of people with dementia, particularly in the latest stages of the disease, is Snoezelen (Baillon et al., 2004; Baker et al., 1997; Baker et al., 2001; Staal et al., 2007; Van Weert, Van Dulmen, Spreeuwenberg, Ribbe, & Bensing, 2005). Snoezelen originated from the idea that adults with severe developmental disabilities will benefit from environments that provide sensory stimulation and relaxation (Hulsegge & Verheul, 1987). Later on, the term Snoezelen was used for interventions that focused on pleasurable multi-sensory stimulations, and an atmosphere of trust, restoration, and relaxation without any (cognitive) pressure (Lancioni, Cuvo, & O'Reilly, 2002). Consequently, Snoezelen environments include various kinds of applications, like projectors, music equipment, vibrating beds, soft toys, bubble machines, and aroma dispensers, which together provide stimulation possibilities for all sensory channels (Baker at al., 2003; Klages et al., 2011; Lanciono et al., 2002; Martin et al., 1998; Van Weert et al., 2003). Snoezel sessions are normally conducted by an occupational therapist or another caregiver and can be individual, meaning that a caregiver is interacting one-on-one with a patient, or in a group setting, meaning (that) one or more caregivers interact with several patients in a Snoezel environment. To give a concrete example of care center practices: At Vitalis Vonderhof, a care facility in Eindhoven, the group Snoezel sessions took generally about two hours, while at Vitalis Engelsbergen, another care center in Eindhoven, individual Snoezel sessions could have any duration from five minutes up to a maximum of 30 minutes. In addition to separate Snoezel sessions, Snoezel principles and activities can be integrated in the daily care program of carecentres (see also the study by Van Weert, Van Dulmen, Spreeuwenberg, Ribbe, and Bensing (2005)). An example of this, practised at both care center Vitalis Engelsbergen and care center Vitalis Vonderhof, is “bad-Snoezelen” (bath Snoezelen), in which soft music, a nice scent, and light projections are used while bathing a patient.

Other popular interventions to promote the wellbeing of this target group are Reminiscence therapy and Activity therapy. Reminiscence therapy was introduced in the 1980s and is based on the assumption that people’s early memories (i.e. from childhood) stay intact until the later stages of dementia. By means of conversation, often with supporting materials like pictures and music, these memories can be recalled and used to communicate with the patient (Cotelli, Manenti, & Zanetti, 2012). Activity therapy sessions are normally guided by an occupational therapist or other staff member of a care centre, and can include any activity, individual or in a group setting, that is meaningful to the dementia patient (Baker et al., 2001; Collier, McPherson, Ellis-Hill, Staal, & Bucks,

2010). Depending on the person's preference and stage of Alzheimer's, examples of activities are grocery shopping, gardening, setting the table for dinner, baking a cake, and reading and discussing the newspaper.

Activity therapy and reminiscence therapy appeal to a person's knowledge, memory, and (social) skills, while the Snoezelen philosophy, by contrast, is about being free from any cognitive pressure and does not demand cognitive performance from a participant. For this reason Snoezelen is especially fit for people who cannot rely on their cognitive abilities anymore, like people with very severe dementia.

1.3 Relevant measures

In this section relevant measures for the current study and its target group are discussed. Starting off with measures that are used to establish a person's level of dementia, followed by measures that assess quality of life and wellbeing of people with dementia, and concluding with measures that have been used to assess the effects of interventions for promoting the wellbeing of people with dementia, the studies they have been used in, and the results they have produced.

1.3.1 Measuring the severity of dementia

Mini Mental State Examination

The Mini Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975) is an interview based test for the assessment of cognition, which takes about 10 to 15 minutes to administer. It assesses orientation, memory, concentration, language, and motor skills and consists of 11 questions with a maximum total score of 30 (normal cognition) and a minimum of 0 (severe cognitive decline). Questions include counting backwards from 100 with steps of 7, memorizing three simple words, and recalling the current year, season, month, date, and day. The MMSE is a general measure for cognitive decline, but has been commonly adopted by dementia related studies to give an indication of the level of dementia of its participants (Baillon et al., 2004; Baker et al., 1997; Baker et al., 2001; Baker et al., 2003; Klages, Zecevoc, Orange, & Hobson, 2011; Staal et al., 2007). It has been shown that MMSE scores can distinguish between different levels of severity of dementia. An MMSE score of 30 represents no, 26 to 29 questionable, 21 to 25 mild, 11 to 20 moderate, and 0 to 10 severe dementia (Perneczky et al., 2006).

"Belevingsgerichte zorg" scale

The care center Vitalis Vonderhof, in which the experiments for the present study were carried out, does not administer the MMSE, but indicates the level of dementia of its inhabitants based on a "belevingsgerichte zorg" (emotion-based care) scale, consisting of the "bedreigde-ik" (threatened-self), "verdwaalde-ik" (lost-self), "verborgen-ik" (hidden-self), and "verzonken-ik" (sunken-self) stage. Hamer (2003) linked these stages to MMSE scores of 8-23, 4-7, 0-3, and 0-3 respectively.

The fact that the MMSE ranges for the last two stages of the emotion-based care scale are equal, indicates that there can be quite some variety in cognitive abilities within a group of people with a certain MMSE score. Based on observations at the dementia ward of the Vitalis Vonderhof care center, most people in the hidden-self stage would not be able to answer any question of the MMSE, and hence would get a MMSE score of 0. However, some of them would be able to understand the majority of the questions (like What date is it today?), but just don't know the answer to it, while others won't even understand what the question means and would therefore be unable to give any answer. Likewise, some people with an MMSE score of 0 will take the initiative to interact with you, e.g. when you walk into the living room, while others will only respond when you really make an effort to interact with them, and even then this response might be no more than just one word, sound, or facial expression.

1.3.2 Measuring Wellbeing and Quality of Life in dementia

Since there is no cure for dementia, promoting wellbeing and maintaining an optimal quality of life has become the main focus in dementia care (Ettema et al., 2005). Since 1990 an impressive number of measures has been created or used to measure Quality of Life (QoL) of people with dementia, which are summarized in table 1. Detailed descriptions of these measures and the studies they have been used in can be found in Ettema et al. (2005) and Schölzel-Dorenbos (2007).

Since someone's own perception of his or her quality of life is key in the appraisal of QoL, most QoL measures rely on self-report (Cooper et al., 2012; Barofsky, 2012). Some QoL measures even make use of only a single global item, like the widely used self-assessed health status item: "How would you rate your health, excellent, very good, good, fair, or poor?" (Barofsky, 2012), and the Visual Analogue Scale (VAS); a single-item measure in which the patient indicates his or her quality of life on a line or scale, in which the anchors are usually 'best possible quality of life' and 'worst possible quality of life' (De Boer et al., 2004).

A single item measure might not provide the insight in QoL that a study is looking for and scoring it asks a lot from the participant: he or she needs to take all aspects of the phenomenon in

consideration and evaluate them according to his or her personal state. This can already be a tough task for the average person, as QoL is a subjective concept and lacks a well-accepted operational definition (Kelley-Gillespie, 2009; Katschnig, 1997 in Ettema et al., 2005), thus scoring such a single item is simply not an option for people whose cognitive capabilities decline, like dementia patients. Hence, the majority of dementia-specific QoL measures are questionnaires (DEMQOL, DHP, DQoL, HSQ, NHP, SF-12, QOL-AD, QOLAS, SEOQoL, WHOQOL 100). Because of the declining cognitive capabilities of the target group, most QoL measures for people with dementia are interview based, or a caregiver is available for assistance when filling out the questionnaire (CBS, DEMQOL, DHP, DQoL, HSQ, NHP, QOLAS, SEIQol, SF-12). Other dementia specific QoL measures rely on interviews with relatives or guardians of the patient (CBS, QOL-AD, QOL-D)s, or make use of observations, either by guardians (ADRQL), by caregivers (QUALIDEM), or by objective observers (DCM, DS-DAT, PRS), to assess QoL of persons with dementia. More information about these measures and the areas of assessment they cover can be found in Table 1.

Although two types of measurements, proxy and self-report, are used to assess the quality of life of people with dementia, these are not equally applicable. While proxy measures like observations and interviews with caregivers have been applied throughout all stages of dementia, self-report measures like questionnaires and interviews have been successfully conducted only with people with mild to moderate dementia (MMSE 10-25) (see Ettema et al., 2005 table 1, p. 678, and table 3, p.681). This implies that even interview based self-reports are not an option for people with severe dementia (MMSE 0-10), let alone people with very severe dementia (MMSE = 0).

1.3.3 Measures to validate interventions for people with dementia

Several studies attempted to validate interventions for people with dementia (Baillon et al., 2004 Baker et al., 1997; Baker et al., 2003, Collier, McPherson, Ellis-Hill, Staal, & Bucks, 2010; Klages, Zecevoc, Orange, & Hobson, 2011; Martin, Gaffan, & Williams, 1998; Staal et al., 2007; Van Weert et al., 2005) or other mental disabilities (Kaplan, Clopton, Kaplan, Messbauer, & McPherson, 2006; McKee, Harris, Rice, & Silk, 2006). See Lancioni, Cuvo, and O'Reilly (2002) for a comprehensive review of Snoezelen studies (either targeted at people suffering from developmental disabilities or dementia) up to 2002, and Cotelli, Manenti, and Zanetti (2012) for a review of Reminiscence therapy for dementia studies.

When zooming in on the studies concerning Snoezelen for people with dementia, it became apparent that the (type of) applied measurements differed greatly in these studies. In the nine studies that were reviewed more than 20 different measurements were used. Additionally, the studies did not

Table 1

Quality of Life measures for people with dementia

Quality of Life Measure	Author(s)	Stage of dementia	Areas of assessment	Conducting procedure
Alzheimer's Disease Related Quality of Life scale(ADRQL)	Rabins, Kasper, Kleinman, Black, & Patrick (1999)	All stages	Efficacy of behavioral interventions, environmental settings, and drug treatments. 5 subscales: Social interaction. Awareness of self, Enjoyment of activities, Feelings and mood, Response to surroundings. (47 items) Can give separate scores, but can also be combined into 1 score	Trained interviewer/caregiver for data collection
Cornell-Brown Scale (CBS) for Quality of Life in Dementia	Ready, Ott, Grace, & Fernandez (2002)	Mild to moderate	Negative affect, Positive affect, Physical complaints, Satisfaction (weight satisfaction; restful sleep)	Clinicist interviewing both patient and caregiver
Dementia Care Mapping (DCM)	Kitwood & Bredin (1992)	All stages	Well or illbeing combined with the level of activity. The observers categorize activities that patients engage in (e.g., having a meal, sleeping, playing a game), and rate the level of well-ill-being every 5 min during a 6 hours	Trained observers
Dementia Quality of Life Instrument (DQoL)	Brod, Stewart, Sands, & Walton (1999)	Mild to moderate MMSE>12	Self-esteem, Positive Affect, Feelings of belonging, Sense of aesthetics	Patient self-report through interview
Health-related quality of life in dementia, DEMQOL	Smith et al., (2005)	All stages, for severe dementia (MMS 10) only proxy measure	5 domains: Daily activities, looking after yourself health, wellbeing, cognitive functioning, social Relationships, self-concept.	Interview based, also proxy version with caregiver and carer
Discomfort Scale – Dementia of Alzheimer Type (DS-DAT)	Hurley, Volicer, Hanrahan, Houde, & Volicer (1992)	Severe dementia MMSE <3	9 behavioral indicators, seven negative, 2 positive: Noisy breathing, Negative vocalization, Content facial expression, Sad facial expression, Frown, Relaxed body language, Tense body language, Fidgeting	Trained observers, systematic observation.
Duke Health Profile (DHP)	Novella et al. (2001)	Mild to severe	17-items, 5 independent health concepts: physical health, 5 items; mental health, 5 items; social health, 5 items; perceived health,1 item; disability, 1 item.	Self-report, questionnaires, with help if needed (more than 80%)
Health Status Questionnaire (HSQ-12)	Pettit, Livingston, Manela, Kitchen, Katona, & Bowling (2001)	Mild to moderate (lived at home)	Health perception, Physical functioning, Mental health, Role-physical, Role-Mental, Social functioning, Bodily pain, Energy	Self-report in interview
Health Utility Index: Mark 2 (HUI-2)	Neumann et al. (1999)	All stages	Sensation, Mobility, Emotion, Cognition, Self-care, Pain, Fertility, scored on levels of severity. Attribute and global index	Caregiver reports

