ENGAGE-DEM

A Model of Engagement of People with Dementia

GIULIA PERUGIA

ENGAGE–DEM:
A MODEL OF ENGAGEMENT OF PEOPLE WITH DEMENTIA

Giulia Perugia

ISBN: 978-90-386-4609-1

Cover design: ProefschriftMaken
Cover pictures: ©pngtree.com
Lay-out: RON Graphic Power, www.ron.nu
Printing: ProefschriftMaken || www.proefschriftmaken.nl

ENGAGE–DEM:
A MODEL OF ENGAGEMENT OF PEOPLE WITH DEMENTIA

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens,
voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op woensdag 24 oktober 2018 om 11:00 uur

door

Giulia Perugia
geboren te Rome, Italië

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	Prof.dr. L.L. Chen
1^{e} promotor:	Prof.dr. A. Català-Mallofré (Universitat Politècnica de Catalunya)
2^{e} promotor:	Prof.dr. G.W.M. Rauterberg
1^{e} copromotor:	Prof.dr. M. Díaz-Boladeras (Universitat Politècnica de Catalunya)
2^{e} copromotor:	Dr.ir. E.I. Barakova
leden:	Prof.dr.ir. P. Markopoulos
	Prof.dr.ir. L.M.G. Feijs
	Prof.dr. A.S. Rigaud MD (Université René Descartes – Paris 5)

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

The work in this dissertation has been produced under the auspices of the Erasmus Mundus Joint Doctorate Program in Interactive and Cognitive Environments. The research was conducted towards a joint double PhD degree affiliated with the following partner universities:

UNIVERSITAT POLITÈCNICA DE CATALUNYA
TECHNISCHE UNIVERSITEIT EINDHOVEN

Acknowledgments

This PhD Thesis has been developed in the framework of, and according to, the rules of the Erasmus Mundus Joint Doctorate on Interactive and Cognitive Environments EMJD ICE [FPA n° 2010-0012] with the cooperation of the following Universities:

Alpen-Adria-Universität Klagenfurt – AAU

Queen Mary, University of London – QMUL

Technische Universiteit Eindhoven – TU/e

Università degli Studi di Genova – UNIGE

Universitat Politècnica de Catalunya – UPC

According to ICE regulations, the Italian PhD title has also been awarded by the Università degli Studi di Genova.

To Bibi.

Cognitive Communication Conference, London 2012

Lotje: I think one of the challenges, probably, for therapists is to deal with the fact that you have to work with somebody who's being assessed and defined by its limitations (…) and the experience of continually being defined by what you can no longer do or how you're sort of limited is devastating.

Therapist: Is there anything you could advise us, as therapists, to help you to get through that? Or do you think there was no other way that we could have done it?

Lotje: To maybe focus on things that a patient might find along the way. Where they're always initially defined by what they can't do, they may discover something that they weren't expecting at all, which is certainly what happened in my case.

Lotje Sodderland (1 year from stroke)
From the documentary "My beautiful broken brain", 2016

The word emotion stems from the Latin *movere*, meaning to move.
When emotions are intense, people move: they act, they react,
sometimes dramatically, as in crimes of passion.
(Bradley and Lang, 2007)

And yet *they* move.

Summary

ENGAGE-DEM: A Model of Engagement of People with Dementia

Dementia is a neurodegenerative disease that affects *cognition*, producing a reduction in thinking, problem-solving, and mnemonic abilities, *functioning*, preventing affected people to care for themselves and carry out activities of daily living (e.g., self-feeding, dress oneself), and *psychosocial well-being*, causing the appearance of disorders of thought content, mood, and behavior (e.g., depression, apathy, anxiety).

The reduction of cognition, the disorientation in space and time, and the inability to complete basic tasks and function independently are the major causes of institutionalization in dementia. Care facilities are extremely efficient in meeting the physical and environmental needs of persons with dementia (e.g., food, self-care, drugs). However, they often fail in addressing mental and social needs. Several studies show that people with dementia living in institutionalized contexts spend most of their time inactive and isolated. Just as many studies demonstrate that engagement in playful activities is crucial to ensure quality of life and psychosocial well-being in dementia.

On the one hand, the centrality of the medical aspects of dementia is due to the legacy of the biomedical approach to care. On the other hand, it is caused by the objective difficulty of understanding the needs of somebody who struggles to communicate. As a matter of fact, we are in great need of models enabling us to make meaning of how people with dementia express their psychological states.

According to the literature, engagement is the psychological state of proactive involvement with an object (e.g., a game, an interactive system) or an agent (e.g., a person, a social robot) that has a positive affective nuance. In healthy adults, engagement can be measured on three different levels, according to three distinct response systems: experiential/subjective (i.e., self-reports), behavioral/expressive (i.e. overt behavior), and peripheral-physiological

(i.e., human bodily reactions). In adults with dementia, the *experiential/subjective* level is rarely accessible due to cognitive impairment, the *behavioral/expressive* level might be blunted by motivational disorders like apathy and depression, and the *peripheral-physiological* level might provide insightful results, but is often overlooked. As the three response systems are not always equally accessible, it becomes crucial to combine them to achieve the most exhaustive possible measurement of engagement in people with dementia.

The objective of this dissertation is twofold. First, it aims at exploring new techniques to assess engagement in dementia with the help of unobtrusive physiological sensors and systematic behavior observation. Second, it focuses on the development of a model of engagement of people with dementia that could formalize the relationships among these assessment techniques and outline their relative meaning in the economy of the overall engagement state.

In order to pursue these two goals, the doctoral research was organized in three studies. First, we carried out an extensive *ethnographic study* to understand people with dementia in their context of living and get acquainted with the activities proposed by nursing homes. Second, we conducted an *exploratory study* to investigate the reactivity of people with dementia to an experimental setting and deploy a sensible research protocol for data collection. Third, we performed an *experimental study* and collected a database of multimodal data (e.g., video recordings, electrodermal activity signals, accelerometer signals) while people with dementia were involved in two types of activities: a game-based cognitive stimulation (i.e., jigsaw puzzles, shape puzzles, and a match with dominoes) and a robot-based free play (with the dinosaur robot Pleo).

As a first result, we came up with three techniques to measure different aspects of engagement in people with dementia: electrodermal activity (EDA), the Ethographic and Laban-Inspired Coding System of Engagement (ELICSE), and quantity of movement. EDA – which is the variation in the skin conductance derived from the activation of the sympathetic nervous system – accounts for the arousal of the person with dementia during the activity. The ELICSE – which is a coding system of engagement based on nonverbal behavior – permits the measurement of different body configurations that account for different levels of engagement. Quantity of movement – which is the amount of movement on the non-dominant wrist gauged with a triaxial accelerometer – captures the proactive engagement of the person with dementia during the activity (i.e., holding and manipulating objects, reaching out others).

As a second result, we built a model of engagement – the ENGAGE-DEM – which specifies the components of engagement, how these are measured through the deployed measurement techniques – EDA, the ELICSE, and quantity of movement – and which relationships they entertain. The model was tested with structural equation modeling using the data collected during the third study and achieved an excellent goodness of fit. The ENGAGE-DEM is the result of a process of testing and progressive refinement of a model of engagement drawn from the literature. This process brought to the refutation of the widely accepted definition of engagement as a compound of positive affect and proactive involvement and to the promotion of a more data-consistent definition of engagement. According to the ENGAGE-DEM, engagement is the degree of proactive participation of the person with dementia in an activity that can take different hedonic tones and achieve different levels of energy mobilization.

The ENGAGE-DEM could contribute to several domains of knowledge. It could benefit the field of nursing research since it could prompt a better understanding of the person with dementia and enable a more informed choice of meaningful activities. It could be an aid for designers aiming to create compelling playful technologies for people with dementia. Last, it could be used to enable socially interactive robots and interactive technologies to detect the engagement state of the person with dementia online and react accordingly.

Table of contents

Summary **xi**

List of figures **xx**

List of tables **xxii**

Acronyms **xxiii**

1 Introduction **1**

1.1 Dementia: a Global Health Priority 2
1.2 The Role of Meaningful Activities 2
1.3 The Complex Measurement of Engagement 3
1.4 Problem Statement 4
1.5 Research Objectives and Questions 4
1.6 Performed Studies 5
1.7 Research Outcomes 6
1.8 Research Contribution 7
1.9 Structure of the Thesis 8

2 State of the Art **11**

2.1 Introduction 12
2.2 Definitions of Engagement 12
2.2.1 Engagement with an Object 13
2.2.2 Engagement with an Agent 15
2.2.3 Co-Activity and Social Support 17

2.3 Components of Engagement 18
2.3.1 Attention 20
2.3.2 Affect 21
2.3.3 Time 21
2.3.4 Other Components 22
2.4 Measures of Engagement 22
2.4.1 Experiential/Subjective Measures 23
2.4.2 Peripheral-physiological Measures 25
2.4.3 Behavioral/Expressive Measures 27
2.5 Conclusions 33

3 Theoretical Framework 35
3.1 Introduction 36
3.2 Provisional Definition of Engagement for Dementia 36
3.3 Co-activities for People with Dementia 36
3.4 Components of Engagement in Dementia 37
3.5 Measures of Engagement for Dementia 38
3.6 Tentative Theory-driven Model of Engagement for Dementia 38
3.7 Conclusions 40

4 *Data Collection:* The Experimental Study 41
4.1 Introduction 42
4.2 Participants 42
4.3 Experimental Design 43
4.4 Experimental Activities 44
4.4.1 Game-based Cognitive Stimulation 45
4.4.2 Robot-based Free Play 46
4.5 Measurement Instruments 46
4.5.1 Observational Rating Scales 47
4.5.2 Video recordings 48
4.5.3 E4 wristband 48
4.6 Setting 49
4.7 Procedure 50
4.8 Ethical Approval 52
4.9 Conclusions 52

5 *Measure Development:* **Electrodermal Activity** 55
5.1 Introduction 56
5.2 Research Gap 56
5.3 Identification and Exclusion of Invalid Measurements 57
5.4 Synchronization and Labeling 58
5.5 Pre-processing and Feature Extraction 59
5.6 Concurrent Validity 61
5.7 Discussion 63
5.7.1 Research Questions 63
5.7.2 Dynamics of Psychophysiology 64
5.7.3 Selection of EDA Features 65
5.8 Conclusions 67

6 *Measure Development:* **A Coding System of Engagement-related Behavior** 69
6.1 Introduction 70
6.2 Methodological Issue 70
6.3 Construction of the Coding System 71
6.3.1 Note on the Video Recordings 71
6.3.2 Ethographic Approach: Development of the Ethograms 72
6.3.3 Laban-Inspired Approach: Structuring of the Ethograms 73
6.4 The Coding System 74
6.4.1 Advantages of the Coding System 76
6.5 Inter-rater Reliability 77
6.6 Concurrent Validity 79
6.7 Discussion 85
6.7.1 Research Questions 85
6.7.2 Behavioral Dynamics of Engagement 85
6.7.3 Adaptability of ELICSE 86
6.7.4 Meaning of Behaviors 87
6.8 Conclusions 88

7 *Construct Modeling:* **The Theory-driven Model of Engagement** 89
7.1 Introduction 90
7.2 Problem Framing and Model Contribution 90
7.3 Components and Measures of Engagement 92

	7.4	Model Testing	94
		7.4.1 Analysis of the Causes of Negative Error Variance	95
		7.4.2 Model Modification	97
	7.5	Model Re-testing	97
		7.5.1 Analysis of the Causes of Negative Error Variance	98
		7.5.2 Potential Sources of Misspecifation	99
	7.6	Discussion	100
	7.7	Conclusions	100
8	***Construct Modeling:* The Hierarchical Organization of Engagement-related Behavior**		**103**
	8.1	Introduction	104
	8.2	Patterns of Body Part Organization	104
		8.2.1 Patterns of Body Part Organization: Focused Attention	104
		8.2.2 Patterns of Body Part Organization: Valence	109
	8.3	Hypothetical Regression Paths	110
	8.4	Results of SEM	111
	8.5	Discussion	113
		8.5.1 Consequences for a Model of Engagement	113
		8.5.2 Consequences for the Measurement of Behavior	114
	8.6	Conclusions	116
9	***Measure Development:* Quantity of Movement**		**117**
	9.1	Introduction	118
	9.2	Measurement Rationale	118
	9.3	Identification and Exclusion of Invalid Measurements	119
	9.4	Feature Extraction	119
	9.5	Concurrent Validity	120
	9.6	Discussion	121
	9.7	Conclusions	123
10	***Construct Modeling*: The Final Model of Engagement**		**125**
	10.1	Introduction	126
	10.2	Stock of the Situation	126
	10.3	Final Model of Engagement	127
	10.4	Hypothesized Relations between Components	128

10.5 Model Testing 130
10.5.1 Model Testing with Single Activities 132
10.6 Discussion 134
10.6.1 Reflections on Engagement 135
10.7 Conclusions 137

11 Conclusions 139
11.1 Introduction 140
11.2 Research Questions 140
11.2.1 RQ1. How can engagement in dementia be measured? 140
11.2.2 RQ2. What are the dynamics of engagement in dementia? 143
11.3 Limitations and Future Work 144

Appendices 145
Appendix A. The Ethnographic Study 147
Appendix B. The Exploratory Study 150
Appendix C. Script for Facilitators 153
Appendix D. OME and OERS 158
Appendix E. The Coding System (ELICSE) 160
Appendix F. Further Analyses 164
Appendix G. SEM Glossary 171

Bibliography 172
Acknowledgements 184
Curriculum Vitae 188
Publications 189

List of figures

Figure	1	Design products for people with dementia	7
Figure	2	Flow channel	13
Figure	3	Circumplex model of affect	27
Figure	4	Tentative theory-driven model of engagement	39
Figure	5	Board games	45
Figure	6	Pleo, overview of sensors and degrees of freedom	46
Figure	7	E4 wristband, overview of sensors	49
Figure	8	Activity room layout	50
Figure	9	Phases of experimental session	51
Figure	10	Datasets, overview of missing data	53
Figure	11	Example of invalid EDA signal	58
Figure	12	Data synchronization in Matlab	59
Figure	13	Filtering of EDA signal	60
Figure	14	Chronology of an EDA session	66
Figure	15	Example of adjustment	72
Figure	16	Coding system in Observer XT	76
Figure	17	Scoring of ELICSE	77
Figure	18	Examples of body configurations	77
Figure	19	Boxplots of percentages of behavior and modifiers duration	80
Figure	20	Theory-driven model of engagement	93
Figure	21	Correlation between focused attention and rapport	96
Figure	22	Modified version of theory-driven model of engagement	98
Figure	23	Main pattern of body part organization of focused attention	106
Figure	24	Patterns of body part organization of focused attention and valence	107
Figure	25	Variations to main pattern of body part organization of focused attention	108