Health Utility Index: Mark 3 (HUI-3)	Neumann et al.(2000)	All stages	Vision, Hearing, Speech, Ambulation, Dexterity, Emotion, Cognition, Pain, scored on levels of severity, Attribute and global index	Caregiver reports
Nottingham Health Profile (NHP)	Bureau-Chalot et al. (2002)	Mild to severe	Physical mobility, Social insolation, Emotional reactions, Pain, Sleep, Energy (38 items)	Self-report if possible, otherwise caregiver
Pleasant Event Schedule-AD	Teri & Logsdon (1991)	Mild to moderate	domains: Possible activities are rates don frequency, availability and enjoyability (53 items in original, 20 in short version)	Caregivers
Positive response schedule	**Perrin (1997)**	Sever dementias	effect of short, individualized interventions on the well-being of people with advanced dementia. 10 behavioral categories, focusing on behavioral components (smile, gesture)	Trained observers
Quality of Life Assessment Schedule (QOLAS)	Selai, Trimble, Rossor, & Harvey (2001)	Mild to moderate	Domains: Physical, Psychological, Social/family, Work, Cognitive (10 items)	Self-report through interview
Qualidem	Ettema, Droes, de Lange, Mellenbergh, & Ribbe (2006)	Up to severe dementia	9 subscales: care relationship, positive and negative affect, restless tense behavior, positive self-image, social relations, social isolation, feeling at home, having something to do (40 items).	Observation-scale
Quality of Life for Dementia (QoLD)	Terada et al.(2002) (Japanese)	Mild to severe	Efficacy of behavioral interventions, environmental settings, and drug treatments. 6 domains: Positive affect, Negative affect and action, Ability of communication, Restlessness, Attachment with others, Spontaneity and activity (31 items)	Care giver report
Quality of Life in Alzheimer's Disease (QOL-AD) 34)	Logsdon, Gibbons, McCurry, & Ten (1999)	Mild to moderate, MMSE<10	Appraisal of: Physical condition, Mood, Interpersonal relations, Ability to participate in meaningful activities, Financial situation.	Both patient and caregiver reports
Schedule for the Evaluation of Individual Quality of Life (SEIQoL)	O'Boyle (1994)	Mild to moderate,	30 items, 5 domains, different per subject	Semi-structured interview
Short Form Health Survey (SF-12)	Pettit et al. (2001)	Mild to severe	Mental, physical, social, perceived health, self-esteem, anxiety, depression, pain, general health	Self- report in interview
Vienna List	Porzsolt et al. (2004)	Severe dementia	Description of wellbeing in severe dementia. 5 factors communication/negative affect, bodily contact, aggression, mobility	Observations of professionals
World Health Organization Quality of Life, WHOQOL 100 and WHOQOL 24	The WHOQOL Group, (1998).	Moderate MMSE>15	Activity in: Psychological function, Physical state, Autonomy, Social Relations, Religion, Environment (100 or24 items)	Self-report

have wellbeing or affect as main subject, but focused on agitation (Baillon et al., 2004; Robbins et al., 2011; Staal et al., 2007; Van Diepen et al., 2002; Van Weert et al., 2005), functional performance (Collier et al., 2010; Klages et al., 2011), or behavior, mood, and cognition (Baker et al. 1997; Baker et al., 2003; Van Diepen et al., 2002; Van Weert et al., 2005). An overview of the measures used in these nine Snoezelen for dementia studies can be found in Appendix A. This variety of applied measures makes it difficult to compare the studies, their reported results, and hence the effectiveness of the assessed interventions. Moreover, these measures did not assess the effects on the behavior of the participants directly, but made use of the observation and interpretation from proxies or observers to measure the state of the dementia patient.

In addition to the accumulation of applied measures, these Snoezelen studies diverged on other study characteristics. The number of participants in the studies ranged from 10 to 125 (the study of Baker et al. (2003) had 136 participants, but combined three studies with N=94, N=26, and N=16 respectively), the number of sessions per participant ranged from 3/6 to 12/10-15, with 2 or 3 sessions a week. One study (Van Weert et al., 2005) integrated Snoezelen in daily care for 18 months, and hence did not have fixed sessions. Moreover, although all studies included people with dementia only, the ranges of the participant's MMSE scores within the studies was quite large: most studies included people with moderate to severe dementia (MMSE 0-20), and some also people with mild to severe dementia (MMSE 0-25).

With regards to the levels of dementia (and accompanying MMSE scores), it can be questioned whether the target group "dementia patients" can be treated as a homogeneous group, as the capabilities of people with severe dementia (MMSE<10) differ a great deal from those with mild dementia (MMSE 20-25). Since the MMSE scores in these studies cover the whole range of dementia from mild to severe, even within studies, and the number of participants in these studies is generally small, the potential effect of the tested interventions on the specific target group of the present study, very severe dementia patients (MMSE=0), can not be truly predicted.

While a number of studies measured effects during the intervention (Baillon et al., 2004; Baker et al 1997; Baker et al., 2003; Klages, et al., 2011; Van Diepen et al., 2002), most studies focused on the effects of the intervention on the behavior of the participant directly after the intervention, in comparison to the behavior just prior to the intervention (Baillon et al., 2004; Baker et al 1997; Baker et al., 2003; Collier et al., 2010; Klages, et al., 2011; Staal et al., 2007; Van Diepen et al., 2002), or they focused on long-term effects (Baker et al 1997; Baker et al., 2003; Klages, et al., 2011; Van Diepen et al., 2002; Van Weert et al., 2005).

Because of their cognitive decline, the notion of past and future has faded for most people with dementia, which might be the reason why many measures in these studies did not indicate any effect and most studies reported merely tendencies or indications of effects. Only a few studies found modestly significant effects of the interventions. Below the results of the reviewed studies are summarized per topic.

Functional performance: Klages et al. (2011) did not find any significant effect on balance when comparing multisensory stimulation to the volunteer visits that served as control condition. Collier et al. (2010) found that both the Snoezelen group and the activity group showed an improvement in motor and process skills from first to last treatment session, as well as from the first to the sixth treatment session, and the multisensory stimulation group improved their motor skills significantly over sessions. Staal et al. (2007) did not find any statistical effects on the RADL, or BDP (the full names and specifics of the measures mentioned in this section can be found in Table 1), but did find improved levels of independence in activities of daily living on KI-ADL for the multisensory stimulation group compared to the activity group.

Agitation: In the (pilot) study of Van Diepen et al. (2002) the CMAI showed a tendency to be lower at 4 weeks and follow-up after a Snoezelen intervention, and the ABMI total score for the Snoezelen group demonstrated a light tendency to be lower after than before the session, which was not sustained during the 15 and 30 min. after the intervention, while for the reminiscence group it showed a tendency to increase over the four time-points. In the study of Baillon et al. (2004) no significant results for agitation (CMAI) were found, while the multisensory stimulation group in Staal et al. (2007) showed a stronger decrease in agitation and a better improvement in apathy compared to the activity group, and Van Weert et al. (2005) found significant treatment effects for level of apathetic behavior, loss of decorum, rebellious behavior, aggressive behavior, and depression after Snoezel care. In the study of Van Diepen et al (2002) a change in heart rate appeared to be dependent on the behavior and activities, while in the study of Baillon et al. (2004) no results of heart rate were found. Robbins and Norton (2011) reported a marked decrease in depression indicators after treatment, slightly higher attendance after a course of Snoezelen interventions, and a slight decrease in the total number of PRN medications administered for problem behaviors from pre- to post-measure.

Behavior, mood, and cognition: In the study of Baillon et al. (2004) no significant results for behavior during the sessions (INTERACT) were found and also Baker et al. (1997) and Baker et al. (2003) did not find effects of condition (Snoezelen or activity) on the INTERACT items. Baker et al. (2003) did not find

effects for GIP, and no significant differences between Snoezelen and activity groups on MMSE, but both groups related better to others and were less bored/inactive after sessions than before, while REHAB showed deterioration in behavior by having a significantly higher score at follow-up than at post-trial. In the study of Baker et al. (1997) REHAB speech and deviant behavior scores deteriorated significantly for the activity (control) group from post-trial to follow up, while Snoezelen participants improved significantly according to their BRS social disturbance score. Van Weert et al. (2005) found significant changes in mood (happiness, enjoyment, sadness) and adaptive behavior (responding to speaking, relating to caregiver, normal-length sentences) from pre to post trial based on the Interact measurement applied to videos of morning care. Van Diepen et al. (2002) performed a pilot study to compare the ease of use and effectiveness of various measures. In this study the INTERACT measure indicated that sessions of the interventions had a positive effect on the participants' behavior and changes in heart rate appeared to depend on behavior and activities.

The summarized results above show that when (an indication of) an effect of intervention was found, this often applied to the experimental as well as to the control condition. As studies made use of control conditions with equal social conditions like activity therapy (Baker et al., 1997; Baker et al., 2003; Collier et al., 2010; Staal et al., 2007), reminiscence sessions (Baillon et al., 2004; Van Diepen et al., 2002), and volunteer visits (Klages et al., 2011), this suggests that the primary factor of influence was personal or social attention, or at least that the effect of this variable was bigger than the effect of the content of the intervention.

When looking at the specific content of the interventions, in particular of the Snoezelen interventions, it appears that many Snoezelen studies exposed all participants to the same Snoezelen environment (Baker at al., 2003; Klages et al., 2011; Martin et al., 1998), while the Snoezelen philosophy is based on the belief that different people can have different sensory preferences. Although participants in these studies were able to choose (with help from caregivers) the applications they wanted to interact with, the other applications were still present and could have caused overstimulation and even irritation. As the aim of Snoezelen is reducing negative states, exposing participants to one type of stimulus might be more appropriate (Pool, 1999).

1.4 Lack of an appropriate measurement tool

The review of measures and studies related to quality of life and wellbeing, Snoezelen, and validating interventions for people with dementia showed that the field lacked a sound measure to validate

interventions for promoting the wellbeing of people with severe dementia. Especially people with very severe dementia, the main target group of Snoezelen and hence of future Snoezelen applications, appeared to be a target group which had not been researched much. Studies investigating interventions for people with dementia included participants whose levels of dementia lay within a range from mild to severe (MMSE 0-25), due to which it was not possible to truly anticipate the effect of the study for people with very severe dementia (MMSE≤0). Also, a big number and variety of measures had been used to validate interventions for people with dementia, which made it hard to compare the studies and their effects, even more because many of these measures did not show significant effects and consequently mainly reported trends or indications. In addition, many studies that investigated the effect of interventions for people with dementia did not focus on quality of life, wellbeing or affect but had functional performance or cognition as main subject, which are domains on which very severe dementia patients are not expected to be able to show any improvement or progress anymore. The studies which additionally investigate subjects related to affect, like agitation, mood, and behavior, often looked for long-term effects, while people with very severe dementia live primarily in the moment and have lost their ability to learn. Moreover, existing and often used quality of life measures are not adequate for people with dementia, as these measures rely on self-report which lies not within the (mental and physical) capabilities of people with very severe dementia. Hence, there is a need for a measurement tool with which the affect of people with very severe dementia can be assessed, and which can be applied to validate interventions for this specific target group.

1.5 Rationale

The present study aspired to create a measure with which the current affect, i.e. a person's appreciation of his or her current situation, of people with severe dementia can be assessed, making it possible to validate and compare interventions for this target group. In order to achieve this, the following three types of measurements were planned to be explored and compared to each other: 1. Structured behavioral observations, 2. Physiological indicators (heart and respiration rate), and 3. Proxy measures (Interact questionnaire filled out by caregivers and the researcher). These measures were applied to a visual stimulation study, based on the Snoezelen philosophy.

Because Snoezelen is about providing a pleasant experience by means of the (sensory) stimulation of a person's sense of preference, it was decided to provide one type of stimulus only, hereby ruling out

the possibility that multiple stimuli would cancel out each other's positive effects. Visual stimuli were selected as the stimuli of choice because the Philips prototype has a projection of lighting patterns as main output (although the device also includes an optional sound feature, but this was not used in the current study) and the occupational therapist at Vitalis Vonderhof agreed that many people with severe dementia are (still) susceptible to visual stimuli.

As the initial goal of the current study was to investigate the effects of the Philips prototype, this intervention was compared to another visual stimulation condition; a muted DVD fragment of the movie March of the Penguins, and to a control condition in which no explicit visual stimulus was present (i.e. participants were sitting in a dimmed room). In prior Snoezelen studies a caregiver was present in the Snoezelen sessions as well as in the activity or reminiscence sessions that served as control condition. In order to investigate the effect of primary sensory stimulation only, without the confounding effect of social interaction or attention, the sessions of the current study did not involve the presence of a caregiver. The caregiver brought a participant to and collected her after the session, but during the session the participant was alone in the room, making it possible to observe the effect of the intervention only.

Next to the quantitative part of the study described above, qualitative data was gathered by the researcher in the time she spent at the care center. These observations and finding were used as an additional source to provide insight in the effect of visual stimuli on the affective state of the participants, as input for recommendations for testing interventions with this specific target group and for future developments of Snoezelen equipment.

2. Method

The sensitivity and convergent validity of the three types of measures (structured behavioral observations, physiological indicators, and proxy measures (Interact questionnaires filled in by caregivers and the researcher)) were assessed by applying them to a visual stimulation study with people with very severe dementia, based on the Snoezelen philosophy. In the following sections the study set-up is described by means of its Design, Participants, Setting & Apparatus, Manipulations, Procedure, and Analysis.

2.1 Design

The present study had a 3 (type of intervention: Snoezel box vs. Positive affect DVD vs. control) X 3 (repeated sessions) design. The order in which the participants received the three different types of interventions was assigned pseudo-randomly: Because of the vulnerable target group (see high drop-out rates in studies of Baillon et al. (2004); Collier et al. (2010), Van Diepen et al. (2002), and Van Weert et al. (2005)) and the accompanying uncertainty regarding the total number of sessions a participant would be able to participate in, patients participated in three "rounds" in which the three different conditions were randomly assigned to them. This ensured that the experiment data would include a fair distribution of conditions per participant, even when a participant would drop out early.

2.2 Participants

2.2.1 Participant characteristics

Four (female) participants, aged between 77 and 94 (M=84.3, SD=7.08), either in the "hidden-self" stage or "sunken-self" stage of Alzheimer's disease, participated in this study. All participants were living at the closed dementia ward of care center Vitalis Vonderhof in Eindhoven and were selected by their occupational therapist to join the study, on basis of their daily behavior and anticipated susceptibility to visual stimulation. The participants were visually oriented (reacting on visual stimuli during the day: looking at people when they walk by, enjoying watching TV, etc.) and all were wheelchair-bound.

When the study started it was planned to include a bigger number of participants than the eventual four but, although the population of people with dementia is growing, it appeared to be very difficult to find people that suited the above mentioned requirements. Dementia wards often include only a small number of people, who normally are in different stages of dementia. Due to the progressive nature of the disease, the group with very severe dementia is often small. Additionally, not all severe dementia patients are susceptible to visual stimuli or in health conditions that allow participating in this kind of study.

The guardians of the participants (in this case children or the partner of the participant) were informed in detail about the goal and set-up of the study and gave their consent for letting their relative take part in the study.

2.2.2 Participant ethics and safety

The study design was reviewed and approved by the ethical committee of Philips (ICBE- Internal Committee Biomedical Experiments). Overall, the situation created for collecting data did not differ that much from daily practice in the care center. Leaving patients for some time alone in a bed or living room is quite common in care facilities: it offers patients some quiet time without being disturbed by other patients.

The stimuli presented to the participants were selected after consultation with the occupational therapist of the care facility. Specifications of the Snoezel box condition (such as speed, patterns, etc.) had been discussed with the occupational therapist and settings were selected that were considered suitable for this target group. The movie fragment of March of the Penguins that was used in the video condition was selected on basis of its purely positive content (it starts when the baby penguins hatch from their egg and ends with the toddler penguins cuddling up together) and the fact that nature videos are also used during the existing Snoezelen sessions at Vitalis Vonderhof.

And although the participants were left alone in the room for the experiment, they were constantly monitored during the sessions (via the real-time footage of the camera watched by the researcher on the laptop in the experiment booth), and the researcher was only a sliding door away (the adjacent bathroom was used as monitoring room). If a participant would have had a negative reaction to the testing conditions, staff could be notified and the session could have been stopped right away. In none of the sessions did the participant respond negatively or panicky to the situation.

2.3 Setting & Apparatus

2.3.1 Arrangement of experiment room

The experiment was conducted at Vitalis Vonderhof in the private room of one of the inhabitants. The room accommodated a bed, a chair, a small table, two cupboards, and a TV (Panasonic, 28 inch). For the experiment, a DVD player (AKAI) was installed and the TV set and DVD player were moved to the left side of one of the corners of the room. The Snoezel box and the accompanying laptop (HP Elitebook 8440p) were placed on two small side tables at the right side of the corner. The camera stand with camera (uEye UI-2220-C, 50 fps, 768 x 582 pixels) was put in the corner of the room and the USB microphone (Logitech USB Desktop Microphone) was put in front of it on the ground. A duct tape cross was stuck to the ground at two meters from the camera stand to mark the place where the wheelchair of the participant should be positioned. The camera and microphone were connected

with a 6 meter USB cable to a laptop (HP Elitebook 8440p) that was positioned on a table in the adjacent bathroom which functioned as observation booth during the experiment. By turning off all lights and closing the curtains, the room was dim but pleasantly lit, as the sessions were scheduled during daytime. Additionally, the personal pictures were removed from the two walls that the participant would be facing.

2.3.2 Snoezel box

The Snoezel box is a prototype developed by Philips, about the size of a shoebox, which is designed to be a portable Snoezel application. At the moment of the study it included a sound option (built-in MP3 player with speakers) and two types of visual output, namely a color changing tube and a color projection wheel with LEDs which project moving colored light patterns on the wall and ceiling adjacent to the device. Its settings, e.g. the speed, brightness and color scheme, can be adjusted by the researcher to the preferences of occupational therapist (and hence the patient's preferences). For the current study only the projection feature was used and one basic scenario suiting the target group was created in agreement with the occupational therapist. This scenario was used for all participants in the study.

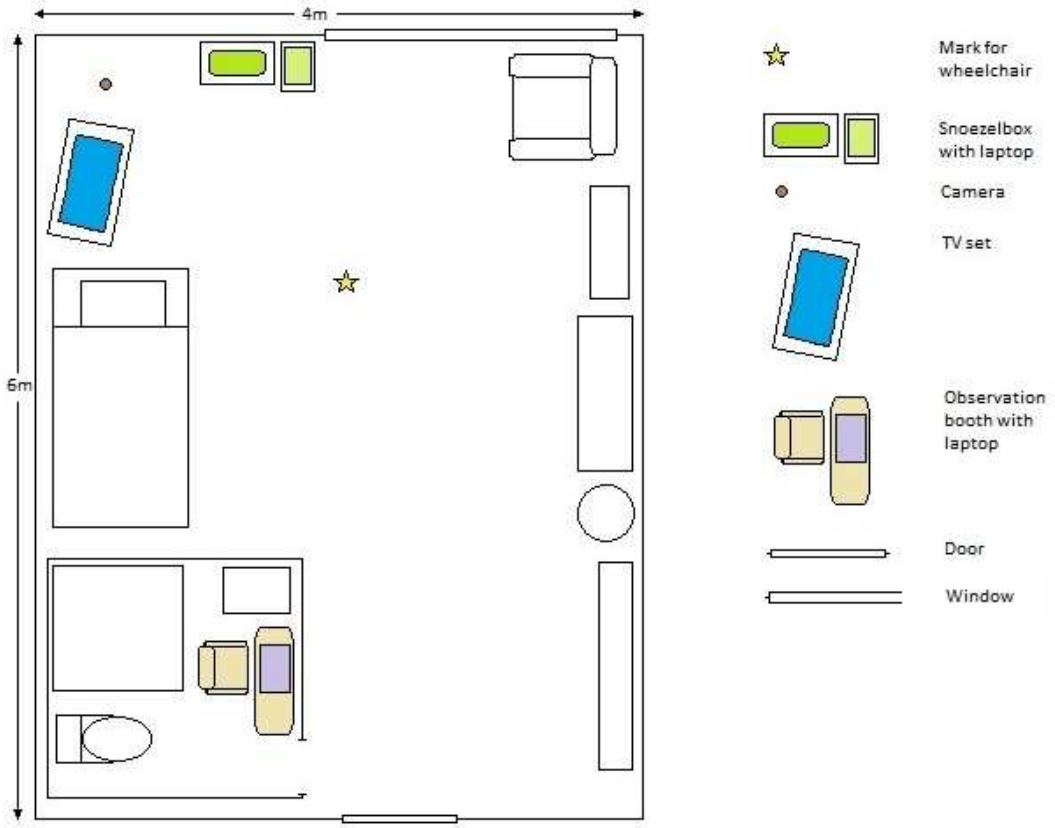

Figure 1. Schematic view of experiment room.

Figure 2, 3, 4, & 5. The Snoezel box - Range of settings.

Figure 6 & 7. Experiment room.

Figure 8 & 9. Arrangement of equipment (pilot test setting).

Figure 10 & 11. Experiment booth in bathroom adjacent to experiment room.

2.4 Manipulations

The experiment consisted of three conditions: the Snoezel box condition, the TV condition, and the Control condition. Every session started with a 6-minute baseline in which the participant was sitting in a dim room without visual stimulation from either the Snoezel box or the TV.

In the *Snoezel box condition*, after the six minutes of baseline measurement, the Snoezel box was turned on, projecting different forms of colored dynamic light in a slow pace on one of the walls and the adjacent part of the ceiling.

In the *TV condition*, after the six minutes of baseline measurement, a segment of 14 minutes from the movie March of the Penguins (2006), in which mainly young penguins are shown (00:52:00–01:06:00), was displayed on the television (nature videos like these are also used during the existing Snoezelen sessions at Vonderhof). Movies are a recommended means to elicit positive affect (Gross & Levenson, 1995; Philippot, 1993; Westermann, Spies, Stahl, & Hesse, 1996) and hence have been used in various studies (Codispoti, Surcinelli, & Baldaro, 2008; Fernandez et al., 2012; Gomez, Zimmermann, Guttormsen-Schär, & Danuser, 2005; Gross & Levenson, 1995).

In the *Control condition*, after the six minutes of baseline measurement, the Snoezel box and the television were both left turned off. The participant was sitting in a quiet room with dimmed lighting.

2.5 Measurements

As literature research showed that a specific tool to measure the affect of people with severe dementia in the moment did not exist yet, the current study made a first attempt to create such a measure, in the form of a structured behavioral observation scheme. This tool was planned to be validated by physiological indicators (but this measure had to be excluded from the study) and the existing proxy measure Interact. In this section the development and application of these three measures are described.

2.5.1 Observation measure

In order to be able to measure positive affect in the moment, a structured behavioral observation measure was developed in the form of an observation scheme, including behaviors that could be linked to demonstrating appreciation (or the opposite).

Existing observation measures were reviewed, but those appeared to be not sufficient for people with very severe dementia. Reasons for this were that they measured the behavior over a longer period, e.g. over the last two weeks: Dutch Behaviour Observation Scale for Psychogeriatric In-patients (BIP; Verstraten & Van Eekelen, 1987), Cohen-Mansfield Agitation Inventory (CMAI; Cohen-Mansfield, Marx, & Rosenthal, 1989), and Qualidem (Ettema, et al., 2006), or over the last week: Cornell-Brown Scale for Depression and Quality of Life in Dementia (CBS; Ready, Ott, Grace, & Fernandez, 2002) and Cornell Scale for Depression in Dementia (CSDD; Alexopoulos, Abrams, Young, & Shamoian, 1988). Some measures focused on negative affect only, e.g. on agitation: CMAI, depression: Cornell Scale for Depression in Dementia (CSDD; Alexopoulos et al., 1988), or discomfort: Discomfort Scale - Dementia of Alzheimer Type (DS-DAT; Hurley, Volicer, Hanrahan, Houde, & Volicer, 1992). And finally, some included units with a high level of interpretation: BIP, e.g. “seems to be happy with visits of relatives”, Interact (Baker & Dowling, 1995), e.g. “Enjoying self, active or alert”, and Qualidem, e.g. “Feels at home at the ward”, while e.g. the widely used Facial Action Coding System (Cohn, Ambadar, & Ekman, 2006) records purely objective units like “Head tilt right” and “Lid droop”.

The developed observation scheme was created by taking and combining suitable items from the measurements mentioned above and adjusting them to a level of interpretation considered appropriate for measuring affect in the moment and for this specific target group, e.g. “Hand movements, quick/slow” and “Talking, positive/neutral/negative”. The selection of units for the final set was based on pilot observations of the target group during Snoezel group sessions and in the

living room at Vitalis Vonderhof, and on two of the three factors of emotions proposed by Russel and Mehrabian (1977). Their three emotion factors are: pleasantness-unpleasantness, ranging from happiness or ecstasy to unhappiness or extreme pain, degree of arousal, ranging from frenzied excitement, via alertness and drowsiness, to sleep, and dominance-submissiveness, ranging from feeling influential and in control to feelings of total lack of control or influence. As the stimuli provided in the study did not ask for any influence or control from the participant but were expected to be able to capture the attention of the participant, the units in the developed behavioral observation scheme (Appendix B) cover behaviors in the domains Attention (alertness-apathy), Arousal (arousal-calmness), and Valance (positive affect-negative affect).

The behavioral observations were conducted by the researcher on basis of video tapes of the sessions capturing the participant's face and upper body. The behaviors were rated in time windows of 2 minutes, which had shown to be a reasonable time span during pilot observations. The session took 20 minutes; 6 minutes of baseline and 14 minutes of intervention. For every unit (e.g. "Talking") in the observation scheme, the researcher ticked the box of the item (e.g., "positive", "neutral", or "negative"), that represented the behavior of the participant best during the last two minutes. For every domain scale, Attention, Arousal, and Valence, the items of the units were assigned a value of 0 (low), 0.5 (medium), or 1 (high), resulting in a unit value between 0 and 1. For the Valence scale an alternative scoring of -1 (negative affect), 0 (neutral), and +1 (positive affect) was explored as well. For an overview of the items per scale and their ratings, see figures 3-5 in section 3. Results. Per participant per session the behavioral scores of the units were combined into an Attention, Arousal, and two Valence scores, resulting in 8 behavior scores per session: 4 for the baseline and 4 for the intervention measure.