Figure 26	Regression paths between indicators of focused attention and valence	112
Figure 27	Plots of effects of motivational disorders on quantity of movement	122
Figure 28	Final model of engagement	129
Figure 29	Final model of engagement with path estimates (without outliers)	131
Figure 30	Path estimates for game-based cognitive stimulation	133
Figure 31	Path estimates for robot-based free play	133
Figure 32	Modification of final model of engagement for robot-based free play	135
Figure A1	Activities of residents	148
Figure C1	Positioning of E4 wristband	153
Figure C2	Positioning of E4 wristband	156
Figure F1	Effects of motivational disorders on items of OME and OERS	165
Figure F2	Effects of motivational disorders on behaviors of the ELICSE	167/168
Figure F3	Effects of motivational disorders on quantity of movement	169
Figure F4	Effects of motivational disorders on EDA	170

List of tables

Table	1	Overview of components of engagement in the literature	19
Table	2	Other components of engagement	23
Table	3	Observational rating scales of engagement for people with dementia	29
Table	4	Ethograms of engagement for people with dementia	30
Table	5	Coding schemes of engagement for people with dementia	31
Table	6	Overview of sessions and order of presentation of games	44
Table	7	Set of EDA features and corresponding equations	61
Table	8	Concurrent validity of EDA for game-based cognitive stimulation	62
Table	9	Concurrent validity of EDA for robot-based free play	62
Table	10	EDA features included in theory-driven model of engagement	65
Table	11	The coding system	75
Table	12	Inter-rater reliability results	78
Table	13	Concurrent validity of the behaviors in the ELICSE for game-based cognitive stimulation	83
Table	14	Concurrent validity of the modifiers in the ELICSE for game-based cognitive stimulation	83
Table	15	Concurrent validity of the behaviors in the ELICSE for robot-based free play	84
Table	16	Concurrent validity of the modifiers in the ELICSE for robot-based free play	84
Table	17	Assignment of behaviors to components of engagement	88
Table	18	Aggregation of modifiers in ELICSE	93
Table	19	Aggregation of behaviors in ELICSE	97
Table	20	Path Estimates of model	113
Table	21	Concurrent validity of quantity of movement for game-based cognitive stimulation and robot-based free play	121
Table	22	Factor loadings and Cronbach's α	130
Table	23	Path estimates of final model of engagement	132
Table	E1	Operational descriptions of head behaviors	160
Table	E2	Operational descriptions of torso behaviors	161
Table	E3	Operational descriptions of arms/hands behaviors	162

Acronyms

ADI	Alzheimer's Disease International
CFI	Comparative Fit Index
EDA	Electrodermal Activity
EEG	Electroencephalography
EFA	Exploratory Factorial Analysis
ELICSE	Ethographic and Laban-Inspired Coding System of Engagement
EMODEB	Evidence-based MODel of Engagement-related Behavior
EMG	Electromyography
ESM	Experience Sampling Method
FACS	Facial Action Coding System
fNIR	functional Near-Infrared Spectroscopy
GBCS	Game-Based Cognitive Stimulation
GEQ	Game Experience Questionnaire
HCI	Human-Computer Interaction
HHI	Human-Human Interaction
HR	Heart Rate
HRI	Human-Robot Interaction
HRV	Heart Rate Variability
IRR	Inter-Rater Reliability
KMO	Kaiser-Meyer-Olkin
LMA	Laban Movement Analysis
MCAR	Missing Completely At Random
MCI	Mild Cognitive Impairment
MEC	*Mini-Examen Cognoscitivo* (Spanish version of MMSE)
MMSE	Mini-Mental State Examination
MPES	Menorah Park Engagement Scale
NFI	Normed Fit Index
NPI-NH	Neuropsychiatric Inventory Nursing Home Version
OERS	Observed Emotion Rating Scale

OME	Observational Measurement of Engagement
QoL	Quality of Life
PNFI	Parsimony Normed Fit Index
PNS	Parasympathetic Nervous System
PPG	Photoplethysmogram
RBFP	Robot-Based Free Play
Reisberg GDS	Reisberg Global Deterioration Scale
RFI	Relative Fit Index
RMSEA	Root Mean Square Error of Approximation
SEM	Structural Equation Modelling
SIQ	Subjective Impression Questionnaire
SNS	Sympathetic Nervous System
UES	User Engagement Scale
VC-IOE	Video Coding Incorporating Observed Emotions
VIF	Variance Inflation Factor
WHO	World Health Organization

1

Introduction

1.1 Dementia: a Global Health Priority

The worldwide population is growing older. In 2015, the only country with the 30% of the population exceeding the 60 years of age was Japan. By 2050, this proportion will be reached by Europe, North America, Chile, China, the Islamic Republic of Iran, the Republic of Korea, the Russian Federation, Thailand, and Vietnam (World Health Organization, 2015). As age is a risk factor for several diseases, the unprecedented growth of age of the worldwide population will have an enormous impact on health. It will trigger an increase in the prevalence rates of age-related illnesses and, consequently, inflate health care costs. In this sense, one of the major health challenges regards dementia (World Health Organization and Alzheimer's disease International, 2012). Dementia is an umbrella term for a set of neurodegenerative disorders (e.g., vascular dementia, Alzheimer's disease, dementia with Lewy bodies) that cause a reduction in *cognition* (e.g., impairment of thinking, problem solving, and mnemonic ability), *functioning* (e.g., inability to carry out activities of daily living, such as self-feeding and getting dressed), and *behavior* (e.g., disorders of perception, thought content, mood, and behavior, such as depression, agitation, and apathy). If in 2015, the number of people living with dementia amounted to 46.8 million, this number is expected to reach 74.7 million by 2030 and peak to 131.5 million by 2050 due to global aging (Prince et al., 2015). The cost of dementia – which in 2010 was estimated in USD 604 billion – is going to reach USD 1.2 trillion in 2030, risking to demise social and economic growth (World Health Organization, 2015).

1.2 The Role of Meaningful Activities

As dementia is a global health priority, several countries have already put forward national dementia strategies. These are focused on promoting early diagnosis, developing a range of services to properly address care needs, and improving quality of life (QoL). Sube Banerjee and Jenny Owen, in the preface of 2009 UK National dementia strategy (UK Department of Health, 2009), underlined that, albeit dementia is a set of incurable diseases that cause a substantial burden in the patient, there is still room to "*live well with dementia*" and the key ingredient to cope with the disease is a satisfactory QoL.

In the last two decades, several studies have focused on examining the drivers of QoL in dementia. Most of these studies found out that cognition and functioning have no significant effect on QoL and that it is perfectly possible to have a good QoL at any stage of dementia (Logsdon et al., 1999; Banerjee et al., 2006; Beerens et al., 2013). Apparently, what affects QoL the most is the presence of behavioral disturbances (Logsdon et al., 1999; Beerens et al., 2013) – especially agitation (Cohen-Mansfield, 2001; Cohen-Mansfield and Werner, 1997; Scherder et al., 2010) and depression (Logsdon et al., 1999; Beerens et al., 2013) – and the lack of exposure to social interactions (Moyle et al., 2011) and meaningful activities (Bryant et al., 2001; Ritchey et al., 2001; Lysack and Seipke, 2002; McIntyre and Howie, 2002; Silverstein and Parker, 2002; Phinney et al., 2007; Moyle et al., 2011). Qualitative studies like those of Phinney et al. (2007) and Bryant et al. (2001) reported that, in the view of people with dementia, it is through *doing* that a person finds life to be meaningful and that healthy aging is the process of "*going and doing something meaningful*".

1.3 The Complex Measurement of Engagement

In this perspective, the promotion of meaningful activities – both inside and outside the nursing home – assumes a crucial meaning. Also, the measurement of engagement plays an important role. Indeed, one thing is offering a wide range of playful activities, another is understanding whether these activities are really meaningful for the person with dementia. For this latter objective, a framework to measure engagement in dementia is essential. According to the literature, engagement is the psychological state of proactive involvement with an object (e.g., a game, an interactive system) or an agent (e.g., a person, a social robot) that has a positive affective nuance.

Engagement can be measured on three levels, using three different response systems or modalities: *experiential/subjective* (e.g., self-reports), *behavioral/expressive* (i.e., overt verbal and non-verbal behavior), and *peripheral-physiological* (e.g., electrodermal activity, heart rate, heart rate variability, inter-beat interval; Izard, 1971, 1991; Izard et al., 1972; Tomkins, 1984; Mauss et al., 2005; Eifert and Wilson, 1991; Bradley and Lang, 2000). In dementia, the experiential/subjective level of measurement is often inaccessible due to cognitive impairment. Indeed, reporting one's own emotional and psychological states requires a combination of introspection, concept classification, and mnemonic ability, all cognitive faculties that get gradually lost with the progression of the disease. With regards to the behavioral/expressive level, this is sometimes impaired, too. Behavioral disturbances like apathy and depression – which have high prevalence rates in dementia (apathy: 55.5%; depression: 44.9%; Robert et al., 2005) – may blunt the expression of emotions at a behavioral level. Concerning the peripheral-physiological level, to date it has been only rarely studied in dementia and, to the best of our knowledge, never collected in the field.

1.4 Problem Statement

In order to properly and exhaustively measure engagement in dementia, it is decisive to combine measures coming from different response systems. In this respect, the literature is far from being conclusive. There are still some open research problems in need of further exploration by the research community, which are listed below.

1. Self-reports are only rarely used as tools to measure engagement in dementia (Wada et al., 2005). Thus, it is hard to make inferences regarding their feasibility as a reliable form of assessment of engagement.

2. Assessment techniques of engagement in dementia rely exclusively on behavior observation (Lawton et al., 1996; Cohen-Mansfield et al., 2009 and 2011; Jones et al., 2015), an approach that is risky as dementia is often accompanied by behavioral disturbances (e.g., apathy, depression, agitation).

3. State of the art behavioral assessment techniques are either too general to grasp the natural flow of engagement (i.e., observational rating scales) or too complex to trace behavior back to an overall engagement state (i.e., ethograms).

4. Available studies on the physiology of engagement of people with dementia are scarce and involve costly and invasive procedures, such as electroencephalography (EEG), urinalysis, hormone analysis, and functional Near-Infrared Spectroscopy (fNIRS; Wada et al., 2005; Wada and Shibata, 2008; Kawaguchi et al., 2012).

5. No study combines different measures of engagement. Consequently, the picture of engagement in dementia that the literature draws is incomplete.

1.5 Research Objectives and Questions

The present thesis has a twofold objective. First, it aims at exploring and developing novel techniques to accurately assess engagement in people with dementia. Second, it aims at creating a model of engagement that formalizes the relationships among these assessment techniques and outlines their relative meaning in the economy of the engagement state.

The research objectives of the present thesis are also described in terms of research questions (RQ). RQ1 refers to the first research objective (i.e., explore new assessment techniques of engagement), while RQ2 refers to the second research objective (i.e., develop a model of engagement that formalizes the relationships between new assessment techniques of engagement).

RQ1. How can engagement be measured in people with dementia?

a. Can self-reports be employed as a gold standard of engagement in dementia? If not, which validated measures of engagement can be used as a gold standard?

b. Which peripheral-physiological measures can be employed to assess engagement in people with dementia? Are these peripheral-physiological measures appropriate to assess engagement?

c. Which behaviors externalize engagement in people with dementia? Are these behaviors appropriate to measure engagement?

RQ2. What are the dynamics of engagement in dementia?

a. What are the relationships between the different components of engagement? Which conclusions can be drawn from these relationships regarding the functioning of engagement in dementia?

1.6 Performed Studies

In order to reply to these research questions, the present doctoral work was organized in three studies. First, we carried out an extensive *ethnographic study* (duration: one month) to familiarize with the context of the nursing home and determine the requirements for the collection of multimodal data. Second, we conducted an *exploratory study* (duration: one month and a half) to deploy a sensible research protocol for the collection of multimodal data, investigate the viability of self-reports as a gold standard for the assessment of engagement, and establish the correct set-up for data collection and the correct temporal division of the experimental sessions. Third, we carried out an *experimental study* (duration: eight months, five of which of data collection) to collect a multimodal database (e.g., videos and psychophysiological data) while people with dementia were involved in two types of activities: a game-based cognitive stimulation (i.e., jigsaw puzzles, shape puzzles, and a match with dominoes) and a robot-based free play (with the dinosaur robot Pleo). This database served to develop and assess the concurrent validity of the newly deployed measures and build and test the model of engagement for dementia.

As the experimental study was central to the development of the different measures of engagement and of the final model, it is given prominence in the thesis. The ethnographic and exploratory studies are presented in the appendixes. They did contribute to the pursuit of the research objectives. However, they did so at a methodological and practical level, as they served to identify the right protocol for data collection given the shortage of examples provided by the literature.

1.7 Research Outcomes

Within this dissertation, we show the development of three measures of engagement for people with dementia: electrodermal activity (EDA), the Ethographic and Laban-Inspired Coding System of Engagement (ELICSE), and quantity of movement. Some of these measures – the ELICSE and quantity of movement – were developed *ex novo* in the context of this research. Others – EDA – were just explored for the first time in this particular user group.

1. EDA – the variation in the skin conductance derived from the activation of the sympathetic nervous system (Dawson et al., 1990) – accounts for the arousal of the person with dementia during the activity (i.e., degree of energy mobilization).

2. The ELICSE – a coding system based on nonverbal behavior (i.e., gaze, postures, arms/hands behaviors) – permits the measurement of the engagement in the activity through different behavioral modalities and the association of different body configurations to different levels of engagement.

3. Quantity of movement – the amount of movement on the non-dominant wrist gauged with a triaxial accelerometer – captures the proactive participation of the person with dementia during the activity (i.e., holding and manipulating objects, reaching out others).

These novel assessment techniques converged into the ENGAGE-DEM. The ENGAGE-DEM is a model that formalizes the different components of engagement, how these are measured with the tools deployed along the dissertation (i.e., EDA, the ELICSE and quantity of movement), and how they relate to each other. Such a model is the result of a *process* of progressive refinement of a model of engagement drawn from the review of the literature.