2.5.2 Physiological measure

The physiological indicators heart rate (Baillon et al., 2004; Fernandez et al., 2012; Vos et al., 2011), respiratory responses (Bloch, Lemeignan , & Aguilera, 1991; Boiten, Nice, Frijda, Cornelis, & Wientjes, 1994; Frazier, Strauss, & Steinhauer, 2004; Gomez et al., 2005) and a combination of the two measures (Codispoti et al., 2008; Herring, Burleson, Roberts, & Devine, 2010; Rainville, Bechara, Naqvi, & Damasio, 2006) have been used to distinguish between different affective states. For the current study it was planned to deduce these two indicators from the videotaped session by Philips' Vital Signs software. Vital Signs was not used in the end due to the software not being ready in time and hence this measure will not be further discussed in the rest of this report.

2.5.3 Proxy measure

The Interact Short (Baker et al., 2001; Baker et al., 2003) was used as proxy measure in the present study, filled out by the care center staff. The Interact (Baker & Dowling, 1995) is a proxy questionnaire especially created for assessing Snoezel sessions (and comparable sessions like activity sessions). It was designed to record behavior during these sessions and covers the domains Mood, Speech, Relating to person, relating to environment, Need for prompting, and Stimulation level with 22 items. For each of these items the frequency of occurrence of a particular behavior (e.g. "Co-operated") is assessed on a five point Likertscale (1 = not at all, to 5 = nearly all the time). The Interact Short is a 12-item version of Interact, and can be completed by care center staff concerning a participant's behavior the 10 minutes immediately before a session and the 10 minutes immediately after a session, to establish any observable changes. Although this measure assesses the behavior of the participant before and after the intervention, and hence not the actual immediate behavior of the participant during the session, it was added to the set of measures of the current study because it was designed especially for assessing the effect of Snoezel applications, has been used successfully in recent Snoezelen studies (Baillon et al., 2004; Baker et al., 2001; Van Weert et al., 2005), and has the potential to validate the other measures by means of cross-validation. Additionally, the Interact During (the full 22 item version of Interact), completed by the researcher right after watching the videotaped session, was added to the set of measures. Although the Interact During measure is based on the participant's behavior during the session, it does not directly measure the behavior or appreciation of the participant, but depends, just like the Interact Short, on the interpretation of the participant's behavior according to another person.

2.6 Procedure

A caregiver -blind to condition- placed the participant (all participants were wheelchair bound) on the marked spot at a distance of about 2 meters from the camera, TV, and Snoezel box and left the room. The caregiver was asked to fill in an Interact Short questionnaire and to return in 20 minutes to collect the participant. After six minutes of baseline measurement, the researcher entered the room and turned on either the TV (TV condition) or the Snoezel box (Snoezel box condition), or fiddled a bit with the Snoezel box without actually doing anything (Control condition), and went back to the observation booth. After another 14 minutes, the researcher returned to the room, turned off the TV (TV condition) or the Snoezel box (Snoezel box condition), or fiddled a bit with the Snoezel box (Control condition) and went back to the observation booth. By 'fiddling with the Snoezel box' in the

Control condition, the differences in procedure between the three conditions were minimized and the potentially confounding variable social attention was kept equal in all three conditions. Around the time that the researcher returned to the observation booth, the caregiver knocked on the door and -still blind to condition- collected the participant. The video recording was stopped when the participant had left the area of the room captured by the camera. On the way out the caregiver received a second Interact Short questionnaire which she was asked to fill out about the behavior of the participant during the 10 minutes after the session. The researcher collected the questionnaires later. Hereafter the researcher -blind to Interact Short scores- filled out the Interact During questionnaire and prepared the experiment room for the next session. At the end of the last session of the day, the video sequence and sound files were transferred to the encrypted volume on the external hard disk.

2.7 Data Analysis

In order to fulfill the research goal of this study to create a measure with which the affect in the moment of people with severe dementia can be assessed, the data from the two explored measures (1. Structured behavioral observations, and 2. Proxy measures (Interact questionnaire filled out by caregivers)) was prepared and analyzed in the following way:

First, the two measures were analyzed and reconstructed to maximize their reliability. For the structured behavioral observation measure and the proxy measure this entailed analyzing the internal consistencies for the groups of items of each subscale of the measure. Subsequently, the sensitivity of the measures for Stimulus condition was assessed by applying Linear Mixed Model Analysis with participant number as subject, intervention measure score as dependent variable, Stimulus condition as factor, and, if applicable, the baseline measure score as covariate. Session number was not included as a repeated measure in the analysis because it was assumed that severe dementia patients live in the moment and do not have the ability to remember and connect different events anymore. For this reason, no effect of sequence was expected and every session was treated as a measure independent of order. In case of a found effect of Stimulus condition, pairwise comparisons were applied to find out in which way the three Stimulus conditions differed from each other. Third, the convergent validity of the measures was assessed by exploring correlations between their Z-scores.

3. Results

In the first part of this section the construction and reliability testing of the different scales are described for both the structured behavioral observation and the Interact measures. For the structured behavioral observation for each scale, the measure of the baseline period and the measure of the intervention period are assessed separately. In the second part of this section, the sensitivity of the scales is investigated, and in the third part their convergent validity is assessed.

3.1 Construction and consistency of the scales

3.1.1 Structured behavioral observation measure

The researcher watched the videotaped sessions and for every subscale on the behavior list (Appendix A) she ticked the box of the item that represented the participant's behavior during the last two minutes best. For every session, the baseline had a duration of 6 minutes and the intervention of 14 minutes. Consequently, the item scores were based on 3 measures for the baseline and on 7 measures for the intervention.

The items Yelling and Teary eyes were excluded from the analysis because no teary eyes or yelling behavior was observed during any of the sessions.

In order to create a coherent scale for Attention, Arousal, and Valence, first a reliability analysis was run with all selected items for a certain scale included. If the Cronbach's α was above .7, the scale was considered coherent, provided that none of the items had a negative corrected item-total correlation. In case these conditions were not met, two procedures were applied. Those items that would raise the overall Cronbach's α of the scale if they were excluded (according to the 'Cronbach's alpha if item deleted'-value) were excluded until an acceptable Cronbach's α of .6-.7 was produced. Additionally, items that correlated negatively with the total scale were excluded from the scale.

3.1.1a Attention scale

When all Attention items were included in the Attention scale, the Cronbach's α for the baseline measurement was .663 (5 items) and .795 (6 items because of the additional item 'Attention directed towards intervention') for the intervention measure. All corrected item-total correlations were > .3, except for Touching in the intervention measure (corrected item-total correlation = .28). An overview if the items included in the final Attention scale can be found in Table 3.

Table 3

Final Attention scale

Items	Value	α
Head	Passive (0) /active (1)	Baseline α = .663
Eyes	Open (1) / open-closed (.5) / closed (0)	(5 items, not including the item
Eyes	Following (1) / blank(0)	Attention for application)
Talking	Positive (.5) /neutral (.5) /negative (.5) / not (0)	
Touching	With (1) / without intent (0) /not touching (0)	Intervention α = .795 (6 items)
Attention for application	Yes (1) / no (0)	

3.1.1b Arousal scale

The reliability analysis for the 10 items that were selected as Arousal items resulted in a Cronbach's α of .536 for the baseline measure, and of .553 for the intervention measure. Because Laughing correlated negatively with both scales (corrected item-total correlation of -.05 for the intervention measure, and -.22 for the baseline measure) and Mouth activity had a low correlation (.02) with the total scale for the intervention measure and a low negative correlation (-. 03) for the baseline measure, first Laughing was excluded from the scales resulting in an α of .571 for both the intervention and the baseline measures (both 9 items). Subsequently, as excluding Laughing made the correlations of Mouth activity more negative, the Mouth activity item was excluded as well, resulting in an 8 item Attention scale with an α of .668 for the baseline measure and an α of .663 for the intervention measure. Respiration had a low correlation with the scales from the start, but due to the exclusion of the items Laughing and Mouth activity, it correlated negatively (-.22) with the total scale for the intervention measure and had only a very small (.10) correlation with the total scale for the baseline measure. After exclusion of the Respiration item, the final 7-item Arousal scale had an α of .695 for the intervention measure and an α .677 for the baseline measure. The items Talking and Sounds without a goal, and for the baseline measure also the item Head, correlated less than .3 with the total scale, but were kept based on face validity and low gain in α if deleted. An overview if the items included in the final Arousal scale can be found in Table 4.

Table 4

Final Arousal scale

Items	Value	α
Talking	Positive (1) / neutral (1) / negative (1) / not (0)	Baseline α = .695
Sounds without intent	Positive (.5) / neutral (.5)/ negative (.5) / not (0)	(7 items)
Head	Passive (0) / active (.5)	
Head movement	Quick (1) /slow (.5) / not moving (0)	Intervention α = .677
Hands movement	Quick (1) /slow (.5) / not moving (0)	(7 items)
Touching	With (1) / without intent (.5) / not touching (0)	
Body movement	Rocking (1) /moving (1) / not moving (0)	
Laughing	Excluded	
Mouth activity without a goal	Excluded	
Respiration	Excluded	

3.1.1c Valence scale

For the Valence scale we explored two alternatives: one only included (potential) indicators of positive affect, so the scores ran from 0 (neutral) to 1 (positive valence); the second included indicators for both negative and positive affect, so the scores ran from -1 (negative affect) to +1(positive affect).

Both the Positive Valence scale and the Positive-negative Valence scale did not manage to lead to α values near .7. The original 10-item Positive-negative scale had only an α of .002 for the baseline measurement and even a negative α, -.437, for the intervention measurement. Excluding the item Mouth corners, raised the α's to .367 and .074 respectively (9 items). In the baseline measure, the only item that still correlated negatively (-.23) with the total scale was the item Laughing. When excluded (8 items remaining), the α of the baseline measure for Positive-negative Valence became .406 and the α of the intervention measure Positive-negative Valence scale went up to .163. Excluding the item Eyebrows created a 7-item Positive-negative Valence scale with an α for the baseline of .531, and for the intervention measure of .340. Removing the item Touching created two 6-item scales with an α of .592 for the baseline measure and .527 for the intervention measure. Removing the item Talking led to a 5-item Positive-negative Valence scale with α = .649 for the baseline and α= .618 for the intervention measure. However, the latter included still one negatively correlating item, Sounds with no intent (corrected total-item correlation = -.315). When this item was excluded, a 4-item scale with α= .686 for the baseline measure and α= .699 for the intervention measure remained.

The positive Valence scale with all 5 items included got an α of .089 for the baseline measure and an α of .222 for the intervention measure (the other 5 Valence items: Mouth, Mouth activity, Hands, Touching, and Body were not included, because these only indicated negative valence behavior). In the intervention scale all but one item (Talking, corrected total-item correlation=.42) correlated < 2 with the total scale, but none of them had a negative correlation. For the baseline measure however, Sounds without an intent and Mouth corners correlated negatively with the total scale. By excluding the Mouth corners item, the scale would only contain four items (α =.248), but none of them would have a negative correlation with the total scale for the baseline measure. However, excluding the Mouth corners item from the intervention measure scale would lead to a drop in Cronbach's α from .222 (5 items) to .095 (4 items). The total item-correlation matrix for the intervention measure indicated that excluding the item Eyebrows instead would lead to a 4-item scale with an α of .394, but doing the same for the baseline scale would result in a 4-item scale in which two items (see above) were still negatively correlating with the total scale and hence the scale would have a negative α (α =-.113).

Altogether, the reliability of the Valence scales was not really convincing, so the scales were explored by means of factor analysis to see if there might be two (or more) factors underlying the supposed Valence scale.

Factor analysis of Valence scales

With 39 cases (4 participants times at least 9 independent sessions) the dataset did not fulfill the rule of thumb requirement for a reliable factor analysis of 200 cases, but taking the assumption of at least five cases per variable into account, the dataset of the current experiment was sufficient to do a factor analysis with about 8 items, and as the factor analysis was exploratory only, all ten items for the Positive-negative Valence scale were included.

Unrotated factor analysis , as well as factor analysis with Orthogonal (Varimax) and Oblique (Oblimin) rotation were applied to the baseline and the intervention measures of the Positive Valence scale (5 items) and the Positive-Negative Valence scale (11 items) in order to explore the possibility of two or three factors underlying the scales. Unfortunately no clear factors could be derived from the Valence items, neither on basis of extracting 2 nor on basis of extracting 3 fixed factors.

Construction of Valence scale

Based on all of the above, a final 6-item Positive-negative Valence scale was constructed (Cronbach's α =.527 for the intervention measure and Cronbach's α = .592 for the baseline measure). Excluding the items Eyebrows, Mouth corners, Laughing, and Touching, and including the items Mouth, Mouth

activity with no goal, Sounds without intent, Hands, Body, and Talking. Even though removing the items Talking and Sounds without intent would make the scale more reliable (α= .686 for the baseline measure and α= .699 for the intervention measure), this would mean that all sound-related items would be excluded from the scale. Based on the face validity of these items in this scale, it was decided to keep them included. However, because the item Talking had a corrected total-item correlation of -.315 in the intervention measure, the 5-item Positive-negative Valence scale without this item (Cronbach's α of .649 for the baseline measure and a Cronbach's α of .618 for the intervention measure) was also selected to be explored as dependent variable in the subsequent mixed model analysis. The two final Positive-negative Valence scales can be found in table 5.