1.8 Research Contribution

The ENGAGE-DEM might benefit several domains of knowledge, namely nursing research, social robotics, and user experience design. In the context of *nursing research*, it could be employed by practitioners to better detect engagement in people with dementia and identify meaningful activities. Moreover, it could be used to investigate whether engagement is a mediating variable interposed between the participation in playful activities and their clinical benefit (e.g., improvement of cognitive functioning, reduction of challenging behaviors, improvement of QoL). In the framework of *social robotics*, it could be exploited for affect-based co-adaptation – the bi-directional process of real-time synchronization and adaptation of social robots to the affective states of humans (Gao et al., 2017). The model can be used to enable social robots to recognize the engagement state of the person with dementia online and adapt their interaction accordingly. This approach has been attempted with healthy participants (Castellano et al., 2009a; Castellano et al., 2009b; Rich et al., 2010; Sanghvi et al., 2011; Salam et al., 2017; Gao et al., 2017) and recently with children with autism (Rudovic et al., 2018), but, to the best of our knowledge,

Figure 1. Design products for people with dementia. Top left: Tactile dialogues; top right: Tovertafel; bottom left: Closer to nature; bottom right: Dynamorph

not with persons with dementia. With regards to *user experience design,* the model can be used to assess the user experience of people with dementia with interactive entertainment technologies. In the last decade, several designers have focused on the development of interactive technologies for people with dementia (see Figure 1). Some examples are *Tactile dialogues* (Schelle et al., 2015) – a textile pillow provided with a haptic feedback aimed at stimulating interpersonal contact, *Tovertafel* (Anderiesen, 2017) – a set of six playful interactive animations to be played on any dining table aimed at reducing apathy, *Dynamorph* (Feng et al., 2017) – a living table activated by mean of four zoomorphic robot-like shapes aimed at reducing boredom and loneliness, and *Closer to nature* (Feng et al., 2018) – an interactive installation that projects a farm on a 87-inch-display aimed at connecting residents with the outdoors. The ENGAGE-DEM can provide designers with a measurement framework to assess the quality of user experience in people with dementia.

1.9 Structure of the Thesis

This dissertation is divided into eleven chapters.

- This chapter introduced the problem investigated in the dissertation – the measurement of engagement in dementia – and outlined its importance and contribution to the fields of nursing research, social robotics, and user experience design.

- Chapter 2 provides a comprehensive and critical review of the state of the art regarding engagement. It does so in three steps: definitions of engagement, components of engagement, and measures of engagement.

- Chapter 3 narrows down the state of the art to the context of dementia and builds up the theoretical framework of this dissertation. It provides a tentative definition of engagement drawn from the literature, identifies the components of engagement measurable in people with dementia, and determines which measures can be used to gauge engagement in people with dementia among those identified in the literature. Finally, it summarizes all this knowledge in a tentative theory-driven model of engagement.

- Chapter 4 illustrates the experimental study that served for data collection: participants, design, measures, setting, and procedure of data collection.

- Chapter 5 describes the first of the three measures deployed to assess engagement, EDA. It presents the process of synchronization, data filtering, and feature extraction of the signal and discusses the adequacy of EDA as a measure of engagement.

- Chapter 6 presents the second of the three measures of engagement developed along the dissertation, the ELICSE. It portrays the process of development of the coding system according to the guidelines of Ethology and Laban Movement Analysis (LMA), discusses the results of inter-rater reliability, and examines the adequacy of the ELICSE as a measure of engagement.

- Chapter 7 finalizes the tentative theory-driven model of engagement sketched in Chapter 3 by detailing its final metrics and employs structural equation modeling (SEM) to test its validity. As the theory-driven model of engagement proves to be invalid, a number of potential causes of misspecification are provided and discussed at the end of chapter 7.

- Chapter 8 follows up on the preceding chapter by examining and describing one of the potential causes of misspecification of the theory-driven model of engagement, the hierarchical ordering of the behaviors in the ELICSE.

- Chapter 9 focuses on the development of the last of the three measures of engagement, quantity of movement[1]. It describes the treatment of the triaxial accelerometer signal, the feature extraction process, and discusses the adequacy of quantity of movement as a measure of engagement.

- Chapter 10 presents the final model of engagement, the ENGAGE-DEM – a data-driven modification of the theory-driven model of engagement. It describes the assumptions on which the model is based, its construction, and testing.

- Chapter 11 reports the concluding remarks, the limitations, and future applications of the present research work.

1 This measure comes in late in the dissertation, as it was not suggested by the literature on engagement, but by the data collection and analysis.

2

State of the Art

2.1 Introduction

This chapter revises the state of the art regarding engagement in three steps: (i) definitions of engagement, (ii) components of engagement, and (iii) measures of engagement. In this respect, it does not limit itself to studies regarding engagement in people with dementia. On the contrary, it encompasses research coming from different scientific fields focusing on a multitude of user types. This decision is motivated by three facts: (1) the literature on engagement in dementia is scarce and mainly confined in the work of few authors, (2) engagement as a construct is often victim of an over-simplification when studied in the context of dementia, and (3) the complexity of engagement can be captured exclusively by adopting different angles of view.

The present chapter aims at providing the reader with a thorough idea of how engagement is defined beyond scientific boundaries, what its main components are, and which tools have been developed to measure it at the levels experiential/subjective, behavioral/expressive, and peripheral-physiological.

2.2 Definitions of Engagement

At present, there is no consensus on the definition of engagement. The literature is filled with attempts and with partially overlapping notions that are called with different names: engagement, enjoyment, immersion, engrossment, flow. The interest surrounding engagement in the recent years is mostly the result of a shift of focus in the fields of human-computer and human-robot interaction (HCI and HRI). In the former context, researchers have started evaluating the design of technologies beyond the usual functional and performance-based concepts of usability, efficiency, and effectiveness and got progressively involved with the "*non-utilitarian aspects of the interaction*", the user experience (Hassenzahl, 2003). In the latter context, researchers have grown an interest in understanding the connection that humans create with artificial agents – especially with social robots – and using this knowledge to make these agents always more spontaneous, empathetic, affectionate, and compliant with social rules. The birth of engagement as a research topic, however, is not to be ascribed to HCI and HRI, but rather to Positive Psychology, the "*science of positive subjective experience, positive individual traits, and positive institutions*" (Seligman and Csikszentmihalyi, 2000). It is no coincidence that the most widely accepted and eminent framework of engagement – the *flow theory* – was produced by one of the most renowned representatives of Positive Psychology, Mihaly Csikszentmihalyi.

The review of the literature that we performed restituted a double definition of engagement. On the one side, engagement is described as the involvement with an *object*: a stimulus, a product, a book, a game, an interactive system, a physical activity. On the other side, it is presented as the involvement with an *agent*: a social robot, a person. We decided to report both connotations of engagement, as they can and in fact do co-exist.

2.2.1 Engagement with an Object

As anticipated, one of the most prominent definitions of engagement in the literature is Csikszentmihalyi's definition of flow (Csikszentmihalyi, 1989; Csikszentmihalyi, 2014). As a word, *flow* was chosen out of consensus, it is the way most interviewees described the subjective experience of being engaged in autotelic activities – activities rewarding for themselves and not for the results they generate (Csikszentmihalyi, 1975; Getzels and Csikszentmihalyi, 1976; Csikszentmihalyi; 2000). Running, painting, dancing, rock-climbing are some of the activities enlisted as evoking flow. All these activities entail an intense and focused concentration, the union of awareness and action, a sense of control of one's own actions, the loss of self-consciousness, the distortion of the temporal axis, and the perception of intrinsic reward. Central to the notion of flow is the fragile balance between challenges and skills. When a person is in flow, the activity is just-manageable. The imbalance between challenges and skills can either lead to *anxiety* – when challenges exceed skills – or to apathy and *boredom* – when skills exceed challenges (see Figure 2).

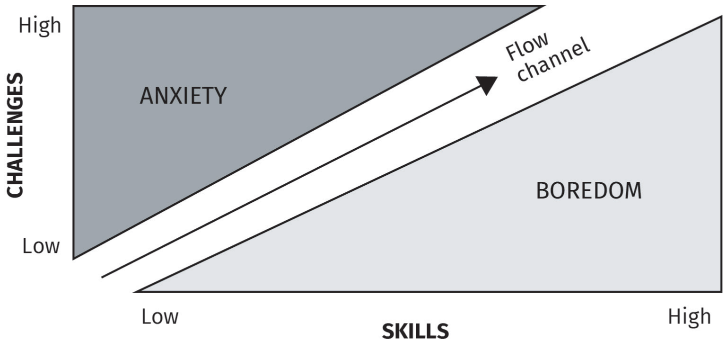


Figure 2. Flow channel (adapted from Schell, 2014)

Brown and Cairns (2004) use the term *immersion* to describe a concept similar to flow, but not exactly overlapping, the experience of getting lost in a game and being out of contact with reality. According to the authors, immersion is closely related to the concept of *presence*, "*the extent to which a person's cognitive and perceptual systems are tricked into believing they are somewhere other than their physical location*" (Lombard and Ditton, 1997; Slater et al., 2009). Immersion has three level of intensity: (i) *engagement*, (ii) *engrossment*, and (iii) *total immersion*. When engaged, gamers invest their time, effort, and attention in the game. When engrossed, their emotions are directly affected by the game. When totally immersed, they are cut off from reality, all that matters is the game.

Douglas and Hargadon (2000) also describe immersion, but in novels. In contrast with Brown and Cairns, who situated engagement in the continuum of intensity of immersion, they differentiate engagement from immersion. Douglas and Hargadon declare: "*immersive novels and films require virtually no engagement from their readers and viewers, since we can simply follow the plot and enjoy the ride*". What this statement suggests is that engagement has to do with the proactive effort of the reader in the activity of reading, while immersion just entails the act of being passively carried away by the novel. "*Reading Jan Eyre is immersive*", while "*reading Ulysses is engaging*". *Jamming* or flow are yet other concepts to Douglas and Hargadon. They occur when the reader, in its proactivity, is performing both incredibly well and effortlessly. Once again, the balance challenge/skills is key to flow.

More design-oriented definitions of engagement focus on the aesthetic qualities of a product or interactive system (Jacques et al., 1995; Overbeeke et al., 2003; O'Brien and Toms, 2008; Attfield et al., 2011). They situate the state of engagement in the *system*, as well as in the *user*. It is the capacity of the system to attract the attention of the user with its aesthetic qualities and the experience of the user that comes with it. To make an example, Chapman et al. (1999) define engaging systems as *enticing* for the users to the extent that they *draw* the user into the activity, *seduce* and *spur* him/her, *capture* and *captivate* his/her interest.

In contrast with these views, engagement in dementia is described in a far more basilar way. Cohen-Mansfield et al. (2011) define it as the *"act of being occupied or involved with an external stimulus*", and, by extension, as "*the antithesis of apathy*". Judge et al. (2000) present it as the motor or verbal behavior exhibited in response to the activity. Compared to the conception of engagement for healthy persons, the definition of engagement in dementia is deprived of any positive connotation and of any subjective quality.

In conclusion, every reviewed definition of engagement includes the following three elements: (i) a *person* – the user, the reader, the player, (ii) an *object* – a game, a physical activity, a book, a product, an interactive system, and (iii) an *interaction* between the person and the object. In plain words, engagement is described by the literature as the *positive quality of the interaction* of a person with an object. This quality is influenced by the attributes of the person, by the characteristics of the object, but also by the context where the interaction takes place (Cohen-Mansfield et al., 2011).

If the object is a *book*, the characteristics of the object influencing the quality of the interaction with the user can be the plot, the characters, the quality of the writing, the narrative style. If the object is a *computer game*, the characteristics of the object influencing the quality of the interaction with the user can be the game design elements: the game space (e.g., narrow pathways, maze), the game components (e.g., enemies, points, game levels), the game mechanics (e.g. jumping, collecting), the game goals, and the game rules (Deterding et al., 2011). If the object is an *activity*, the characteristics influencing the quality of the interaction with the user can be the goal of the activity, the tools involved in it, and the persons participating to the activity (Carvalho et al., 2015).

Another element characterizing any definition of engagement is the voluntariness of the interaction. Engagement is unanimously described as intrinsically motivated. As a side note, it is noticeable that almost all the reviewed definitions of engagement with an object – except those for dementia – describe it as the *subjective experience* of the quality of interaction, posing a lot of attention on how the interaction is subjectively perceived and felt and neglecting its observable features.

2.2.2 Engagement with an Agent

In this section of the state of the art, we present definitions of engagement that regard human and artificial life-like agents, such as social robots. The decision to include social robots in the category agents instead of objects was dictated by the fact that social robots – as opposed to interactive systems - are endowed with the quality of *social agency*. They adhere to social norms and prompt social interactions comparable to those described by human-human (HHI) and human-animal interaction and, in fact, comply with models of interaction entirely distinct from those of HCI.

In general, the definition of social engagement (or interaction) is rather settled and refined in the context of social sciences. One of the most eminent formalizations of social engagement between humans is to be ascribed to Tickle-Degnen and Rosenthal (1990). They used the term *rapport* to describe the dynamic structure of three interrelating components: mutual attentiveness, positivity, and coordination. The authors claim that when people experience a high degree of rapport, they are other-involved, they form a cohesiveness with each other through the expression of mutual attention. Also, they feel a mutual sense of friendliness and caring (i.e., positivity) and are fine-tuned with each other to the extent that they react simultaneously, sympathetically, and sometimes in a synchronized way (i.e., coordination). As an example, when people experience a high level of rapport – when they click – they might find themselves mirroring each other’s postures, gestures (La France and Ickes, 1981), and physiological states (Levenson and Gottman, 1983; Järvelä et al., 2014).

Opposed to HHI, social HRI has not yet developed a clear and globally accepted definition of engagement. Most of the studies regarding engagement with a social robot abstain from presenting a definition of the construct and rely on the reader’s common sense to fill the void. Few of the definitions that can be retrieved in the literature are those of Sidner et al. (2005), Rich et al. (2010), Castellano et al. (2009a and 2009b), and Díaz-Boladeras (2017). They utilize notions of social engagement borrowed from social sciences and use terms such as individual and participant to refer both to human and robotic interactors. To them, social engagement is:

- The process by which *individuals* in an interaction start, maintain, and end their perceived connection to one another (Sidner et al., 2005; Rich et al., 2010).

- The value that a *participant* in an interaction attributes to the goal of being together with the other participant(s) and continuing the interaction (Castellano et al., 2009a and 2009b).

- The observable component of bonding (e.g., time spent, joint activity, attention, proximity, absorption) and its behavior-inferred or self-reported emotional correlate (i.e., the feeling of closeness; Díaz-Boladeras, 2017).

Due to the lack of a formal definition of social engagement in the context of social HRI, engagement is often confused with attention. As Castellano et al. (2009b) correctly report, social engagement comprises an affective component on top of the attentional one. She and colleagues also note that engagement implies *proactivity*, as to say "*a high action tendency*" of the user.

With regards to the definition of social engagement in the context of dementia, the literature is quite limited. Cohen-Mansfield et al. (2010b) promoted the same definition of engagement both for the engagement with an object and for the engagement with an agent. In their view, sociality is an *attribute* of the stimulus. On the opposite, Jones et al. (2015) define social engagement (with a social robot) as a social connection/interaction, posing the attention on the process of interaction, rather than on the characteristics of the stimulus. They extend the work of Cohen-Mansfield et al. to include an affective connotation of social engagement. In spite of the great interest towards promoting engagement in dementia – especially of a social type and especially with social robots (Pino et al., 2015) – definitions that describe the complexity of the phenomenon are still scarce.

In conclusion, the reviewed definitions of social engagement include three components: (i) a *person*, (ii) an *agent* – human or artificial, and (iii) an *interaction* between the person and the agent. Social engagement is described by the literature as the *positive quality of the social interaction* between a person and an agent – being the agent artificial or human. The quality of the social interaction is influenced by the attributes of the person, the characteristics of the agent, and the context where the interaction takes place (e.g., school, hospital, nursing home).

If the agent is a *person*, the attributes influencing the quality of interaction can be the length of the relationship (early encounters or later interactions), the character of the person, the aesthetics of the person (Dion et al., 1972), the degree of mutuality of his/her behavior. If the agent is a *social robot*, the characteristics influencing the quality of interaction can be the term of exposure to the interaction (i.e., short-term vs long-term), the social cues that the robot displays (Ghazali et al., 2017a and 2018a), the language it uses (Ghazali et al., 2017a and 2017b), its appearance (Wu et al., 2012; Wrobel et al., 2013; Ghazali et al., 2018b), the reciprocity of its behavior (Leite et al., 2014), and its contingent responsiveness (Díaz-Boladeras, 2017).

According to this analysis, engagement with an object and engagement with an agent are structurally similar. However, while in the type of interaction prospected by the latter, reciprocity is key – meaning that the agent and the person in the interaction are bound to respond to each other's behaviors – within the type of interaction prospected by the former, reciprocity is optional. It can appear in the form of a feedback in an interactive system, but can as well be absent, as when reading a book.

Another difference between engagement with an agent and engagement with an object that we need to register regards the relationship between challenges and skills. Albeit for some users social interaction is particularly demanding (e.g., persons with autism or with dementia), in most cases it does not entail a substantial level of complexity. As a side note, we would like to point to the fact that – unlike engagement with an object – most of the definitions of social engagement are focused on the *external representation* of the quality of interaction and not on its subjective experience.

From this review of the definitions of engagement, it becomes evident the lack of a proper definition of engagement – both with an object and with an agent – in the context of dementia. The literature on the engagement of people with dementia should be enriched with the complexity and depth that the construct of engagement has when it regards healthy subjects. Hence, one of the first requirements of this doctoral work is to *deploy a definition of engagement capturing its true connotation*.

2.2.3 Co–Activity and Social Support

As anticipated, object- and agent-directed engagement can co-exist. Most of the playful activities with which we get involved include the presence of others. Brandtzæg et al. (2003) introduce the concept of *co-activity*, an activity that implies a collective action. In co-activities, users do not engage with an object/agent alone. Instead, they do so with one or more social partners. Multi-player video games, running groups, study groups, robot-mediated interaction exist because we, as humans, find it more rewarding to pursue activities together with others rather than in isolation (Zajonc, 1965).

Social affiliation has a boosting effect on engagement also in another sense, it provides social support. Karasek and Theorell (1990) define *social support* in the work environment as "*the overall level of helpful social interaction available on the job from co-workers and supervisors*" and claim that social support can moderate the negative impact of high job strain situations on health. Vygotsky (1987) presents social support (or scaffolding) as the most efficient way to learn. The zone of proximal development is in fact the difference between what a learner can do without help and what s/he can do with the help of a teacher or more competent peer. Put it in engagement terms, social support can enhance the abilities (or skills) of a person to the extent that s/he can achieve far more compelling objectives (challenges).

2.3 Components of Engagement

The take-away message from the previous section is that engagement is the positive quality of the interaction between a person and an object or agent, which can be influenced by the person attributes, by the object/agent characteristics, as well as by the environment where the interaction takes place.

Engagement can be captured from two perspectives: subjective and objective. These can be represented as an iceberg. The *subjective* facet of engagement can be grasped through the first person report of a subjective experience, thus it is the portion of the iceberg located under water. Without a person able to report his/her experiences, it remains inaccessible. The *objective* facet of engagement can be recorded through the observation of its visible elements, hence it is the portion of the iceberg situated above water. It is accessible even without the report of the person experiencing the engagement state.

In this section of the chapter, we identify the building blocks of engagement. To do so, we make reference to frameworks that refer to one or both of the following elements: person attributes that influence engagement, components of the subjective experience of engagement, observable components of engagement, and object/agent characteristics that affect engagement (see Table 1). Since we have identified a common structure between engagement with an object and engagement with an agent, in this section, we treat them together.

Table 1. Overview of components of engagement in the literature. In regular, the papers regarding engagement with an object. In italics, the papers regarding engagement with an agent

	PERSON ATTRIBUTES		**ENGAGEMENT SUBJECTIVE EXPERIENCE COMPONENTS**			**ENGAGEMENT OBSERVABLE COMPONENTS**			**OBJECT / AGENT CHARACTERISTICS**			
	Intrinsic Motivation/ Curiosity /Interest / Proactivity	Past and Present Preferences	Time Perception / Endurability	Skills / Challenges Control / Richness	Loss of Self and External Consciousness	Attention	Affect	Time Investment / Duration	Aesthetics	Novelty	Reputation / Trust	Feedback / Reciprocity / Interactivity
Chapman et al., 1999	X			X		X	X					
Brown and Cairns, 2004			X	X	X	X	X	X				
O'Brien and Toms, 2008	X		X	X	X	X	X		X	X		X
Cohen-Mansfield et al., 2009						X	X	X				
Cohen-Mansfield et al., 2010a and 2010b		X				X	X		X			
Attfield et al., 2011	X		X	X		X	X		X	X	X	
Nakamura & Csikszentmihalyi, 2014	X		X	X	X	X						
Tickle-Degnen & Rosenthal, 1990						X	X					X
Sidner et al., 2005						X						X
Castellano et al., 2009b	X					X	X					
Rich et al., 2010						X						X

Chapter 2

2.3.1 Attention

Attention is unanimously recognized as the basic component of engagement (see Table 1). Indeed, it appears in all the reviewed frameworks. Nakamura and Csikszentmihalyi (2009) state that attention plays a key role in entering and staying in flow, as it shapes a person's experience. "*What to pay attention to, how intensely and for how long, are choices that will determine the content of consciousness, and therefore the experiential information that is available to the organism*" (Csikszentmihalyi, 1978). During flow, attention is allocated to a limited stimulus field, thus it becomes *focused*, and transforms itself into concentration. "*Concentration is this intensely focused attention on a narrow range of stimuli. It is a prerequisite for (…) working at the peak of one's capacity*" (Csikszentmihalyi, 2014). Attfield et al. (2011) agree with Csikszentmihalyi. According to them, engagement in an experience involves focusing attention to the exclusion of other things. To them, it is specifically this focused attention that leads to the loss of time perception during heightened engagement. Interestingly, Bakker et al. (2010) describe a different form of attention, opposite to the focused one, *peripheral attention*. They describe attention as "the division of resources over potential activities" and define a *center* of the attention – the activity to which most of the resources are allocated – and a *periphery* – the activities that are not in the center regardless of the number of resources allocated to them. In this sense, focused attention can be seen as the exclusive allocation of resources to the *center* to the detriment of the *periphery*.

O'Brien and Toms (2008) describe four steps of engagement all modulated by attention: (i) *point of engagement*, (ii) *period of engagement*, (iii) *disengagement*, and (iv) *re-engagement*. To them, the process of engagement starts when the user's attention is drawn by the aesthetic and information composition of the interactive system (i.e., point of engagement). Then, it continues for the period of time in which the user maintains his/her attention and interest towards the interactive system (i.e., period of engagement). Finally, it ends when the user makes an internal decision to stop the activity and directs his/her attention somewhere else (i.e., disengagement). At this point, engagement can be re-initiated and start all over again (i.e., re-engagement). Chapman et al. (1999) describe a similar process to that of O'Brien and Toms by borrowing the definition of engagement of Jacques (1995) which states that engagement is the positive interactive state where the attention of the user is willingly given and held.

Also in the context of dementia, attention is considered one of the main dimensions of engagement (Cohen-Mansfield et al., 2009). However, it is more an observational concept, it is the amount of focus that the person with dementia pays to a stimulus in terms of gaze allocation, manipulation of the stimulus, and verbal behavior regarding the stimulus.

Attention can be unidirectional or mutual. In the context of engagement with an object, attention is usually *unidirectional* and exerted by the person on the object. However, when engagement becomes social, attention might become mutual and bi-directional. It can be exerted by all participants in the interaction – being those persons and/or artificial agents. This is the case of the attention described by Tickle-Degnen et al. (1990), Sidner et al. (2005), and Rich et al. (2010). As unidirectional attention, also mutual attentiveness occurs in the form of a process with a phase of establishment, maintenance, and end (Sidner et al., 2005; Rich et al., 2010).

2.3.2 Affect

In Csikszentmihalyi, affect is not a necessary condition of flow. However, most of the frameworks that we reviewed feature affect – of a positive nature – as a crucial dimension of engagement. Attfield et al. (2011) and Chapman et al. (1999) affirm that engaged users are affectively involved. Castellano et al. (2009b), Peters et al. (2008), and Peters et al. (2009) regard engagement as a compound of attention and affect. Douglas and Hargadon (2000) talk about a *pleasure principle* in the title of their paper. Brown and Cairns (2004) present emotional involvement as the differentiating element discriminating between engagement and engrossment. O'Brien and Toms (2008) describe an emotional thread built in the state of engagement. Last, Tickle-Degnen and Rosenthal (1990) feature positivity as a component of rapport.

With regards to dementia, Cohen-Mansfield et al.'s framework of engagement (2009) also features an affective component, the *attitude* towards the stimulus. This refers to the amount of excitement/expressiveness towards the stimulus that the person with dementia displays. Interestingly, the frameworks of engagement for dementia – as opposed to those for healthy subjects – feature the whole range of affective states and not just positive ones. For instance, Jones et al. (2015) incorporate Lawton's Observed Emotion Rating Scale (OERS) in their Video Coding - Incorporating Observed Emotion (VC-IOE) protocol. The OERS features both negatively and positively-valenced affective states (i.e., pleasure, anger, anxiety/fear, sadness, and general alertness).

2.3.3 Time

Time is a clear component of engagement. Across frameworks, it appears to have a twofold connotation. On the one hand, there is the subjective perception of time. On the other hand, there is its objective passage. When speaking about the former conception of time, researchers have described the state of engagement as characterized by a progressive spatiotemporal detachment that causes a distortion of the time perception during the experience of engagement (Brown and Cairns, 2004; O'Brien and Toms, 2008; Nakamura and Csikszentmihalyi, 2009). When talking about the latter conception of time, a number of frameworks have paid attention to the temporal progression of engagement and identified its different phases (Sidner et al., 2005; O'Brien and Toms, 2008; Rich et al., 2010). Time has also a role on the *endurability* and retention of the experience of engagement. O'Brien and Toms (2008), O'Brien (2010), and Attfield et al. (2011) describe endurability as one of the components of engagement. There are two sides of endurability (Read et al., 2002). One is the likelihood of remembering an experience and is called *remembrance*. The other is the willingness to repeat that experience and is called *returnance*. Positive engaging experiences are more likely remembered and repeated by the user. This principle is called the *Pollyanna principle*.

2.3.4 Other Components

In the previous section, we mostly focused on the observable components of engagement. In this section, we describe person attributes influencing engagement, components of the subjective experience of engagement, and object/agent characteristics affecting engagement. As some of these concepts have already been partially treated in the chapter, we summarize them in Table 2.

In light of the different views on the composition of engagement, one of the requirements that a measurement framework on engagement should satisfy is to *identify the components of engagement and their boundaries*. In the context of dementia, this becomes especially difficult as the researcher needs to differentiate among those components of engagement that are actually accessible and those that cannot be reached due to the action of the disease.

2.4 Measures of Engagement

In this section of the chapter, we discuss the available measures of engagement. As anticipated in the introduction of the thesis, engagement can be measured on three different levels, according to three response systems: experiential/subjective, behavioral/expressive, and peripheral-physiological.

We first enlist some of the available self-report measures of engagement, then we describe how engagement can be measured through physiology. Last, we present an overview of the behaviors used in the literature to assess engagement and inventory the behavioral assessment tools used for people with dementia.

Table 2. Other components of engagement: person attributes, experiential/subjective components, object/agent characteristics

TYPE OF COMPONENT	COMPONENTS	DESCRIPTION
Person Attributes	Intrinsic motivation / Curiosity / Interest / Proactivity	When activities are autotelic, they are more likely to elicit curiosity, interest, and, in general, proactive behavior (Nakamura and Csikszentmihalyi, 2009).
	Past and present preferences	These are the past and present interests of a person. Cohen-Mansfield et al. (2010a) found out that activities leaning on these interests are more engaging.
Experiential/ Subjective Components	Skills / Challenge - Control / Richness	Attfield's balance between control and richness (2011) is a synonym for Csikszentmihalyi's balance between skills and challenges (2014).
	Loss of self and external consciousness	Loss of self and external consciousness refers to the lack of contact with oneself and with the surrounding environment experienced during flow and immersion (Brown and Cairns, 2004; O'Brien and Toms, 2008; Nakamura and Csikszentmihalyi, 2009).
Object/Agent Characteristics	Aesthetics	The sensory and visual appeal of an interface (Attfield et al., 2011; O'Brien and Toms, 2008). The appearance and qualities of a stimulus have an effect on engagement (Cohen-Mansfield et al., 2010a and 2010b).
	Novelty	The surprising, unfamiliar, and unexpected effect that interactive technologies are likely to elicit. The novelty of an interactive system is a feature that is likely to positively affect engagement (Attfield et al., 2011).
	Reputation / Trust	Reputation is the trust users globally invest in a resource. Trust is the reliability of that resource (Attfield et al., 2011).
	Feedback / Reciprocity / Interactivity	Feedback is the information communicated to the users as a result of their actions (O'Brien and Toms, 2008; Stone et al., 2005). Reciprocity is the bi-directionality of social engagement. Interactivity is the ability of an interactive system to respond and reciprocate user's inputs (O'Brien and Toms, 2008).

2.4.1 Experiential/Subjective Measures

As follows, we introduce three self-reports of engagement. These were selected for two reasons: (i) they assess the subjective experience of engagement in a quantitative manner and (ii) they are related to frameworks of engagement. The Experience Sampling Method (ESM) attempts to measure *flow* (Hektner et al., 2007; Nakamura and Csikszentmihalyi, 2009). The Game Experience Questionnaire (GEQ) aims at quantifying *game immersion* (Brown and Cairns, 2004). The User Engagement Scale (UES) is targeted at assessing O'Brien and Toms' conception of *engagement* (2008).