3.1.1d Final observation measure

The final observation measure consisted of four scales: 1. the 6-item Attention scale (Table 3), 2. the 7-item Arousal scale (Table 4), and 3.& 4. the 6-item Positive-Negative Valence scale & the 5-item Positive-Negative Valence scale (Table 5).

Table 5

Final (Positive-negative) Valence scale

Items	Value	α
Sounds without intent	Positive (1) / neutral (0) /negative (-1)	Baseline α = .592 (6 items)
Mouth activity without a goal	Yes (-1) / no (0)	
Hands	Tensed (-1)/ normal (0)	Intervention α = .527 (6
Mouth	Tensed (-1) / normal (0)	items)
Body movement	Rocking (-1) / moving (0) / not moving (0)	
Talking	Positive (1) / neutral (0) / negative (-1) / not (0)	
Eyebrows	Excluded	*Excluding the item Talking*:
Mouth corners	Excluded	Baseline α = .649 (5 items)
Laughing	Excluded	
Touching	Excluded	Intervention α = .618 (5
		items)

3.1.2 Proxy measure: Interact Short and Interact During

The items of the Interact scale are normally reported on item-level, but in an attempt to minimize the number of analyses to be conducted on the Interact data, subscales for the items of the Mood,

Speech (for Interact During only), and Stimulation level domains were explored before the effect of condition on the Interact measure was investigated.

3.1.2a Interact Short measure

The caregiver who brought the participant to and collected her from the experimental sessions filled in the Interact Short concerning the 10 minutes before and the 10 minutes immediately after every session.

The items 'Confused' (Mood domain) and 'Wandering, restless or aggressive' (Stimulation Level domain) were never reported during the study, neither in the Interact Before nor in the Interact After measurement, and hence were excluded from further analysis. The 3 remaining items for Mood did not produce a reliable subscale, Cronbach's α =.165 for the Interact Before and Cronbach's α =.162 for the Interact After measure. The 3-item subscale for Stimulation Level was not reliable either, producing Cronbach's α =.212 for the Interact Before and Cronbach's α =.349 for the Interact After measure.

3.1.2b Interact During measure

The researcher completed the Interact During immediately after watching the video of the session. The Interact (Baker & Dowling, 1995) forms were especially developed for assessing the effects of interventions like Snoezelen, but due to the severity of the dementia level of the participants in the current study in combination with its experimental design which did not include the presence of a caregiver during the sessions, some of the (22) Interact During items showed to be non-applicable to the current study. Hence, the following items were excluded from further analysis: 'Recalled memories', 'Held eye contact appropriately', 'Touching', 'Related well', 'Co-operated', 'Touched objects/equipment appropriately', 'Comments or questions about activities/objects', and 'Did things from own initiative'.

Additionally, the items 'Recalled memories' and 'Bored, inactive or sleeping inappropriately' were never scored different from 'not at all' (the lowest item on the 5-point Likert scale) and thus excluded from further analysis. The 4-item subscale for Speech showed to be reliable (Cronbach's α= .801), the 4-item subscale for Mood had a Cronbach's α of .542, and the 3-item Stimulation Level subscale produced a Cronbach's α of .596.

3.2 Sensitivity of measures

The sensitivity of the two types of measures used in this study, 1. Structured behavioral observations, and 2. Proxy measures (Interact questionnaire filled out by caregivers and by researcher), was assessed by applying Linear Mixed Model Analysis to each of the measures to see if they were able to capture an effect of Stimulus condition.

In order to assess whether the variations in behavior during the sessions could be mainly explained by inter- or intra-personal variations of the participants, the above described LMM analyses were preceded by analysis of the corresponding null models to obtain the intraclass correlation (ICC). As the behaviors of the target group were expected to be idiosyncratic, a high ICC for the measures was anticipated. A high ICC would indicate that much of the variance of the model was due to variance within-participants, and hence reduce chances to detect an effect of Stimulus condition, as only a small amount of variance would be left to account for variations between participants.

3.2.1 Structured behavioral observation measure

To assess the sensitivity of the behavioral observation measure, the effect of Stimulus condition (Snoezel box vs. TV vs. Control) on each of the three observation scales, Attention, Arousal, and Valence, was explored by means of Linear Mixed Model analyses including Stimulus condition as factor and controlling for individual differences between participants by using participant number as subject variable.

Preceding the above described analysis, the null model was run in order to obtain the ICC for each of the observation scales (intervention measure). For the Attention scale ICC= 0.03, indicating that variations within participants accounted for only about 3 percent of the variation of the Attention scale. For the Arousal scale: ICC= 0.40, indicating that about 40 percent of the variation on the Arousal scale was due to intrapersonal differences and 60 percent due to differences between participants. For the Affect scale: ICC= 0.89 (6-item scale) and IC=0.88 (5-item scale) indicating that about 90 percent of the variation of the Affect scales was due to intrapersonal differences, due to which only about 10 percent of the total variation remains to detect an effect of Stimulus condition.

As expected, none of the baseline measures showed an effect on any of the three scales. For the intervention measures, no significant effect was found on the Attention scale, but when using the baseline measure as a covariate ($F(2,34)= 13.21$, $p=.001$) , the effect of Stimulus condition was

significant, F(2, 35) = 3.43, p=.043). When comparing means, the TV condition (M=22.8, SD=1.90) had the highest Attention score followed by the Snoezel box condition (M=18.8, SD=1.83) and the Control condition (M=15.8, SD=1.96). Pairwise comparisons showed that the difference between TV and control condition was significant (p= .013). The Snoezel box condition did not differ from either the TV condition (p=.132) or the Control condition (p=.265).

The effect of condition on Arousal showed a trend towards significance (F(2,35)= 2.75, p=.078), with the highest mean for the Snoezel box (M=12.4, SD=1.97), the lowest for the TV condition (M=9.0, SD=1.99), with the control condition (M=11.5, SD=2.00) in between. The difference between the TV condition and the Snoezel box condition was significant (p=.029), the difference between the TV and control condition was not (p=.116), and neither was the difference between the Snoezel box condition and the Control condition (p=.556). As the main effect of Stimulus condition was not significant these findings have to be treated as indications only. No effect of Stimulus condition was found on Valence (neither on the 6-item nor on the 5-item scale).

3.2.2 Interact Short (Before and After)

Based on the unsuccessful explorations for combining several Interact Short items into one subscale item, it was decided to explore the effect of Stimulus condition on each of the relevant 10 (of the original 12) Interact short items separately.

Linear mixed models analysis was applied with the Interact After item as the dependent, the appropriate Interact Before item as covariate, Stimulus condition as factor and participant number as subject variable. The null model that was run for all Interact Short items returned ICC values ranging from .00 to .36.

The only item that showed an effect of Stimulus condition was 'Talked spontaneously' (F(2,38)= 3.41, p =.043[2)]). The Snoezel box (M=2.6, SD=0.24) condition showed the highest speech score, followed by the Control (M=2.6, SD=0.24) and the TV (M= 1.8, SD=0.25) condition. Pairwise comparisons showed that the TV condition was significantly different from the Snoezel box condition (p=.025) and the Control condition (p=.033), while the Snoezel box condition and the Control condition were not significantly different (p=.886).

2) The final Hessian matrix was not positively definite although all convergence criteria were satisfied.

3.2.3 Interact During

Based on the explorations of the Interact During subscales, the effect of Stimulus condition was explored for the Mood, Speech, and Stimulation level subscales and for the remaining individual Interact During items by means of a linear mixed model analysis. The null model that was run returned ICC values <.1 for all three subscales. The null model that was run for the separate Interact During items returned ICC values ranging from .00 to .33.

The Speech, Mood, and Stimulation level subscales did not show an effect of Stimulus condition. The only individual item that showed an effect of Stimulus condition was 'Attentive/responding to/focused on activity/objects' ($F(1,22)= 5.783$, $p=.025$), in favor of the TV condition ($M=3.4$, $SD=0.51$) compared to the Snoezel box condition ($M=2.4$, $SD=0.51$). The control condition was excluded from this comparison as in that condition no stimulus was presented to the participant.

3.3 Convergent validity

In order to find out if the scores of the observational measure and the proxy measures related to each other, items that were presumed to be related (e.g. the behavioral Arousal scale and the Stimulation level subscale of the Interact) were standardized (all scores were turned into Z-scores) and added to a LMM with participant as subject variable, one of the behavioral scales as dependent variable, and the appropriate Interact subscale, or set of items, as covariate. An overview of the correlations between the three observational scales, Attention, Arousal, and Valence, and the corresponding Interact subscales can be found in Table 6.

None of the Interact Short subscales or items correlated with the observation scales. Both the 5-item ($F(1, 35)= 10.434$, $p=.003$, $\beta= .17$) and the 6-item ($F(1, 35)= 16.426$, $p<.000$, $\beta= .19$) Valence scale did correlate with the Interact During Mood subscale. The Arousal scale correlated with the Interact Stimulation Level subscale: $F(1,34)= 15.180$, $p<.000$, $\beta= .40$, and when adding the three separate items of this scale together as a covariate, the correlation of the item 'Enjoying self, active or alert' was significant ($F(1, 35)= 7.087$ $p=.012$, $\beta= .34$). When using the Attention scale as dependent variable and the Interact During items 'Attentive/responding to/focused on activity/objects' and 'Tracked observable stimuli' together as covariates, the correlation of 'Tracked observable stimuli' was significant ($F(1, 22)= 98.799$, $p<.000$, $\beta= .89$).

Table 6

Convergent validity of observation scales and Interact subscales

Observation scale	**Interact During**	**Interact Short**	**Observation & Interact D**	**Observation & Interact S**	**Interact D & Interact S**
Attention	Relating to environment: 14. Tracked observable stimuli 16. Attentive/responding to/focused on activity/objects	Relating to environment: 7. Attentive/responding to/focused on environment	14. Tracked observable stimuli: F(1, 22)= 98.799, p<.000, β= .89. 16. Attentive/responding to/focused on activity/objects: F(1,35)= 3.070, p=.092, β= .17	F(1, 35)= 0.269, p=.608, β=-.08	14. Tracked observable stimuli: F(1,23)= 0.124, p=.728, β= .09 16. Attentive/responding to/focused on activity/objects: F(1,23)= 1.231, p=.279, β= -.30
Arousal	Stimulation Level: 9. Wandering, restless or aggressive 10. Enjoying self, active or alert 12. Relaxed, content or sleeping appropriately	Stimulation Level: 10. Enjoying self, active or alert 11. Bored, inactive or sleeping inappropriately 12. Relaxed, content or sleeping appropriately	F(1,34)= 15.180, p<.000, β= .40	F(1, 34)= 1.210, p=.279, β=-.15	F(1,36)= 2.644, p=.113, β= -.28
Valence	Mood: 1. Tearful/Sad 2. Happy/Content 3. Fearful/anxious 4. Confused	Mood: 1. Tearful/Sad 2. Happy/Content 3. Fearful/anxious	5-items: F(1, 35)= 10.434, p=.003, β= .17 6-items: F(1, 35)= 16.426, p<.000, β= .19	5-items: (F(1, 33)= 2.992, p=.093, β= .08) 6-items: F(1, 33)= 2.497, p=.124, β=.08	F(1,36)= 0.078, p=.826, β= .04

4. Discussion

The goal of this study was to explore the possibilities for measuring the affective state of people with very severe dementia in the moment. In the literature many different measures were found that could be used to investigate related effects, such as mood or agitation, but no measure could fulfill this particular aim. The two measures, structured behavioral observations and proxy measures (Interact questionnaires), that were explored in this study did not fully succeed in this either, but their exploration produced valuable insights for future development of an affect measure for people with

very severe dementia. In this section first the results of the explorations of the sensitivity of the measures are discussed, followed by a reflection on how they relate to each other. Subsequently, the implications of these findings for the future development of the aspired tool for assessing the affective state of people with very severe dementia are discussed. Additionally, findings concerning providing this kind of visual stimuli to severe dementia patients are described, and finally insights and recommendations for future tests in a similar environment and for this specific target group are discussed.

4.1 Structured behavioral observation measure

The structured behavioral observation measure included three affect scales: an Attention, Arousal, and Valence scale. In this study, only the Attention and Arousal scales were able to capture an effect of Stimulus condition, although for the Arousal measure the effect was not significant but remained a trend ($p=.078$). The Valence measure did not capture an effect of Stimulus condition in this study.