The *ESM* (Kubey et al., 1996) is a self-report tool that enables the study of the subjective experience of flow in natural environments and everyday life. When measuring flow with the ESM, the respondents are typically equipped with a pager that emits an alarm at random intervals (usually 10 times per day). When the pager rings, the respondent is asked to note down pieces of information regarding his/her momentary psychological state. The ESM contains open questions regarding the location, social context, primary and secondary activity, content of thought, and time of the experience. Moreover, it features a number of Likert scales aimed at assessing the respondent subjective state regarding affect (i.e., happy, cheerful, sociable, friendly), activation (i.e., alert, active, strong, excited), cognitive efficiency (i.e., concentration, ease of concentration, self-consciousness, clarity of mood), and motivation (i.e., wish to do the activity, control, feeling of involvement).

The *GEQ* (IJsselsteijn et al., 2008) is a self-report assessment tool of the quality of game experience. It is made of three modules: (i) the *core questionnaire* which measures the game experience based on seven items (five-point Likert scale): immersion, flow, competence, positive and negative affect, tension and challenge; (ii) the *social presence* module which assesses the psychological and behavioral involvement of the player with a virtual, mediated, or co-located co-player; and (iii) the *post-game questionnaire* that gauges how the player feels after playing. The GEQ has also an in-game version which can be used to score engagement while the game is played. However, it is important to note that IJsselsteijn and colleagues did not publish any evidence on the psychometric reliability and validity of the GEQ, thus the properties of this assessment tool are still debatable (Norman, 2013).

The *UES* (O'Brien et al., 2018) is a 31-item experiential questionnaire that evaluates engagement in several settings, such as education, consumer applications, haptic technologies, and video games. The 31 items of the UES are organized around six dimensions: focused attention, perceived usability, aesthetic appeal, endurability, novelty, and felt environment.

As one can note, the ESM and GEQ include items that concern the presence of a social interactor during the engaging experience, while the UES does not. In general, as pinpointed before, rapport is not measured through its subjective self-reported correlates, but rather through its behavioral expression.

Self-reports of engagement for people with dementia are mainly qualitative and those that aspire to be quantitative are in general obtained by means of verbal accounts. One of the few self-reports for people with dementia is the Subjective Impression Questionnaire (SIQ; Wada et al., 2005). The *SIQ* is a four-item questionnaire administered verbally to people with dementia after the interaction with a social robot, it aims to gather the overall impression of the person with dementia regarding the activity. The questions in the SIQ are the following: (i) Is the robot cute/ugly? (ii) Do you like/dislike the robot? (iii) Is playing with the robot fun/boring? (iv) Do you want to play with the robot again? These questions are filled out on a five-point Likert scale by a facilitator or caregiver based on the content of the verbal report of the person with dementia. The SIQ has a number of defects: (1) it relies on closed questions that do not allow the rating of the magnitude of the dimensions cute/ugly, like/dislike, fun/boring on a Likert scale, (2) it relies on a proxy interpreting the meaning of a short subjective report, and (3) it is highly prone to an acquiescence bias as the negative reply to any of the four questions is likely to sound socially inappropriate. Albeit the SIQ has several drawbacks, the drought of self-assessment instruments of engagement targeted to people with dementia has brought several researchers to adopt it (Yu et al., 2015). This brings us to the third requirement of a measurement framework of engagement in dementia, *investigate the use of self-reports*.

2.4.2 Peripheral–physiological Measures

Most of the studies regarding the use of peripheral-physiological measures to assess engagement concern the fields of game experience (Ravaja et al., 2008; Nacke et al., 2010; Drachen et al., 2010; Negini et al., 2014), entertainment technology design (Silveira et al., 2013; Mandryk et al., 2006a and 2006b), social HRI (Leite et al., 2013; Henriques et al., 2013), and media psychology (Van den Bosch et al., 2013). Across these studies, EDA appears to be the most widely employed psychophysiological measure. EDA is the electric change in the skin deriving from the activation of the sympathetic nervous system (SNS). This activation – which is triggered by episodes of excitement, attention, anxiety, and high cognitive load (Andreassi, 2010) – leads to the secretion of sweat from the eccrine glands, a type of glands situated under the palms of hands and soles of feet, that react to psychological stimulation alongside temperature.

The extensive use of EDA as a measure of engagement is to be ascribed to the fact that this psychophysiological measure is straightforward to understand, low-cost if compared to other psychophysiological measures, and minimally intrusive. Also, it is more reliable in terms of engagement recognition with respect to other measures, such as heart rate (HR) and heart rate variability (HRV). As an organ, the skin is exclusively innervated by the SNS, while the heart is dually innervated by the sympathetic and the parasympathetic nervous system (PNS). Sympathetic activation – which is related to emotional arousal – causes the heart to beat faster; parasympathetic activation – which is related to information intake and attentional allocation – causes the heart to beat slower (Ravaja, 2004: Ravaja et al., 2008). As engagement is a compound of both emotional and attentional processes, the use of cardiac measures to assess engagement might lead to somewhat counterintuitive results. To tackle this issue, HR and HRV are often employed in combination with other physiological measures that can help in discriminating whether the source of cardiac activation is sympathetic or parasympathetic. HRV is particularly interesting in the context of engagement as it can be used to assess and visualize arousal, but also mental stress (Taelman et al., 2009; Liang et al., 2018; Yu et al., 2018). Unfortunately, most of the studies employing HR and HRV are conducted with young participants – mostly university students (e.g., Mandryk et al., 2006a and 2006b, Ravaja et al., 2008, Drachen et al., 2010; Henriques et al., 2013; Negini et al., 2014; only Nacke et al., 2010 contemplated a slightly older population with an age of 18 to 41 years) - and are hence hard to generalize to the dementia population which is usually composed of people older than 65 years of age. With the increase of age, the incidence of cardiovascular disease grows exponentially. In people aged between 20 and 39 years, the prevalence of cardiovascular disease is of the 11.9% (15.9% in men, 7.8% in women), in persons with an age comprised between 60 and 79 years and with an age over 80 years, it is respectively of the 73% (73.3% in men, 72.6% in women) and of the 82.6% (79.3% in men, 85.9% in women; Yazdanyar and Newman, 2010; Prince et al., 2015). Cardiovascular disease might intrude the reliability of HR and HRV in elderly people. Moreover, patients with dementia present a decreased variability in their HR as noted by Giblin (2016), Giblin et al. (2013), and Phillips (2011) and further confirmed by Treusch et al. (2015).

Peripheral-physiological measures of engagement lean on dimensional models of affect to infer engagement. The most eminent of these is the *circumplex model of affect* (Russell, 1980; Posner et al., 2005). According to this model, any emotion resides in a two-dimensional space defined by two orthogonal axes: arousal and valence (see Figure 3). *Arousal* describes the degree of activation of a certain affective state (deactivated-activated), *valence* describes its pleasantness (unpleasant-pleasant). EDA – which univocally measures arousal – has been used to detect engagement in conjunction with other physiological measures, more suited to measure valence. One example is facial electromyography (EMG). Facial EMG measures the electrical activity of three muscles of the face: orbicularis oculi, zygomaticus major, and corrugator supercilii. The *zygomaticus major* and *orbicularis oculi* are activated by the act of smiling and laughing. The *corrugator supercilii* is activated by the act of frowning. Facial EMG is particularly suited to discriminate positive and negative affective states (Mandryk et al., 2006a; Mandryk et al., 2006b; Ravaja et al., 2008; Nacke et al., 2010). However, it is extremely intrusive as it implies that electrodes are to be placed on participants' faces. The placement of electrodes on participants' faces is not only likely to produce artifacts (e.g., noise) in the signal as it causes nervousness, but also prevents the experimental subjects to forget that they are taking part in a study, thus producing a threat to the internal validity of a study called 'artificiality'.

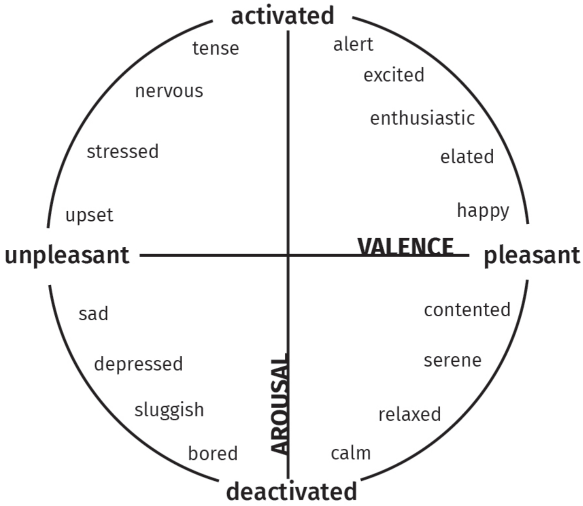


Figure 3. Circumplex model of affect (adapted from Posner et al., 2005)

Studies on the physiology of engagement of people with dementia are scarce and involve costly or invasive procedures, such as EEG (Wada et al., 2005), urinalysis and hormones analysis (Wada and Shibata, 2008), and fNIRS (Kawaguchi et al., 2012). In the context of dementia, we found only two studies conducted using EDA, both for purposes beyond our interest (Treusch et al., 2015; Alam et al., 2016). The only real antecedent to our work is a study carried out with healthy seniors and seniors with Mild Cognitive Impairment (MCI) during interactions with a telepresence robot (Giraff) using cardiac measures, such as HR and heart rate variability (HRV; Tiberio et al., 2012). As a consequence, it is extremely important for the study of engagement in people with dementia *to explore the measurement of the physiology of engagement in the field with unobtrusive sensors*.

2.4.3 Behavioral/Expressive Measures

Behavioral measures of engagement follow a different trend with respect to experiential/subjective and peripheral-physiological measures: they are scarce for healthy subjects and abundant for people with dementia. Apparently, when the subjective facet of engagement is easily reached, its behavioral correspondents are neglected.

Most behavioral metrics of engagement for healthy subjects come from the field of social psychology (i.e., interaction studies on children) and social HRI. In this latter context, researchers seek for behavioral indicators that robots can track with sensors and use to infer engagement states. In this sense, *gaze* is one of the most exploited behavioral cues of engagement (Peters et al., 2005, Castellano et al., 2009b; Nakano and Ishii, 2010; Bednarik et al., 2012) as it provides the robot with a clear idea of what the user is paying attention to.

When a social robot takes up the role of a game companion, it is crucial for it to keep track of whether its human interactor is paying attention to the game or not (Castellano et al., 2009b). Also, it can use gaze as a social hint to direct the attention of the human interactor towards a particular target in the game environment (Mwangi et al., 2018).

Rich et al. (2010) identify two gaze behaviors that regulate social interactions: *directed gaze* and *mutual gaze*. The former occurs when a person (the initiator) looks at an object in the immediate environment and, with a short delay, the other interactor (the responder) looks at the same object. The latter starts when the initiator looks at the responder's face and then, after a short lag, the responder looks back at him/her. Other extensively used behavioral cues of engagement are *backchannel events* (Tickle-Degnen and Rosenthal, 1990; Rich et al., 2010; Inoue et al., 2016). These are, for instance, nods and "saying uh, huh" and are typically used in conversation to notify the responder's comprehension of the initiator's communication.

Also *facial expression* and *postures* are commonly used to quantify engagement states. With regards to the first, the majority of the reviewed studies employ the Facial Action Coding System (FACS; Ekman and Rosenberg, 1997) – an anatomical taxonomy that describes all observable facial movements – to train machine learning models aimed at recognizing emotions from facial expressions (Castellano et al., 2009b; Jaques et al., 2016; Monkaresi et al., 2017). With regards to the latter, posture actions (e.g., sitting on the edge, leaning forward, sitting upright, leaning back, slumping back), joint kinematics (i.e., the motion of joints or body segments), and posture features (i.e., body lean angle, slouch factor, quantity of motion, contraction index) are utilized to define engagement levels (Mota and Picard, 2003; Sanghvi et al., 2011; Ge et al., 2016; Jaques et al., 2016).

In the context of dementia, we identified three methodologies to assess engagement through behavior: observational rating scales, ethograms, and coding schemes.

1. *Observational rating scales* (see Table 3) are Likert-type scales that gauge engagement on a number of items operationalized through behavior. Some examples are the Observational Measurement of Engagement (OME; Cohen-Mansfield et al., 2009), the Menorah Park Engagement Scale (MPES, Judge et al., 2000), and the Observed Emotion Rating Scale (OERS, Lawton et al., 1996).

2. *Ethograms* (see Table 4) are comprehensive and accurate inventories of engagement-related actions observed in context and used to annotate videos. These are for instance developed by Olsen et al. (2016), Jøranson et al. (2016), and Jones et al. (2015).

Table 3. Observational rating scales of engagement for people with dementia

OBSERVATIONAL RATING SCALE	ITEMS	OPERATIONAL DESCRIPTION / BEHAVIORS
OME Cohen-Mansfield et al., 2009	*Duration* (in seconds or minutes)	Length of time that the resident is occupied or involved with a stimulus.
	Attention (4-point Likert scale)	Amount of attention the resident is paying to an object during the engagement (manipulating/holding/content of talking about object). Following staff instructions without any change in affect is still attention. Attention can be visual or physical: (i) *visual engagement*: eye contact, eye tracking, visual scanning, (ii) *physical engagement*: i.e., stroking cat even if looking away.
	Attitude (7-point Likert scale)	Amount of excitement/expressiveness toward stimulus: smiling, frowning, energy, excitement in voice.
	Refusal (yes or no)	Whether or not the participant refuses the stimulus.
MPES Judge et al., 2000	*Non-engagement* (1st point of the Likert scale)	Staring off into space or another direction away from the activity, sleeping, or any motor and/or verbal behavior in response to an activity that the client is not currently participating in. Comparable to disengagement and apathy
	Self-engagement (2nd point of the Likert scale)	Motor, verbal, listening and/or looking behavior during a transition period when an activity is not offered or when the client chose not to participate in the scheduled activity.
	Passive engagement (3rd point of the Likert scale)	Listening and/or looking behavior exhibited in response to the activity the client is participating in. Listening to a discussion or a speaker, watching others paint or color in an art therapy project, and listening to music.
	Constructive engagement (4th point of the Likert scale)	Motor or verbal behavior exhibited in response to the activity in which the client is taking part. Talking in a discussion group, painting in a creative arts activity, and singing or dancing to music.
OERS Lawton et al., 1996	*Pleasure* (5-point Likert scale)	Laughing, Singing, Smiling, Kissing, Stroking or Gently touching other, Reaching out warmly other.
	Anger (5-point Likert scale)	Physical Aggression, Yelling, Cursing, Berating, Shaking fist, Drawing eyebrows together, Clenching Teeth, Pursing lips, Narrowing eyes, Making distancing gestures.
	Anxiety/Fear (5-point Likert scale)	Shrieking, Repetitive calling out, Restlessness, Wincing/Grimacing, Repeated or agitated movements, Line between eyebrows, Lines across forehead, Hand wringing, Tremor, Leg jiggling, Rapid breathing, Eyes wide, Tight facial muscles.
	Sadness (5-point Likert scale)	Crying, Frowning, Eyes Drooping, Moaning, Sighing, Head in hand, Eyes/Head turned down and Face expressionless.
	General Alertness (5-point Likert scale)	Participating in a task, Maintaining eye-contact, Eyes following object or person, Looking around room, Responding by moving or saying something, turning body or moving towards object or person.

Table 4. Ethograms of engagement for people with dementia

ETHOGRAM	BEHAVIOR GROUPS / BEHAVIORS / OPERATIONAL DESCRIPTIONS (when available)
Olsen et al., 2016	– Conversation (unspecified target) – Look at (other people, the dog activity, other things) – Touch (people, dog) – Smile or laugh at (dog, other things) – Sing/dance/clapping hands – Stereotyped behavior – Wandering around – Agitated behavior – Yawn and sigh – No response/ Asleep – Leaving the room – Off camera
Jøranson et al., 2016	– Conversation with or without PARO – Observe (PARO, other participant/activity leader, other things in the room), – Smile/laughter (PARO, other participant/activity leader) – Physical contact with PARO – Active with PARO – Singing/whistling – Clapping/humming/dancing – Napping – Walking around/ Repetitive movement – Time out of recording – Physical contact (with participant/activity leader) – Signs of discomfort – Leaving the group – No response to contact.
VC-IOE Jones et al., 2015	Facial emotional response The OERS items: Pleasure, anxiety/fear, anger, sadness, general alertness, none. Verbal engagement – *Positive verbal engagement with stimulus*: Participating and maintaining conversation, verbally responding to statement/questions specifically to stimulus and about stimulus. – *Positive verbal engagement with facilitator*: General talking. Participating and maintaining conversation, verbally responding to statements/questions that are not relevant to stimulus. – *Negative verbal engagement*: Verbalizes the desire to leave. Refuses to participate in the activity by verbalizing “no”, “stop”, etc. Makes repetitive generalized somatic complaints. Cursing and swearing. – *No verbal engagement*: Not participating and maintaining conversation. Not responding or talking to the facilitator when prompted. – *Missing*: No audio or distorted audio. Visual alertness/engagement – *Visually engaged with stimulus*: Appears alert and maintaining eye contact with stimulus. Eyes following stimulus or looking at stimulus. – *Visually engaged with facilitator/others*: Appears alert and maintaining eye contact with facilitator or others. Eyes following facilitator or others. Looking at facilitator or others. – *No visual engagement*: Appears inattentive. Blank stare into space. Does not make eye contact with stimulus, facilitator or others. – *Missing visual*: Cannot see face to determine visual engagement.

ETHOGRAM	BEHAVIOR GROUPS / BEHAVIORS / OPERATIONAL DESCRIPTIONS (when available)
	<u>Behavioral engagement</u> – *Positive behavioral engagement*: Touching or attempting to touch stimulus. Stroking, petting, nuzzling, holding and handling stimulus appropriately. – *Negative behavioral engagement*: Hitting, shaking and handling stimulus inappropriately. Shoving stimulus away. Pulls back from stimulus or facilitator. Maintains a stiffness of extremities. – *No behavioral engagement*: No touching; no physical contact with stimulus or not handling stimulus. Rests stimulus on lap or on furniture, but not handling or interacting with stimulus. – *Missing behavior*: Cannot see the body posture of the resident. <u>Collective engagement</u> – *Using stimulus for collective engagement*: Encouraging others to interact with stimulus. Introducing stimulus to others. Using stimulus as a communication channel to interact and talk with others (e.g., family members, staff and other residents). – *No evidence of collective engagement* <u>Agitation</u> Based on Cohen-Mansfield Agitation Inventory – CMAI – *Evidence of agitation*: Restlessness, repeated/agitated movement (frequent non-purposeful movement), moving in chair, picking and fiddling with clothes; repetitive rubbing own limbs or torso; appears anxious. Repeats words or phrases (exclude stumbling over word/phrase), abusive or aggressive toward self or other. – *No evidence of agitation*: No sign of agitation as described above.

Table 5. Coding schemes of engagement for people with dementia

CODING SCHEME	BEHAVIORS / OPERATIONAL DESCRIPTIONS (when available)
Wada et al., 2010	– Emotional expression (laugh, smile, no expression, hate) – Gaze (PARO, staff, user, others) – Talk (PARO, staff, user, others) – Type of interactions with PARO (give, stroke, hold, other)
Cruz et al., 2011	– *Engagement in the task*: The resident moves the body or a body part in order to perform a task related to the activity. – *Interactions with objects*: The resident voluntarily moves the body or a body part in the direction of an object, reaching it. – *Verbal communication*: The resident articulates words or sentences with meaning, voluntarily and purposely, in order to communicate with another person. – Smiling – Laughing – Nodding the head – Closed eyes
Šabanović et al., 2013	– *Visual engagement*: Look at the robot. – *Verbal engagement*: Speak, sing, vocalizations towards the robot. – *Physical engagement*: Pet, hit, hold, kiss, take/offer PARO.
Takayanagi et al., 2014	– Talk (to PARO/lion, to the staff, to him/herself or to nobody) – Touched or stroked (PARO/lion) – Positive, neutral or negative facial expression.

3. *Coding schemes* (see Table 5) are excerpts of ethograms aimed at answering specific research questions (Martin and Bateson, 2007; Bakeman and Quera, 2011). Some examples of coding schemes are those of Wada et al. (2010), Cruz et al. (2011), Šabanović et al. (2013), and Takayanagi et al. (2014).

As one can notice from Tables 3, 4, and 5, observational rating scales, ethograms, and coding schemes feature some of the behavioral metrics identified in social HRI, namely gaze, facial expressions, postures, and backchannels. However, they take on a more multimodal and exhaustive approach and also include *affective touch* (e.g., stroke, hold, clapping), *facial gestures* (e.g., kissing, yawning, wincing/grimacing), *manipulations* (e.g., hold, touch), *vocalizations* (e.g., singing, humming), *content of conversation* (e.g., yelling, cursing, berating), *stereotyped behaviors*, and *agitated behaviors* (e.g., hand-wringing, wandering around) in the measurement of engagement.

Observational rating scales are very useful tools to get a broad idea of the engagement state of the person with dementia during activities. However, they can grasp engagement only at a global level. Indeed, they do not get into the detail of how behavior naturally occurs and unfolds. They collect a general idea of engagement drawn from the meaning of certain behaviors. *Coding schemes* provide a deeper understanding of behavior compared to observational scales. However, they grasp only some characteristics of behavior. Indeed, instead of considering behavior in its natural flow, they fragment it to pick up only the desired pieces of information. In some cases, since the fragmentation of behavior is not performed in a systematic way – as a result of a preexisting ethogram – it results in a cherry-picking of behaviors based on their perceived meaningfulness. *Ethograms* are optimal to study behavior in its complexity as it naturally occurs and flows. However, they produce a measurement of engagement that is not integrated, but segmented into many small pieces of information that cannot be traced back to a whole body dynamic and hence to an overall engagement state.

The main flaw of state of the art assessment tools of engagement-related behavior does not reside in what they measure, but in *how* they measure it. Social HRI metrics conceive behavioral modalities in isolation. Observational rating scales employ behaviors to operationalize the different items of engagement. Ethograms and coding schemes fragment behavior into many small pieces difficult to bring back to an overall engagement state. In general, there is the need for a measurement system that saves the complexity and multimodality of behavior, allows the assessment of behavior in its whole-body configuration, and enables researchers to make inferences regarding the engagement state that a certain body configuration represents. This can be substantiated in the following requirement: *develop a methodology to measure engagement-related behavior in a more consistent way across activities*.

2.5 Conclusions

This chapter has given a systematic and thorough account of the literature surrounding the notion of engagement. First, it has scrutinized the major definitions of engagement coming from different scientific fields. Second, it has enumerated and described the components of engagement. Last, it has identified several measures of engagement and organized them along the three response systems.

According to the state of the art, engagement is the *positive quality of the interaction* between a person and an object (e.g., a game, an activity) or agent (e.g., a person, a social robot) which is influenced by the *attributes* of the person experiencing it, by the *characteristics* of the object/agent with which the person is involved, and by the *context* where the interaction takes place. If we leave aside the elements influencing engagement (i.e., person attributes and object characteristics) and just focus on engagement as a state, we can isolate three observable components of engagement – *attention*, *affect*, and *time investment/duration*[2] – and three experiential components of engagement – *time perception/endurability*, *control* (or balance between challenges and skills), and *loss of self* and *external consciousness*.

In terms of measurement, engagement can be compared to an iceberg, which has a submerged and an emerged part. The submerged part is the subjective side of engagement and can be measured through *self-reports* – such as the ESM, the GEQ and the UES. The emerged part is the observable side of engagement and can be gauged through *psychophysiological measures* – like EDA, EMG, and HR – and *behavioral metrics* – for example, gaze, facial expression, and postures.

When the engagement of people with dementia is at stake, its definition is deprived of the usual positive connotation and of any subjective quality and its components are presented as observational constructs. Moreover, assessment tools of engagement are few and mostly grounded on behavior observation. In order to pursue a more appropriate measurement of the engagement of people with dementia, the framework to build needs to meet the following requirements:

- *Requirement 1*: Deploy a *definition* of engagement capturing its true connotation.

- *Requirement 2*: Identify the *components* of engagement and their boundaries.

2 Time (as duration of engagement) is not a component of engagement that one can infer from the behavior or physiology. It is more an observable consequence of the state of engagement.

- *Requirement 3*: Investigate the use of *self-reports.*

- *Requirement 4*: Explore the measurement of the *physiology* of engagement in the field with unobtrusive sensors.

- *Requirement 5*: Establish a methodology to measure engagement-related *behavior* in its units and as a body configuration across activities.

Also, in order to formalize the knowledge resulting from a definition of engagement and from the study of its components and multimodal measures, a framework of engagement for people with dementia should meet an additional requirement:

- *Requirement 6*: *Profile* the expression of engagement by determining the relationships among its different components and measures.

The next chapter will serve to narrow down the state of the art regarding engagement to the context of dementia, sketch the theoretical framework in which the present doctoral research operates, and present a tentative model of engagement for dementia based on the literature review.

3

Theoretical Framework

3.1 Introduction

This chapter circumscribes the state of the art on engagement to the context of dementia and outlines the theoretical framework of this dissertation. First, it proposes a new *definition* of engagement for dementia. Second, it settles the type of *activities* that fall within the definition. Third, it establishes which *components* of engagement are sensible and accessible in people with dementia. Last, it enumerates the possible *measures* of engagement for people with dementia. Finally, as a result of this process, it sets out a tentative theory-driven *model* of engagement that specifies the hypothesized relationships between the components of engagement and introduces a number of possible metrics to quantify them.

3.2 Provisional Definition of Engagement for Dementia

According to the literature, engagement is the proactive involvement with an object (e.g., a game, an interactive system) or an agent (e.g., a person, a social robot) that has a positive affective nuance. By extension, we provisionally define engagement in dementia as the psychological state of enjoyment and proactive attentiveness experienced by a person with dementia involved in a meaningful activity. The present definition of engagement:

(i) Identifies the person experiencing engagement – the person with dementia – and the object of his/her engagement – a meaningful activity.

(ii) Briefly sketches the main components of engagement identified in the literature – positive affect and attention – from the perspective of the person with dementia: *enjoyment* – the state or process of taking pleasure in something (Oxford dictionary) – and proactive *attentiveness* – the proactive action of paying close attention to something (Oxford dictionary).

This definition of engagement is still tentative. It will turn definitive should the model of engagement encompassing it be confirmed.

3.3 Co-activities for People with Dementia

According to the ethnographic study that we performed at the beginning of this doctoral project (see Appendix A), individual activities are mostly targeted to persons with severe to very severe dementia, while most of the activities for people with a dementia ranging from mild to moderate are co-activities. In the context of this doctoral dissertation, we decided to focus on co-activities. This is because co-activities constitute a non-pharmacological form of prevention of the progression of the disease and also because they respond to three needs of the person with dementia: (i) *the need for occupation*, as to say the necessity to feel useful and participate to meaningful activities, (ii) *the need for attachment*, which refers to the necessity to form bonds with others, and (iii) *the need for inclusion*, which has to do with the necessity to feel part of a social group, and restrain from loneliness and isolation (Vila-Miravent et al., 2012).

As described in Appendix A, within this dissertation, we focus on co-activities that have the following characteristics: (a) they do not entail *physical effort*, (b) they envision a *proactive role* for the person with dementia, and (c) they involve the use of *tangible artifacts*. Physical activities are excluded as they prospect a type of engagement distinct from the one profiled in the state of the art. Proactive activities are preferred because they are more likely to elicit intrinsically motivated engagement. Activities mediated by tangible artifacts are privileged as they provide people with dementia with important multisensory stimulation. Moreover, they are a superset of technology-mediated activities, such as robot-mediated interaction.

3.4 Components of Engagement in Dementia

According to the literature, engagement is a compound of *subjective elements* – loss of self and external consciousness, sense of perceived control over the activity, and time distortion – and of *observable elements* – attention and affect. In dementia, the subjective components of engagement cannot be reliably measured. This is because the detachment from reality and loss of spatial and temporal reference points that is so distinctive of the subjective experience of engagement is in fact the normal condition of most persons with dementia. As a consequence of this, we can just gather the subjective opinion of the person with dementia on the likeability of the activity and potentially use it as gold standard. However, the only components of engagement that can be reliably gauged in people with dementia are the observable ones. These are *attention*, *affect*, and – as we concentrate on co-activities – also *rapport*.

To stress the connection with the state of the art and make the notions of attention, affect, and rapport more clear, hereafter, we report their definitions:

1. Attention, or *focused attention*, is the voluntary focusing of attention on a limited stimulus field (Csikszentmihalyi, 2014).

2. Affect, or *core affect*, is the neurophysiological state accessible to consciousness as a single simple feeling (Russell, 2003) which is a blend of two dimensions, *valence* (displeasure-pleasure) and *arousal* (deactivation-activation).

3. *Rapport* is a meaningful human experience of close and harmonious connection with another that involves common understanding (Tickle-Degnen, 2006).