The assessment of the intraclass correlations of the three scales showed that the Valence scale had a very high (>.85) intraclass correlation, indicating that most of the variance could be found within participants. This finding is possibly due to the idiosyncratic behavior of the participants, as during the sessions it was observed that one participant would for instance often show sucking or licking mouth behavior, while another always had her mouth closed, but did often rub or stroke her hands without a clear intent. This lack of consistency in the behaviors shown by the participants could indicate that the Valence scale is unsuitable as a general measure. A larger scale study has the potential to enhance the overall variety in participant behavior, increasing the probability that all behaviors on the scale will be displayed by more than one participant (which was clearly not the case in the present study). However, considering that the range of behaviors that a person with severe dementia is able to show will decline over time, it is to be expected that there will always be big variation in displayed behavior between participants and hence a between-subject design is not recommended for future studies with this target group.

While Attention and Valence scores are easily translated into their contribution to the level of positive affect, i.e. both high attention and high valence are contributing to a positive affect state, the way arousal relates to positive affect is less obvious. The Arousal score seems only interpretable in combination with an accompanied Valence score. A high level of Arousal in combination with a high level of Valence could indicate a highly positive affect state, while a high level of Arousal in

combination with a low (or negative) level of Valence could indicate a negative affect state. Hence, the level of arousal seems to work as an amplifier of the Valence score. In this study we were trying to evoke positive affect behaviors, indicating appreciation, like relaxation (low arousal) and happiness (high arousal), as opposed to negative affect behaviors like boredom (low arousal) or anger (high arousal). As the Valence measure was not sensitive enough to capture an effect of Stimulus condition in this study, it is difficult to interpret the found indication of an effect on the Arousal score.

In sum, it seems promising that two of the three scales of the structural observation measure were sensitive enough to capture an effect, or at least a trend towards an effect, of Stimulus condition in this small-scale study. In order to assess and improve the full functionality of this measure it should be applied to a study with more participants and more sessions. During the sessions was observed that there was a huge variety in the behaviors that were displayed by the four participants in this study. E.g. one of them was always very tense, would not move a lot and was easily captivated by the TV, while another seemed more restless and spent a big part of the time rearranging and fiddling with her clothes. Although the within-subject design of the study controlled for individual differences like these, a study with a larger number of participants would be helpful to establish a more comprehensive set of appropriate behaviors. Additionally, this would be a means to explore the weights that should be appointed to the items on the Attention, Arousal and Valence scales in order to make them more reliable.

4.2 Proxy measures: Interact Short and Interact During

Of the 12 items of the Interact Short questionnaire, only one showed an effect ($p = .043$) of Stimulus condition, namely 'Talked spontaneously'. Although spontaneous talk in general could be characterized as a positive behavior for this target group, the question is whether spontaneous talk after the intervention is also automatically an indicator of positive affect during the session. It might for instance be that the participant had a really nice experience during the sessions that made her enthusiastic, which made her more talkative afterwards. On the other hand, it could be that a participant had a nice experience, was satisfied and hence did not feel the need to talk, or that the participant did not particularly enjoy the session and felt kind of relieved when she was taken out of the situation and therefore started talking. This kind of uncertainty is inherent to measurement tools that measure the effects of an intervention after the intervention and not during the intervention itself. For most people behavior after an intervention will correlate with the behavior during the

session, but for people with very severe dementia this can not be assumed, as their sense of time has disappeared and they only live in and react on what happens in the moment. This might also be the reason why most items of the Interact Short did not manage to capture effects of Stimulus condition: caregivers who had filled out the Interact Short forms mentioned that they often did not observe a difference between the participant's behavior in the 10 minutes before and the 10 minutes after the session.

The original Interact During had 22 items, but due to the typical study set-up and specific target group, only 14 of them were applicable to this study. Of these 14 items, only the item 'Attentive/responding to/focused on activity/objects' showed an effect of Stimulus condition ($p=.025$). The Interact During was deliberately conducted in the way it was intended: the researcher scored the items after she had watched the patient's complete session on video. Even though the researcher had quite some experience with observing the behaviors of this particular target group, it was difficult to recall the occurrence and especially the approximate duration of the 14 different behavior items. (The items of the Interact have to be scored either Not at all, A bit of the time, Some of the time, Most of the time, or Nearly all the time.) This might be a reason for the low sensitivity of the Interact During in this study. The additional indication of change of the items during the session (Increasing, Decreasing, or Without a clear pattern) was not used during this study, mainly because it would have complicated the measure even more. In addition, in some sessions short moments of annoyed or irritated behavior were observed, a type of behavior which did not correspond to any of the items of the Interact.

In conclusion, it is questionable if proxy measures like the Interact Short are an appropriate means to assess the affect of people with very severe dementia, because of the uncertainty whether differences in a participant's behavior before and after the intervention are a valid indicator of the affective experience of the participant during the intervention. Moreover, the items of the Interact scales do not seem to contain the appropriate and comprehensive set of behaviors to cover the affect space for people with severe dementia. An explorative study focusing on observing and describing all possible affect related behaviors of this specific target group would be useful for heightening the reliability of this kind of proxy measures.

4.3 Comparing the measures

This study was initiated to find the most suitable measure for affect of people with very severe dementia. When looking at the abilities of the measures to capture affect in the moment, the

successfulness of the Interact Short is highly questionable, as it is not clear if a participant's experience during the session will carry over to his or her behavior after the session. The convergent validity exploration supports this idea, as none of the appropriate Interact Short subscales or individual items correlated with the observation scales. The subscales and appropriate individual items of the Interact During, which are scored based on the mood and behavior of the participant during the session, did correlate with the observation Items. However, the questions of the Interact During are answered after the session has taken place, as a result of which the interpretations of the researcher could be liable to cognitive biases like the peak-end rule (Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993) which makes that we judge experiences mainly upon how they peaked and how they ended. For this reason the structured observational measure is evaluated as the most promising measure for assessing the affect of people with very severe dementia in the moment.

4.4 Conclusion

Altogether, the findings of this study are rather promising. Although the user group was small (N=4) and personal variation in displayed behavior fairly large, three reasonably consistent behavioral item scales were produced. Contrary to most Snoezelen studies, the present study kept the conditions in the three conditions constant except for the type of device that was turned on, resulting in a relatively subtle manipulation. The fact that the developed behavioral item scales were sensitive enough to show an effect of this subtle manipulation is encouraging for the future development of this kind of real-time observational measures for this target group. In section 4.5 reflections on the construction of the structured behavioral observation scale are provided and commendations for the future development are presented.

In addition to the implications of the found effects for the developed measurement tool, the fact that the tool was able to detect a difference implies that the participants behaved differently and thus were sensitive to the manipulation of the experiment. This indicates that even very severe dementia patients are susceptible to these kinds of stimuli, which is an exciting finding by itself, but also a big motivation to carry on with the research to Snoezelen or other sensory stimulating interventions, and the development of supporting applications and products like the Snoezel box.

In sections 4.6 en 4.7 some additional findings of the study are discussed, relating to the provided visual stimuli in 4.5, and to conducting tests within this kind of settings and with this specific target group in 4.6.

4. 5 Future development of the structured behavioral observation scheme

The constructed domain scales of the Affect observation scale had fairly high alpha-values (>.6, <.7) Explorations of the scales showed that higher alpha values could have been obtained if more items were excluded, but this option was discarded, since the remaining scale did not represent the construct (Attention, Arousal, or Valence) it was supposed to measure. E.g. the Arousal scale would raise its alpha from .592 to .699 when excluding the items Talking and Sounds without intent, but the remaining scale lacked face validity, as it did not include any items regarding making sounds, which was considered a very relevant type of behavior for this target group. As an intermediate solution, the scale was explored both with the item Talking included (α= .527) and excluded (α= .618), but neither of these scales were able to show an effect of Stimulus condition. Additionally, the items Teary eyes and Yelling were excluded from analysis from the beginning because they did not occur during the current study. However, these behaviors were initially added to the observation scheme because their occurrence was observed during pilot sessions, just like the other behaviors included in the scheme. Hence, a more thorough exploration and investigation, by means of a larger scale study, of behaviors that cover the domains Attention, Arousal, and Valence should be conducted in order to explore and create an appropriate and valid set of behavioral items for each of the three affect scales.

Additionally, it should be noted that the sensitivity of the behavioral observation scale might have been influenced by the researcher. As the person who conducted the experiments also rated the behaviors on the behavior scale afterwards, the researcher was not blind to condition. However, in a bigger study, the gathering and the analysis of the data could easily be carried out by two different persons. It would then be interesting to investigate if an observation scale like the one that was explored in the current study is most reliable when making use of truly objective observers, i.e. people who are not only blind to condition, but who also have no prior knowledge about the behavior of the people they are rating, vs. observers who are also blind to condition, but have affinity with this particular target group, e.g. caretakers or researchers who have done field research in advance of the observations. This latter point links to the usability of the aspired observational measure. For purely scientific purposes a very extensive observation scheme, including items for all potentially occurring behaviors, might turn out to be the most valid measure to assess the affect or appreciation of people with severe dementia. However, for use in the field a smaller and hence more manageable set of behavioral items would probably more useful. If such an observational scale with high validity and reliability could be created, this measure could ultimately also be used by (in addition to researchers) caretakers in care facilities to validate and support their daily practices. Next to exploring the most

relevant behaviors, the second step towards the realization of such a measure would be that the behaviors are scored by different researchers to determine the inter-rater reliability of the items to create a valid and reliable behavior set and additionally to get insight in the level of interpretation (e.g. 'talks', 'talks positively', or 'tells a story') that is most appropriate for these behavioral items.

4.6 Visual Snoezelen stimuli

Outcomes of the observation measure showed that participants were more attentive in the TV condition than in the Control condition, while the Snoezel box condition did not significantly differ from any of the two conditions. This could be an indication that, even though the Snoezelen philosophy suggests otherwise, meaningful stimuli might have a positive effect on the affective state of people with very severe dementia. Of course it must be kept in mind that the provided meaningful stimulus had purely positive content and was easy to comprehend, i.e. participants did not need to follow a storyline as the scenes of baby penguins were apparently enjoyable by themselves. Another reason for this detected difference might be that a TV provides dynamic stimuli with strong contrasts, far stronger than the soft colors projected by the Snoezel box, which automatically draw visual attention.

The Arousal measure indicated ($p=.078$) that there was a difference in level of arousal between the Snoezel box and the TV condition. Participants in the Snoezel box condition showed more or higher arousal behavior than participants in the TV condition. Because no effect on Valence was found, this indication is difficult to interpret, but because none of the participants showed truly negative behavior during the sessions, one can suggest that the Snoezel box has the potential to stimulate positive affect behavior, by evoking arousal in people with severe dementia.

Additionally, the Interact During showed that participants were more attentive to the TV than to the Snoezel box ($p=.025$). This might be an indication that a TV or other static device providing dynamic stimuli is able to captivate someone's attention, while the Snoezel box, which projections are moving through the room, might stimulate people to look around and hence to be distracted from the provided stimulus to other things in the room. Both types of visual stimulus and their influence on the behavior and affect of people with dementia might be interesting to explore in more depth in future Snoezel studies.

4.7 Testing with this specific target group and in this specific setting

To my knowledge, this study was the first to assess the effect of an intervention on this target group without the confounding presence of a caregiver in the test situation. Previous Snoezelen studies tried to control for this confound by including a control condition with comparable social attention, i.e. a caregiver would be present in this condition as well. It is imaginable though that the different conditions could have influenced the behavior of the caregiver in the sessions, which in turn could have influenced the behavior of the participant. To rule out this possibility, the current study was designed in such a way that the caregiver would only bring and collect the participant, and the participant would be alone in the room during the session. However, while spending time in the living room and attending Snoezel group sessions at care center Vonderhof, it was observed that people with severe dementia respond the most, and usually positive, to social attention. Therefore it is important to carefully consider the pros and cons of setting up a test with or without a caregiver before deciding on the final experiment design.

When planning to conduct a scientific study in a care center environment one has to keep in mind that the design of the study has to suit the daily practice in the care facility. It has to fit into the daily routines of the inhabitants, but also in the schedules of the caregivers. It must be kept in mind that the priority of the caregivers of the care center is providing the inhabitants with the best possible care and not participating in a scientific study. At this moment, the Netherlands deals with a structural shortage of professional caregivers in care centers. The observed general attitude of the caregivers confirmed this national trend: More than once it became clear that they felt like they did not have enough time to provide the (quality of) care that they wanted to give. Hence, their time is precious and should be treated accordingly when designing the study.
Additional note: Even when the study design is set up around the schedules of the caregiver, it should be born in mind that a caregiver's job is never the same and that his or her tasks are dependent on the state and needs of the inhabitants. During the current study for instance, an inhabitant (who did not participate in the study) had been refusing to eat for days and the caregiver that was scheduled to bring and collect a participant that morning had to help solving the situation.

Although people with severe dementia mainly live in the moment, their circadian rhythm in combination with reoccurring events like daily meals and weekly bathing sessions might have an influence on their general mood and behavior, which can add extra noise to the measurement. A future study should try to control for this within participant variability. In the current study this was

done by scheduling participants' sessions at a fixed time and using the first six minutes of the session as baseline.