3.5 Measures of Engagement for Dementia

The measurement framework of engagement that we draw from the literature encompasses all three response systems: experiential/subjective, behavioral/expressive, and peripheral-physiological. However, it attributes them different functions. The *experiential/subjective system* – constrained as described above to the subjective opinion of the person with dementia on the likeability of the activity – is used (if deemed reliable) as a gold standard of engagement. The *behavioral/expressive system* is utilized to gauge the components of engagement *focused attention*, *rapport* and *valence*. The *peripheral-physiological system* is used to quantify the component of engagement *arousal*, which, together with valence, gives an account of core affect.

With regards to the experiential/subjective level of measurement, we employ the only available self-report of engagement for dementia, the *SIQ* (Wada et al., 2005). As this tool does not come with a validity and reliability check, we utilize it in an exploratory study to ascertain its feasibility and are ready to resort to *observational rating* scales in case it does not provide sensible results. With respect to the behavioral/expressive system, we develop a dedicated *coding system* of engagement by adopting a mixed approach. We use *ethograms* to create exhaustive activity-dependent taxonomies of behaviors and then employ *Laban-Movement Analysis* (Laban, 1966) to interpret and organize them in sensible categories workable across activities. With regards to the peripheral-physiological system, we utilize *EDA* as a measure of the psychophysiological correlates of engagement. We reviewed previous studies carried out with healthy adults and children (Henriques et al., 2013; Leite et al., 2013) and identified the following features to extract: signal magnitude area, mean, standard deviation, range, summation of harmonics, number of peaks ratio, kurtosis, and skewness.

3.6 Tentative Theory–driven Model of Engagement for Dementia

In this section of the chapter, we present a tentative version of the model of engagement that we are going to further define and test after data collection. This model is the result of the review of the literature on engagement and of its organization in the theoretical framework specific to dementia that we have just outlined. As the model strongly relies on the literature, we call it *theory-driven*.

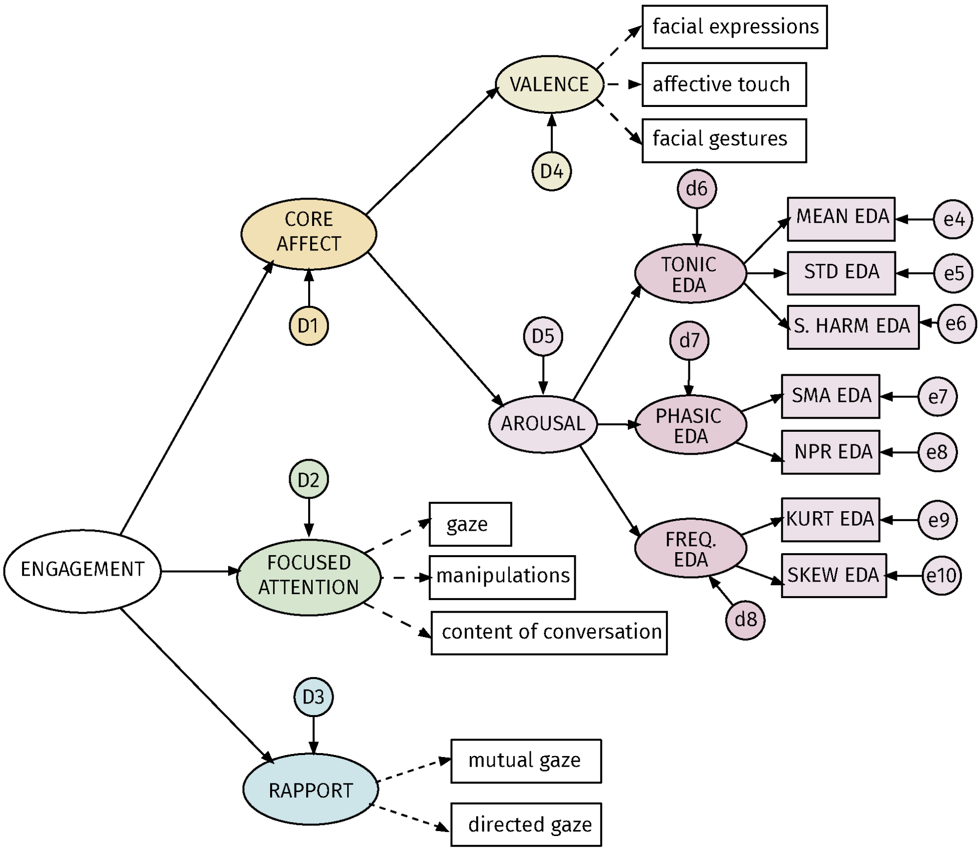


Figure 4. Tentative theory–driven model of engagement

The theory-driven model of engagement is drawn according to path diagram notation, a symbolic language that describes different types of variables and the relationships between them. Per conventional path diagram notation, latent variables – variables that are not directly observable (e.g., intelligence, trust) – are depicted by circles and indicators – variables that are directly measurable (e.g., height of a person) – are depicted by squares or rectangles. Unidirectional arrows (→) between these variables represent direct effects (i.e., regressions), while curved bi-directional arrows (↔) symbolize covariances (i.e., correlations; Brown, 2006).

In the theory-driven model, the state of *engagement* is represented as a latent variable that has a direct effect on another three latent variables. These are the components of engagement in dementia: focused attention, core affect, and rapport. *Focused attention* and *rapport* are measured through a number of behavioral indicators which are not yet established. In Figure 4, we present some possible metrics, but connect them to the latent variables via dotted arrows to stress their temporariness. As for *core affect*, given its blended nature, it bifurcates into two ulterior latent variables, the orthogonal axes of the circumplex model of affect, arousal and valence. Likewise focused attention and rapport, also *valence* is measured through behaviors yet to be determined. With regards to *arousal*, instead, it is assessed through the features of EDA. These are organized around the components of EDA that they reflect: *tonic EDA* (i.e., slower acting components and background characteristics of the signal), *phasic EDA* (i.e., faster changing elements of the signal, such as its peakedness), and the *frequency distribution of EDA*. According to this model, the more engaged is the person with dementia, the more focused attention and rapport increase, valence becomes positive, and arousal rises. This functioning of engagement closely depicts the definition of engagement given by the literature.

3.7 Conclusions

This chapter has established the theoretical framework of the dissertation in terms of definition, components, and measures of engagement for dementia. Moreover, it has proposed a tentative theory-driven model of engagement that systematizes the relationships between the components of engagement and outlines how these components can be measured.

The theory-driven model of engagement that we presented here is still tentative. To deploy the definitive version, we first need to gather a large sample of multimodal data. Then, we need to develop a coding system of engagement-related behavior and prove its reliability. Last, we must verify the adequacy of each of the selected measures of engagement against the gold standard. Only after these three steps, we can finalize the theory-driven model of engagement and test it with structural equation modelling (SEM).

In appendix B, we describe the exploratory study that brought us to the definitive exclusion of self-reports as a gold standard of engagement for people with dementia. In the next chapter, we detail the process of data collection in terms of methods, procedures, and instruments. In Chapter 5, we assess the adequacy of EDA as a measure of engagement. In Chapter 6, we present the coding system of engagement-related behavior, discuss the results of the inter-rater reliability, and test its adequacy as a measure of engagement. Only then, in Chapter 7, we describe the final version of the theory-driven model of engagement and run it with SEM.

4

Data Collection: The Experimental Study

4.1 Introduction

This chapter describes the experimental study that we carried out to collect a multimodal database of engagement-related data. Within this experimental study, we collected data using video cameras, wearable physiological sensors, and observational rating scales. These data served four purposes: (i) *video recordings* enabled us to build a coding system of engagement-related behavior and attest its adequacy as a measure of engagement, (ii) *wearable physiological sensors* allowed us to collect EDA and verify its appropriateness as a measure of engagement, (iii) *observational rating scales* provided us with validated and reliable experts' estimates of engagement usable as gold standard to assess the concurrent validity of the coding system and of EDA. Finally, taken together, (iv) the multimodal data led us to finalize and test the theory-driven model of engagement.

The present chapter focuses on detailing the methods, procedures, and instruments utilized for data collection and motivating them in virtue of the theoretical framework sketched in the previous chapter and of the ethnographic and exploratory studies discussed in Appendix A and B. The results of the experimental study that reply to the research questions of this dissertation are discussed in the subsequent chapters.

4.2 Participants

The multimodal data were collected in two nursing homes in the province of Barcelona (Redós de Sant Josep i Sant Pere and La Mallola). Fourteen participants with an age comprised between 69 and 92 years (M_{age}= 83.93, STD_{age}= 7.28) were selected among the residents of the nursing homes to take part in the study. *Inclusion criteria* for the participation in the study were a diagnosis of mild to moderate dementia and the informed consent of both the participants and their legal guardians. *Exclusion criteria* were severe dementia, acute visual impairment, bedridden condition, reduced motility in the upper limbs, Parkinson's disease, Parkinson's disease dementia, and strong hallucinatory or delusional states.

The decision to exclusively focus on mild and moderate dementia was taken as a result of the ethnographic study. Indeed, during the observation of leisure activities in the nursing homes, we realized that people with severe dementia displayed behaviors that might be given an engagement meaning, but were in fact due to the severity of the clinical condition (e.g., sleeping). At this level, we needed to find unambiguous physiological and behavioral correlates of engagement, thus we resorted to people with mild and moderate dementia. In a second moment, the results of this study can be generalized to people with severe dementia.

The selection of participants was performed by the clinical staff of the nursing homes (geriatrician and psychologist) on the basis of the clinical picture of the residents in three steps: (1) exclusion of residents affected by severe dementia, Parkinson's disease, Parkinson's disease dementia, and mobility issues in the upper limbs; (2) exclusion of residents affected by MCI and diseases other than dementia; and (3) exclusion of residents not willing to participate to the study or sign the informed consent[3]. Out of seventeen participants that were proposed the study, one refused to participate due a privacy issue and two dropped out after one session for different reasons. The first participant fell ill, the second found it difficult to participate to the activity due to a severe form of agitation and wandering.

The residents that were willing to participate in the study were further screened with the Reisberg Global Deterioration Scale (Reisberg GDS; Reisberg et al., 1982), the *Mini-Examen Cognoscitivo* (MEC, the Spanish version of the Mini-Mental State Examination; Vinyoles Bargalló et al., 2002), and the Neuropsychiatric Inventory – Nursing Home version (NPI-NH; Cummings et al., 1994). The Reisberg GDS and the MEC served to assess dementia severity, the NPI-NH to determine whether or not the selected participants suffered from hallucinations and delusions. Residents with a score of 4 or 5 at the Reisberg GDS, a score ranging between 10 and 23 at the MEC, and a score inferior to 4 in the items *delusions* and *hallucinations* of the NPI-NH were included in the study.

As the focus of this dissertation is modeling the engagement of persons with dementia in *co-activities*, the 14 participants meeting the inclusion and exclusion criteria were randomly coupled and took part in the study in pairs. The participants in the couples did not know each other prior to the start of the research.

4.3 Experimental Design

The experimental study followed a repeated measures design and featured two activities as experimental conditions: game-based cognitive stimulation and robot-based free play. In the *game-based cognitive stimulation*, the seven couples of participants were asked to collaboratively complete three board games: jigsaw puzzles, shape puzzles, and a match with the tiles of domino. In the *robot-based free play*, they were asked to interact with the animatronic pet robot Pleo (www.pleoworld.com).

Each activity was presented in a different session and was repeated three times within the study, for a total of six experimental sessions per participant (42 sessions overall). Game-

3 An informative meeting was held in the nursing homes to explain the study in detail.

based cognitive stimulation and robot-based free play were alternated in order and presented to participants every other week. The order of presentation of board games was randomized using a Latin squares technique and was always different across sessions (see Table 6).

Table 6. Overview of sessions and order of presentation of games

1st session	2nd session	3rd session	4th session	5th session	6th session
Jigsaw puzzle 1 (4-6 pcs)	Play with Pleo	Dominoes	Play with Pleo	Shape puzzle 1 (9 slots)	Play with Pleo
Jigsaw puzzle 2 (4-6 pcs)		Jigsaw puzzle 1 (4-6 pcs)		Shape puzzle 2 (8 slots, no color contrast)	
Jigsaw puzzle 3 (6-9 pcs)		Jigsaw puzzle 2 (4-6 pcs)		Shape puzzle 3 (15 slots)	
Shape puzzle 1 (9 slots)		Jigsaw puzzle 3 (6-9 pcs)		Dominoes	
Shape puzzle 2 (8 slots, no color contrast)		Shape puzzle 1 (9 slots)		Jigsaw puzzle 1 (4-6 pcs)	
Shape puzzle 3 (15 slots)		Shape puzzle 2 (8 slots, no color contrast)		Jigsaw puzzle 2 (4-6 pcs)	
Dominoes		Shape puzzle 3 (15 slots)		Jigsaw puzzle 3 (6-9 pcs)	

All sessions of activities were conducted by a *facilitator* – one of the professionals working in the nursing homes (i.e., the psychologist or the social educator of the care facility) – at the presence of an *experimenter* – always the same researcher from the university. Given the impossibility of randomizing the variable facilitator by assigning each facilitator to always different couples, we controlled it and provided facilitators with scripts and guidelines to follow during sessions (see Appendix C).

The experimenter was present during sessions to monitor the instrumentation (i.e., video cameras, sensors) and check for the correct execution of the different phases of the experimental sessions. In order to reduce the reactivity of the participants to his/her presence, the experimenter took part in the activities of the nursing homes for one month prior to the start of the study.

4.4 Experimental Activities

The two activities that constituted the experimental conditions of the study were chosen as they fitted in the definition of *co-activities* that we provided in Chapter 3. Indeed, they envisaged the presence of others (being these the other resident involved in the same activity and the facilitator), did not require physical effort, involved the use of tangible artifacts, and required a certain degree of proactivity. Besides this, game-based cognitive stimulation and robot-based free play were selected as they differed in a number of aspects. They involved: (i) different skills (cognitive vs social and emotional), (ii) different degrees of challenge (right or wrong activity vs failure free activity), (iii) tangible artifacts with different degrees of interactivity (non-animated artifacts vs animated artifacts) and different interactive qualities (non-social artifacts vs social artifacts). As a consequence of these differences, the two activities were likely to prompt engagement states with diverse characteristics and diverse behavioral and physiological correlates. This was of substantial importance as one of the objectives of the experimental study was that of testing the adequacy of the selected measures of engagement across distinct activities.

4.4.1 Game–based Cognitive Stimulation

The board games to complete within the game-based cognitive stimulation were jigsaw puzzles, shape puzzles, and a match with the tiles of domino (see Figure 5). In the *jigsaw puzzles*, the players were asked to collaboratively assembly a set of pieces in a complete picture, usually of an animal. In the *shape puzzles*, they were requested to wedge a set of shapes, usually in wood, in a board with a series of slots. In the match with the tiles of *domino*, the players were requested to down a numbered tile from a set of seven that matched the tile on the table.