Physiological data could potentially be a valuable addition to the more subjective measures like observations and questionnaires that are generally used in studies like the present one. Yet, obtaining physiological data is normally an obtrusive procedure, with sensors that have to be attached to the participant's body, to which you don't want to expose a vulnerable target group like dementia patients. However, Philips is developing a software program (Vital Signs) which can deduce both heart and respiration rate from video material, which makes it possible to obtain this physiological data in a non-obtrusive way. Unfortunately the development of the Vital Signs software did not progress as quickly as expected, as a result of which no physiological data could be assessed in this study, but this application has the potential to be a valuable addition to future studies.

5. References

Alexopoulos, G. A., Abrams, R. C., Young, R. C., & Shamoian, C. A. (1988). Cornell scale for Depression in Dementia. *Biological Psychiatry, 23*, 271-284.

Alzheimer's Association. Seven Stages of Alzheimer's. Retrieved on 10-05-2012 from: http://www.alz.org/alzheimers_disease_stages_of_alzheimers.asp

Alzheimer's Association. (2011). 2011 Alzheimer's Disease Facts and Figures. Retrieved on 10-05-2012 from: http://www.alz.org/documents_custom/2011_Facts_Figures_Fact_Sheet.pdf

Alzheimer's Disease International. (2009). World Alzheimer's Report 2009. London, ADI. Retrieved on 10-05-2012 from http://www.alz.co.uk/research/files/WorldAlzheimerReport.pdf

Alzheimer's Disease International. (2012). Alzheimer's Disease. Retrieved on 10-05-2012 from http://www.alz.co.uk/info/alzheimers-disease

Baillon, S., Van Diepen, E., Prettyman, R., Redman, J., Rooke, N., & Campbell, R. (2004). A comparison of the effects of Snoezelen and reminiscence therapy on the agitated behaviour of patients with dementia. *International Journal of Geriatric Psychiatry, 19*, 1047–1052. DOI: 10.1002/gps.1208

Baker, R., Bell, S., Baker, E., Gibson, S., Holloway, J., Pearce, … Wareing, L. (2001). A randomized controlled trial of the effects of multi-sensory stimulation (MSS) for people with dementia. *British Journal of Clinical Psychology, 40*, 81-96.

Baker, R., Dowling, Z., Wareing, L. A., Dawson, J., & Assey, J. (1997). Snoezelen: Its long-term and short-term effects on older people with dementia. *British Journal of Occupational Therapy*, *60*, 213-218.

Baker, R., Holloway, J., Holtkamp, C. C. M., Larsson, A., Hartman, L. C., Pearce, R., (Scherman, B., Johansson, S., Thomas, P. W., Wareing, L. A.,) & Owens, M. (2003). Effects of multi-sensory stimulation for people with dementia. *Journal of Advanced Nursing, 43(5)*, 465–477.

Barofsky, I. (2012). Can Quality or Quality-of-Life be defined? *Quality of Life Research, 21*, 625-631.

Berg, L. (1988) Clinical dementia rating. *Psychopharmacological Bulletin, 24*, 637-39.

Bloch, S., Lemeignan, M., & Aguilerat, N. (1991). Specific respiratory patterns distinguish among human basic emotions. *International Journal of Psychophysiology, 11*, 141-154.

Boiten, F. A., Frijda, N. H., Wientjes, C. J.E. (1994). Emotions and respiratory patterns: review and critical analysis. *International Journal of Psychophysiology, 17*(2), 103-128. DOI: 10.1016/0167-8760(94)90027-2.

Brod, M., Stewart, A. L., Sands, L., & Walton, P. (1999). Conceptualization and Measurement of Quality of Life in Dementia: The Dementia Quality of Life Instrument (DQoL). *The Gerontologist, 39*(1), 25-35.

Bureau-Chalot, F., Novella, J. L., Jolly, D., Ankri, J., Guillemin, F., & Blanchard, F. (2002). Feasibility, Acceptability and Internal Consistency Reliability of the Nottingham Health Profile in Dementia Patients. *Gerontology, 48*(4), 220-225.

Codispoti, M., Surcinelli, P., Baldaro, B. (2005). Watching emotional movies: Affective reactions and gender differences. *International Journal of Psychophysiology, 69*, 90-95.

Cohen-Mansfield, J., Marx, M. S., & Rosenthal, A. S. (1989). A Description of Agitation in a Nursing Home. *Journal of Gerontology: Medical Sciences, 44*(3), M77-M84.

Cohn, J., Ambadar, Z, & Ekman, P. (2006). Observer-based measurement of facial expression with the Facial Action Coding System. In: J. A. Coan, & J. B. Allen (Eds.), *The handbook of emotion elicitation and assessment* (pp. 203-221). New York, NY: Oxford University Press Series in Affective Science.

Collier, L., McPherson, K., Ellis-Hill, C., Staal, J., & Bucks, R. (2010). Multisensory Stimulation to Improve Functional Performance in Moderate to Severe Dementia - Interim Results. *American Journal of Alzheimer's Disease and Other Dementias, 25*(8), 698-703.

Cooper, C., Mukadam, N., Katona, C., Lyketsos, C. G., Ames, D., Rabins, … Livingston, G. (2012). Systematic review of the effectiveness of non-pharmacological interventions to improve quality of life of people with dementia. *International Psychogeriatrics 24*, 856-878.

Cotelli, M., Manenti, R., & Zanetti, O. (2012) Reminiscence Therapy in Dementia: A Review. *Maturitas, 72*(3), 203-205.

De Boer, A. G. E. M., Van Lanschot, J. J. B. Stalmeier, P. F. M., Van Sandick, J. W., Hulscher, J. B. F., De Haes, J. C. J. M., & Sprangers, M. A. G. (2004). Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? *Quality of Life Research, 13,* 311-320.

De Jonghe, J.F., & Kat, M. G. (1996). Factor structure and validity of the Dutch version of the Cohen-Mansfield Agitation Inventory (CMAI-D). Journal of the American Geriatrics Society, 44, 888-889.

Edelman, P., Fulton, B. R., Kuhn, D., & Chang, C. (2005). A Comparison of Three Methods of Measuring Dementia-Specific Quality of Life: Perspectives of Residents, Staff, and Observers. *The Gerontologist, 45*, 27-36.

Ettema, T. P., Dröes, R. M., de Lange, J., Mellenbergh, G. J., & Ribbe, M. W. (2006). QUALIDEM: Development and evaluation of a Dementia Specific Quality of Life Instrument. Scalability, Reliability and Internal Structure. *International Journal of Geriatric Psychiatry, 22*(6), 549-56.

Fernandez, C., Pascual, J. C., Soler, J., Elices, M., Portella, M. J., & Fernandez-Abascal, E. (2012). Physiological Responses Induced by Emotion-Eliciting Films. *Applied Psychophysiology and Biofeedback, 37*(2), 73-79.

Ferri, C. P., Prince, M., Brayne, C., Brodaty, H., Fratiglioni, L., Ganguli, M., . . . Scazufca, M. (2005). Global Prevalence of Dementia: a Delphi Consensus Study. *The Lancet*, 366(9503), 2112-2117.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-98.

Frazier, T. W., Strauss, M. E, & Steinhauer, S. R. (2004). Respiratory sinus arrhythmia as an index of emotional response in young adults. *Psychophysiology, 41*, 75–83. DOI: 10.1046/j.1469-8986.2003.00131.x

Gomez, P., Zimmermann, P., Guttormsen-Schar, S., & Danuser, B. (2005). Respiratory responses associated with affective processing of film stimuli. *Biological Psychology, 68*, 223-235.

Gross, J. J., & Levenson, R. W. (1995). Emotion Elicitation using Films. *Cognition and Emotion, 9*(1), 87-108.

Hamer, T. (2003). Belevingsgerichte Zorg in de Praktijk: Gewoon Doen! *Tijdschrift voor Verpleeghuisgeneeskunde, 27*(5), 19-23.

Herring, D. R., Burleson, M. H., Roberts, N. A., & Devine, M. J. (2010). Coherent with Laughter: Subjective Experience, Behavior, and Physiological Responses during Amusement and Joy. *International Journal of Psychophysiology, 79*, 211-218.

Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., & Martin, R. (1982) A new Clinical Scale for the Staging of Dementia. *British Journal of Psychiatry, 140*, 566-72.

Hulsegge, J. & Verheul, A. (1987). Snoezelen, Another World. Rompa, Chesterfield.

Hurley, A. C., Volicer, B. J., Hanrahan, P. A., Houde, S., & Volicer, L. (1992). Assessment of Discomfort in Advanced Alzheimer Patients. *Research in Nursing & Health, 15*(5), 369-377.

Kaplan, H., Clopton, M., Kaplan, M., Messbauer, L., & McPherson, K. (2006). Snoezelen Multi-Sensory Environments: Task Engagement and Generalization. *Research in Developmental Disabilities, 27,* 443-455.

Kelley-Gillespie, N. (2009). An Integrated Conceptual Model of Quality of Life for Older Adults based on a Synthesis of the Literature. *Applied Research Quality Life, 4*, 259-282.

Kitwood, T., & Bredin, K. (1992). A New Approach to the Evaluation of Dementia Care. *Journal of Advances in Health and Nursing Care, 1*(5), 41-60.

Kahneman,D., Fredrickson,B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more Pain is preferred to less: Adding a better End. *Psychological Sciene, 4*(6), 401-405.

Klages, K., Zecevoc, A., Orange, J. B., & Hobson, S. (2011). Potential of Snoezelen Room Multisensory Stimulation to improve Balance in Individuals with Dementia: a Feasibility randomized controlled Trial. *Clinical Rehabilitation, 25*(7), 607-616.

Lancioni, G., Cuvo, A., & O'Reilly, M. (2002). Snoezelen: an Overview of Research with People with Developmental Disabilities and Dementia. *Disability & Rehabilitation, 24*(4), 175-184.

Lawton, M. P., Moss, M. S., Fulcomer, M., & Kleban, M. H. (1982). A research and service oriented multilevel assessment instrument. *Journal of Gerontology, 37*(1), 91-99.

Logsdon, R. G, Gibbons, L. E., McCurry, S. M.,& Ten, L. Quality of life in Alzheimer's disease: Patient and Caregiver Reports. *Journal of Mental Health and Aging, 5*(1), 21-32.

Martin, N. T., Gaffan, E. A., & Williams, T. (1998). Behavioural effects of long-term multi-sensory stimulation. *British Journal of Clinical Psychology, 37*(1), 69–82.

McKee, S., Harris, G., Rice, M., & Silk, L. (2006). Effects of a Snoezelen room on the behavior of three autistic clients. *Research in Developmental Disabilities, 28*, 304–316.

Neumann, P. J., Kuntz, K. M., Leon, J., Araki, S. S., Hermann, R. C., Hsu, M-A., & Weinstein, M. C. (1999). Health utilities in Alzheimer's disease: A cross-sectional study of patients and caregivers. *Medical Care, 37*(1), 27-32.

Neumann, P. J., Sandberg, E. A., Araki, S. S., Kuntz, K. M., Feeny, D., & Weinstein, M. C. (2000). A comparison of HU12 and HU13 utility scores in Alzheimer's disease. *Medical Decision Making, 20*(4): 413-422.

Novella, J., Ankri, J., Morrone, I., Guillemind, F., Jolly, D., Jochum, C., ... Blanchard, F. (2001). Evaluation of the quality of life in dementia with a generic quality of life questionnaire: The Duke Health Profile. *Dementia and Geriatric Cognitive Disorders, 12*(2), 158-166.

O'Boyle, C. A. (1994). The Schedule for the Evaluation of Individual Quality of Life (SEIQoL). *International Journal of Mental Health, 23*(3): 3-23.

Perneczky, R., Wagenpfeil, S., Komossa, K., Grimmer, T., Fiehl, J., & Kurz, A. (2006). Mapping Scores onto Stages: Mini-Mental State Examination and Clinical Dementia Rating. *American Journal of Geriatric Psychiatry, 14*(2), 139-144.

Perrin, T. (1997). The positive response schedule for severe dementia. *Aging and Mental Health,1*, 184-197.

Pettit, T., Livingston, G., Manela, M., Kitchen, G., Katona, C., & Bowling, A. (2001). Validation and normative data of health status measures in older people: The Islington study. *International Journal of Geriatric Psychiatry, 16*(11), 1061-1070.

Philippot, P. (1993). Inducing and assessing differentiated emotion-feeling states in the laboratory. *Cognition & Emotion, 7*(2), 171–193.

Pool, J. (1999). The pool activity level (PAL) instrument. London, England: Jessica Kinsley Publishers.

Porzsolt, F., Kojer, M., Schmidl, M., Greimel, E. R., Sigle, J., Richter, J., & Eisemann, M. (2004). A New Instrument to Describe Indicators of Well-being in Old-old Patients with Severe Dementia-The Vienna List. *Health Quality of Life Outcomes, 2*, 10-18.

Rabins, P. V., Kasper, J. D., Kleinman, L. , Black, B. S., and Patrick, D. L. (1999). Concepts and Methods in the Development of the ADRQL: An instrument for Assessing Health-Related Quality of Life in Persons with Alzheimer's Disease. *Journal of Mental Health and Aging, 5*(1), 33-48.

Rainville, P., Bechara, A., Naqvi, N., & Damasio, A. R. (2006). *International Journal of Psychophysiology, 61*, 5-18.

Ready, R. E., Ott, B. R., Grace, J., & Fernandez, I. (2002). The Cornell-Brown Scale for Quality of Life in Dementia. *Alzheimer Disease & Associated Disorders, 16(*2), 109-115.

Robichaud, L., Hébert, R., & Desrosiers, J. (1994). Efficacy of a sensory integration program on behaviors of inpatients with dementia. *American Journal of Occupational Therapy, 48*(4), 355-60.

Rosen, J. M. D., Burgio, L., Kollar, M. M. S. N., Cain, M. B. S. N., Allison, M. R. N., Fogleman, M. B. S. N., … Zubenko, G. S. M. D. (1994). The Pittsburgh Agitation Scale: A user-friendly instrument for rating agitation in dementia patients. *The American Journal of Geriatric Psychiatry, 2*(1), 52-59.

Russell, A. R., & Mehrabian, A. (1977). Evidence for a Three-Factor Theory of Emotions. *Journal of Researcj in Personality, 11*, 273-294.

Schölzel-Dorenbos, C. J. M., Ettema, T.P., Bos, J., Boelens-van der Knoop, E., Gerritsen, D.L., Hoogeveen, F., … Dröes, R. M. (2007). Evaluating the Outcome of Interventions of Quality of Life in Dementia: Selection of the Appropriate Scale. *International Journal of Geriatric Psychiatry, 22*, 511-519.

Selai, C. E., Trimble, M. R., Rossor, M. N., & Harvey, R. J. (2001). Assessing Quality of Life in Dementia: Preliminary Psychometric Testing of the Quality of Life Assessment Schedule (QOLAS). *Neuropsychological Rehabilitation, 11*(3-4), 219-243.

Smith, S. C., Murray, J., Banerjee, S., Foley, B., Cook, J. C., Lamping, D. L., … Mann, A. (2005). What constitutes health-related quality of life in dementia? Development of a Conceptual Framework for People with Dementia and their Carers. *International Journal of Geriatric Psychiatry, 20*, 889-895.

Staal, J. A. (2012). Functional analytic multisensory environmental therapy for people with dementia. International Journal of Alzheimer's Disease, doi:10.1155/2012/294801

Staal, J. A., Sacks, A., Matheis, R., Collier, L., Calia, T., Hanif, H., & Kofman, E. S. (2007). The Effects of Snoezelen (Multi-Sensory Behavior Therapy) and psychiatric care on agitation, apathy, and activities of daily living in dementia patients on a short-term geriatric psychiatric inpatient unit. *International Journal of Psychiatry in Medicine, 37*(4), 357-370. doi: 10.2190/PM.37.4.a

Terada, S., Ishizu, H., Fujisawa, Y., Fujita, D., Yokota, O., Nakashima, H., … Kuroda, S. (2002). Development and Evaluation of a Health-Related Quality of Life Questionnaire for the Elderly with Dementia in Japan. *International Journal of Geriatric Psychiatry, 17*(9), 851-858.

Teri, L., & Logsdon, R. G. (1991). Identifying pleasant activities for Alzheimer's disease patients: The pleasant events schedule-AD. *Gerontologist, 31*(1), 124-127.

The WHOQOL Group. (1998). Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine, 28*(3), 551-558

Trigg, R., Watts, S., Jones, R., & Tod, A. (2011). Predictors of quality of life ratings from persons with dementia: the role of insight. *International Journal of Geriatric Psychiatry, 26*, 83-91.

Van Diepen, E., Baillon, S. F., Redman, J., Rooke, N., Spencer, D. A., & Prettyman, R. (2002). A Pilot Study of the Physiological and Behavioural Effects of Snoezelen in Dementia. *The British Journal of Occupational Therapy, 65*(2), 61-66.

Van Weert, J., Van Dulmen, S., Kerkstra, A., Spreeuwenberg, P., Ribbe, M., & Bensing, J. (2003). Effecten van snoezelen op demente ouderen en zorgverleners: samenvatting van de onderzoeksresultaten (in Dutch). NIVEL. www.nivel.nl

Van Weert, J. C., Van Dulmen, A. M., Spreeuwenberg, P. M., Ribbe, M. W., & Bensing, J. M. (2005). Behavioral and mood effects of snoezelen integrated into 24-hour dementia care. *Journal of the American Geriatrics Society*, *53*(1), 24-33.

Verstraten, P. F. J., & Van Eekelen, C. W. J. M. (1987). Manual for the BIP. Dutch Behavior Observation Scale for Psychogeriatric In-patients [Handleiding voor de GIP. Gedragsobservatieschaal voor de Intramurale Psychogeriatrie]. Deventer: Van Loghum Slaterus, 1987.

Vos, P., De Cock, P., Munde, V., Petry, K., Van Den Noortgate, W., & Maes, B..(2012). The tell-tale: What do heart rate; skin temperature and skin conductance reveal about emotions of people with severe and profound intellectual disabilities? Research in Developmental Disabilities, 33(4), 1117–1127.

Westermann, R., Spies, K., Stahl, G., & Hesse, F. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology, 26*, 557–580.

6. Appendices

Appendix A. Measures in Snoezelen studies

Baker et al. (2001)		**INTERACT** (22)	**INTERACT-SHORT** (12)	**REHAB**: (Baker & Hall, 1983) **General Behaviour** subscale and **Deviant Behaviour** subscale	**Behaviour and Mood Disturbance scale** (BMD scale)	**Clifton Assessment Procedures for the Elderly Behaviour Rating** Scale (BRS)	**Mini Mental State Examination**	**CAPE**: **Cognitive Assessment** Scale (CAS)	
	-	behavior during session, completed immediately after a session by keyworker	baseline mood and behavior 10 min. before & 10 min. after session, completed by keyworker	observation of behavior within the normal regime of the day hospital, by two members of staff	behavior and mood at home, completed by carers, with the aid of a research assistant	change in behavior in the home setting	Cognitive assessment, by psychology research assistants	Cognitive assessment, by psychology research assistants)	
Baker et al. (1997)		" "		" "	" "	" "	" "	" "	
Baker et al. (2003)									**Gedragsobservatieschaal voor de Intramurale Psychogeriatrie**, GIP (NL)
				" "	" "	" "	" "		Three observation scales: apathy, cognition, affect. Completed by psychologist or caregiver
Staal et al. (2007)		**KATZ index** of activities of daily living **(KI-ADL)**	**Refined Activities of Daily Living** assessment scale (RADL)	**Scale for assessment of negative symptoms in Alzheimer's Disease (SANS-AD)**	**Pittsburgh Agitation Scale (PAS**)	**Global Deterioration Scale (GDS)**	**Beck Dressing Performance Scale (BDP)**	**MMSE**	
		Bathing, dressing, toileting etc. Completed post intervention, by nurses	Activities of daily living Completed post intervention, by nurses	Measure negative symptoms	Assess agitation Measured after sessions by researchers	Determine stage of illness	Dressing ability Completed after sessions by researchers	Cognitive functioning (base measure)	
Collier et al. (2010)		**AMPS**					**Gottfries**	**Standardized**	

							Brane Steen scale (**GBS**)	**Mini-Mental State Examination (SMMSE)**	
		Baseline of functional performance Used pre and postsession .					Physical inactivity, intellectual impairment, emotional and cognitive impairment	Degree of physical inactivity, intellectual impairment, emotional and cognitive impairment	
Robbins & Norton (2011)		**MDS** Quarterly Minimum Data Set assessment	**CMAI** (Cohen-Mansfield Agitation Inventory)	**AFABS** Adult Functional Adaptive Behavior Scale	Data from **clinical record**				
			27 agitation-related behaviors		Nr of depression symptoms on MDS, necessary (PRN) medications, activities attended/refused, agitation indicators, falls, etc				
Baillon et al. (2004)		**Mini-Mental State Examination** (Folstein et al., 1975)	**Clinical Dementia Rating Scale** (Hughes et al., 1982)	**Cohen-Mansfield Agitation Inventory** short form (14 items) (Cohen-Mansfield et al., 1989a	**Agitation Behaviour Mapping Instrument**. (Cohen-Mansfield et al., 1989b; Cohen-Mansfield, 1986)—	**Interact scale**. (Baker and Dowling, 1995)— *Secondary outcome*	**Heart rate.** *Secondary outcome*		
		Cognitive impairment	Dementia severity	Frequency of agitated behavior at the care unit	Observational agitation instrument	The subjects' behaviour Measured during each session	heart-rate at one-minute intervals from 15 min before, until 30 minutes after the sessions		

Klages et al. (2011)		**Functional Reach Test**	**Sharpened Romberg**	**Timed Up and Go Test** with and without a cognitive dual task		Secondary outcome measures included **frequency of falls** recorded in the pre-, during and post-intervention periods.			
Van Weert et al. (2005)		Parts of the **Dutch Behavior Observation Scale for Psychogeriatric In-patients** (BIP)	**CMAI (Dutch version)**	**The Cornell Scale for Depression in Dementia** (Dutch version) (CSDD-D)	**(extended) INTERACT**	**Eight study-specific items**, based on the observation form of Bernardus Expertise Center/Fontis and literature	**Three face diagrams** (FACE)		
		Nonsocial, apathetic, distorted behavior, consciousness, loss of decorum, anxious, rebellious, behavior, restless, and disoriented behavior		Mood-related signs, behavioral disturbance, physical signs	Mood and behavior, Based on videos of morning care	Mood and behavior, Based on videos of morning care	Mood		
Van Diepen et al.(2002)		**Agitation Behaviour Mapping Instrument** (ABMI)	**Short-formCohen-Mansfield Agitation Inventory** (CMAI)	**Heart rate recording**	**INTERACT**	**Clinical Dementia Rating scale** (CDR)			
		frequency of agitated behaviour during 3-minute episodes by direct observation, each session	Agitation over previous 2 weeks	Recorded from 10 min before, during the session and until 30 min after the session, at one minute intervals.	Effects of Snoezelen, Completed immediately after each session by the therapist	Dementia severity			

Appendix B. Observation scheme

Attention/ **Ar**ousal/ **Va**lence **H**igh/ **M**edium/ **L**ow	Attention areas		Baseline				Intervention						
			2 min	4 min	6 min		2 min	4 min	6 min	8 min	10 min	12 min	14 min
Ar-M, At-L	Eyes	Open											
Ar-L/M, At-M		Closed/open											
Ar-L, At-L		Closed											
Va-L/H		Teary/watery											
Va-M		Dry/normal											
At-H		Following/focused (directed)											
At-M/L		Unfocused/blank (staring at nothing)											
Va-H		Eyebrows up											
Va-M		Eyebrows neutral											
Va-L		Eyebrows down (frown)											
Ar-H	Attention	Directed towards application											
Ar-L/M		Not directed to application											
Va-L	Mouth	Tensed (open or closed)											
Va-M/H		Normal (open or closed)											
Ar-H, Va-L		Mouth activity, without intenten (licking, sucking)											
Ar-M/L, Va-M/H		No mouth activity without intent											
Va-H		Smiling (corners up)											
Va-M		Neutral (not really up or down)											
Va-L		Unhappy (corners											

		down)											
At-L, Ar-L	Sounds	No sound											
At-H, Ar-H, Va-H		Laughing											
		Not laughing											
At-H, Ar-H, Va-H		Yelling positive											
At-H, Ar-H, Va-L		Yelling negative											
At-H, Ar-M, Va-H		Talking/babbling (intent to communicate) positive											
At-H, Ar-M, Va-M		Talking/ babbling neutral											
At-H, Ar-M, Va-L		Talking/ babbling negative											
At-L, Ar-M, Va-H		Sounds without intent-positive											
At-L, Ar-M, Va-M		Sounds without intent -neutral											
At-L, Ar-M, Va-L		Sounds without intent -negative											
Ar-H	Respiration	Quick breathing											
Ar-L/M		Relaxed/normal breathing											
Ar-M/H, At- M/H	Head	Active/Upright position (chin up)											
Ar-L, At-L		Passive/hanging position (chin down)											
Ar-L		No head movement											
Ar-M		Slow head movement											
Ar-H		Quick head movement											
Ar-H	Hands	Tensed											
Ar-L/M		Relaxed											
Ar-L		No movements											

Ar-M		Slow movements											
Ar-H		Quick movements											
Ar-L, At-L, Va-M		Not actively touching											
Ar-H, At-H, Va-L		Actively touching without a clear intent/goal											
Ar-H, At-H, Va-H?		Actively touching with an intent/goal											
Ar-L, Val-M/H		Not moving upper body											
Ar-H, Val-L		Rocking, rhythmically moving upper body											
Ar-H, Val-M/H		Moving upper body											