Figure 5. Board games: A) jigsaw puzzles, B) shape puzzles, and C) dominoes

The jigsaw puzzles and the shape puzzles to complete were three. They were presented in a progressive order of difficulty, from the easiest to the most difficult, across sessions (see Table 6). The challenge level of jigsaw puzzles was customized according to the cognitive level of the participants in the couples. Couples with one or both participants with mild dementia completed two six-piece puzzles and one nine-piece puzzle. Couples with both participants with moderate dementia completed two four-piece puzzles and one six-piece puzzle. With regards to the dominoes, all couples played one match with dominoes.

4.4.2 Robot-based Free Play

In the robot-based free play, the participants interacted with Pleo. Pleo is a robotic dinosaur developed by UGOBE which acts as a living pet (see Figure 6). It has an array of sensors that allow it to make sense of the surrounding environment and interact with people. For instance, touch sensors to discriminate among different types of touch, microphones to perceive sound and orientate towards it, ground foot sensors to detect surfaces, a camera-based vision system to detect light and navigate, and an internal clock to recognize the time to get up, eat, or sleep. Pleo is also able to display its internal states (e.g., hunger, sleep) and moods (e.g., happy, scared). We chose Pleo among the available robots, because, while being very interactive and responsive, it featured a series of traits that are demonstrated to be appealing to old people (Wu et al., 2012): it is small (in relation to human size), it has animal-like features, and its behavior mimics that of a domestic animal (e.g., cat and dog).

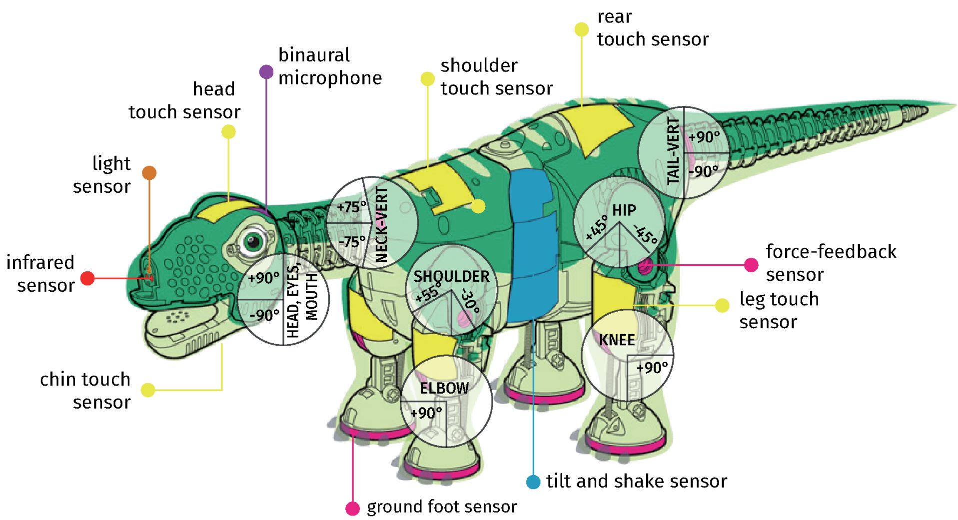


Figure 6. Pleo, overview of sensors and degrees of freedom (adapted from L-DOPA)

During sessions, participants interacted with Pleo in a spontaneous manner. However, due to the unstructured nature of the session, the facilitators were given a script with a list of activities that Pleo could support (e.g., feed Pleo, make Pleo sleep), so that they could prompt further interaction in case of a deadlock (see Appendix C).

4.5 Measurement Instruments

During the experimental sessions, we collected data using three types of instruments: (i) observational rating scales, (ii) video recordings, and (iii) wearable physiological sensors.

Video cameras and *wearable physiological sensors* provided the source of objective data necessary to deploy a sensible coding system of engagement-related behavior and to extract features from the EDA signal. *Observational rating scales* substituted self-reports as a gold standard for the testing of concurrent validity after the exploratory study (see Appendix B). Indeed, within this study, we noticed that self-reports entailed too much cognitive strain for participants and were not reliable in terms of responses. When asked to self-report their psychological states, participants struggled in recalling the activities to which they had participated. Moreover, they found it difficult to retrieve how they felt during them and to rank the different experiences. This result enabled us to meet the requirement *investigate the use of self-reports*. However, it also gave a negative reply to the first research question: *Can self-reports be employed as a gold standard of engagement in dementia?*

4.5.1 Observational Rating Scales

The observational rating scales of engagement employed in the study were the OME and the OERS (see Appendix D). With regards to the former scale, we used the items *attention* (four-point Likert scale; where 1= *not attentive* and 4= *very attentive*) and *attitude* (seven-point Likert scale, where 1= *very negative* and 7= *very positive*), using the latter twice, to obtain scores regarding the attitude of participants *towards the game* and the attitude *towards the partner*. Moreover, we added a further item, *cognitive difficulty* (five-point Likert scale, where 1= *not at all difficult* and 5= *very difficult*), present in further versions of the OME, to keep track of the level of challenge of the proposed activities. With regards to the OERS, we used it in its original version to rate the intensity of five affective states on a five-point Likert scale (where 1= *never* and 5= *more than 5 minutes*): *pleasure*, *anger*, *anxiety/fear*, *sadness*, and *general alertness*.

The choice of the items of the OME and OERS was based on the theoretical framework presented in Chapter 3. We employed:

1. The item *attention* of the OME and the item *alertness* of the OERS as measures of *focused attention*.

2. The item *attitude toward partner* of the OME to quantify *rapport*.

3. The item *attitude toward game* of the OME and the items *pleasure*, *anger*, *anxiety/fear*, and *sadness* of the OERS to gauge the dimension *valence* of *core affect*.

The OME and the OERS were completed after each session by the facilitators. We asked facilitators to fill out one OME and one OERS for each of the board games in the game-based cognitive stimulation and one OME and one OERS for the entire session of robot-based free play. We computed the median of the values of the three OME and three OERS filled out for board games and used it to perform statistical analyses.

4.5.2 Video recordings

All sessions were video recorded with two hand-held video cameras positioned one in front of and one on the side of participants. The video cameras were switched on as soon as the participants reached the activity room and were switched off once they left the room after the activity to go back to their units. Each session lasted around 50 minutes and the activities had a duration of 20-25 minutes. As a result, we collected 35 hours of video-footage, half of which (17.5 hours) were of activities. Albeit the presence of cameras might be thought of as a factor that could affect participants' behavior, we noticed that participants forgot about the cameras as the activity started.

The video footage was used for two purposes: to develop and test the coding system aimed at measuring engagement-related behavior, but also to synchronize the EDA signal and identify the activity windows for feature extraction. In both occasions, we used the video footage of the frontal video camera and resorted to the lateral video camera as a back-up, in case the frontal video camera did not work or objects occluded the visibility of the scene.

4.5.3 E4 wristband

We measured EDA with the E4 wristband (see Figure 7). The E4 wristband is a wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition (Garbarino et al., 2014). It has four sensors embedded in its case: (i) a *photoplethysmography sensor* (PPG) to measure blood volume pulse and derive HR, HRV, and inter-beat interval, (ii) a *triaxial accelerometer* to capture motion-based activity and detect movement patterns, (iii) an *infrared thermopile* to gauge peripheral skin temperature, and (iv) an *EDA response sensor* to measure the electrical conductance of the skin and, consequently, psychological arousal. As it was not possible to deactivate any of the sensors in the E4 wristband, during sessions we obtained data from all of them.

The E4 wristband was selected among the available wearable physiological sensors as it was lightweight and unobtrusive. Moreover, it was the only device measuring EDA that did not entail the positioning of the electrodes on the medial or distal phalanxes. This was of crucial importance as it left participants free to manipulate objects during the activities without jeopardizing the data collection.

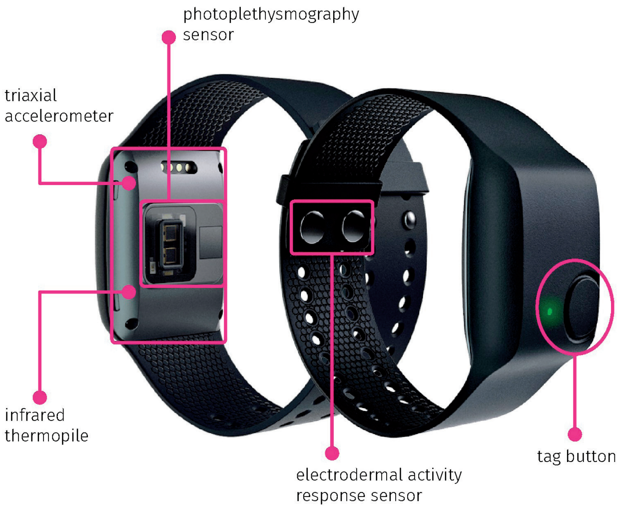


Figure 7. E4 wristband, overview of sensors (adapted from Empatica.com)

Before starting the study, we performed some tests to understand whether the E4 wristband could be accepted by participants and whether it was really as unobtrusive for them as we supposed. We asked few participants to wear the wristband (turned off) and tell us whether they found it uncomfortable and if they would wear it for longer periods. We explained them that the wristband was supposed to collect some physiological parameters, for instance their heart beat. The response to the sensor was optimal, none of the seniors complained about it. During data collection, we noticed that the participants acknowledged the presence of the sensor on their wrist at the beginning of the session. However, they forgot about it immediately after the baseline collection. In this sense, the baseline phase (i.e., listening to a fairytale, see sub-section 4.7) functioned as a distractor.

4.6 Setting

The data collection was performed directly in the nursing homes, in rooms that were usually allocated to recreational activities. As we collected data in two different contexts, a great deal of effort was put into making the locations of data collection resemble each other. A rectangular table was placed on one side of the room and the hand-held video cameras were arranged on top of mini-tripods and positioned one in front and one on the side of participants (see Figure 8). The frontal video camera was positioned on a small table, while the lateral video camera was either hidden on a library shelf or positioned on a desk.

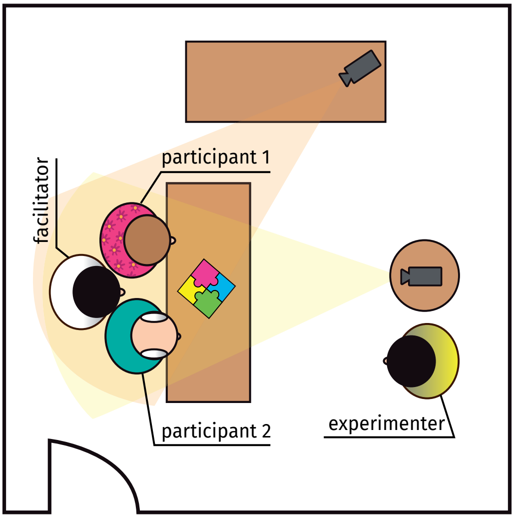


Figure 8. Activity room layout

During activities, the participants sat on the same side of the rectangular table while the facilitator stood up in-between them. The central positioning of the facilitator was meant not to influence the engagement state of the participants. Indeed, in the exploratory study (see Appendix B), we noticed that, when the facilitator was positioned on the side of the participants, it involuntarily spent more time closer to one of them. This affected engagement, as it had a negative influence on the engagement of the furthest participant and a positive influence on the engagement of the closest one.

With regards to the experimenter, s/he was always seated on the left of the frontal video camera. This way s/he could control that the main video camera was properly recording the session, communicate the facilitator the time progression of the session, and readily intervene in case of a malfunctioning of the sensors.

4.7 Procedure

The experimental sessions were divided into phases with different goals. The structuring of the session in phases was in part inspired by the usual subdivision of activities that we observed during the ethnographic study (see Appendix A), in part it was a consequence of the testing of the procedure of data collection in the exploratory study (see Appendix B). Last, it was due to the necessity to synchronize EDA with the video footage and obtain an EDA baseline. During sessions, there were six phases: (1) a *preparation phase* (10 minutes), (2) a *habituation phase* (5 minutes), (3) a *synchronization phase* (2 minutes), (4) a *baseline phase* (5 minutes), (5) an *activity phase* (20-25 minutes), and (6) the *end of the activity* (5 minutes).

Figure 9. Phases of experimental session: A) habituation phase, B) baseline phase, C) and D) activity phase for board games, E) and F) activity phase for robot play

1. During the *preparation phase*, the experimenter set up the room, while the facilitator helped participants to reach it. Once participants reached the room, the experimenter switched on the video cameras.

2. During the *habituation phase*, the experimenter conversed shortly with the participants, while they sat to recover from the effort of walking to reach the room, then s/he helped them to wear the E4 wristband (Figure 9A).

3. During the *synchronization phase*, the experimenter switched on the wristbands of both participants and simultaneously pushed the tag buttons on top of them. This way the E4 produced a timestamp of the particular moment when the tag button was pushed which was then used for synchronization.

4. During the *baseline phase*, the facilitator read a fairytale (5 minutes) to the participants to collect the baseline of physiological data (Figure 9B).

5. During the *activity phase* (20-25 minutes), the participants completed the three board games (Figure 9C and 9D) or interacted with Pleo (Figure 9E and 9F).

6. At the *end of activity,* the experimenter switched off the wristbands in front of the video cameras, removed them from the wrists of participants, and switched off the video cameras once the participants left the room. At this point, the participants were guided back to their units by the facilitator.

4.8 Ethical Approval

The study was conducted according to the Declaration of Helsinki and to Spanish laws number 159/2007 and 41/2002. An informed written consent was signed by all the legal guardians of participants. All participants were informed about the study and gave their personal consent to participate. Both the consent of the legal guardian and that of the participant were required to take part in the data collection.

4.9 Conclusions

In this chapter, we described the methods and procedures that allowed us to collect a large database of multimodal data while participants with dementia were involved in two playful activities: a game-based cognitive stimulation and a robot-based free play. The multimodal data that we collected were observational rating scales, video recordings, and physiological signals (for an overview of the database and missing data, see Figure 10). Observational rating scales replaced self-reports as a form of gold standard to judge the concurrent validity of the yet-to-develop objective measures of engagement. Video recording functioned as supports to develop and test the coding system of engagement-related behavior. Physiological signals were employed to attest whether EDA was an appropriate measure of engagement.

This chapter had an eminently methodological character. The next two chapters, instead, will have a more analytical hallmark. Within them, we will focus on the first research objective of the dissertation – explore and develop novel measures of engagement for people with dementia – and on replying to the following research questions: *Which peripheral-physiological measures can be employed to assess engagement in people with dementia? Are these peripheral-physiological measures appropriate to assess engagement?* And also *which behaviors externalize engagement in people with dementia? Are these behaviors appropriate to measure engagement?*

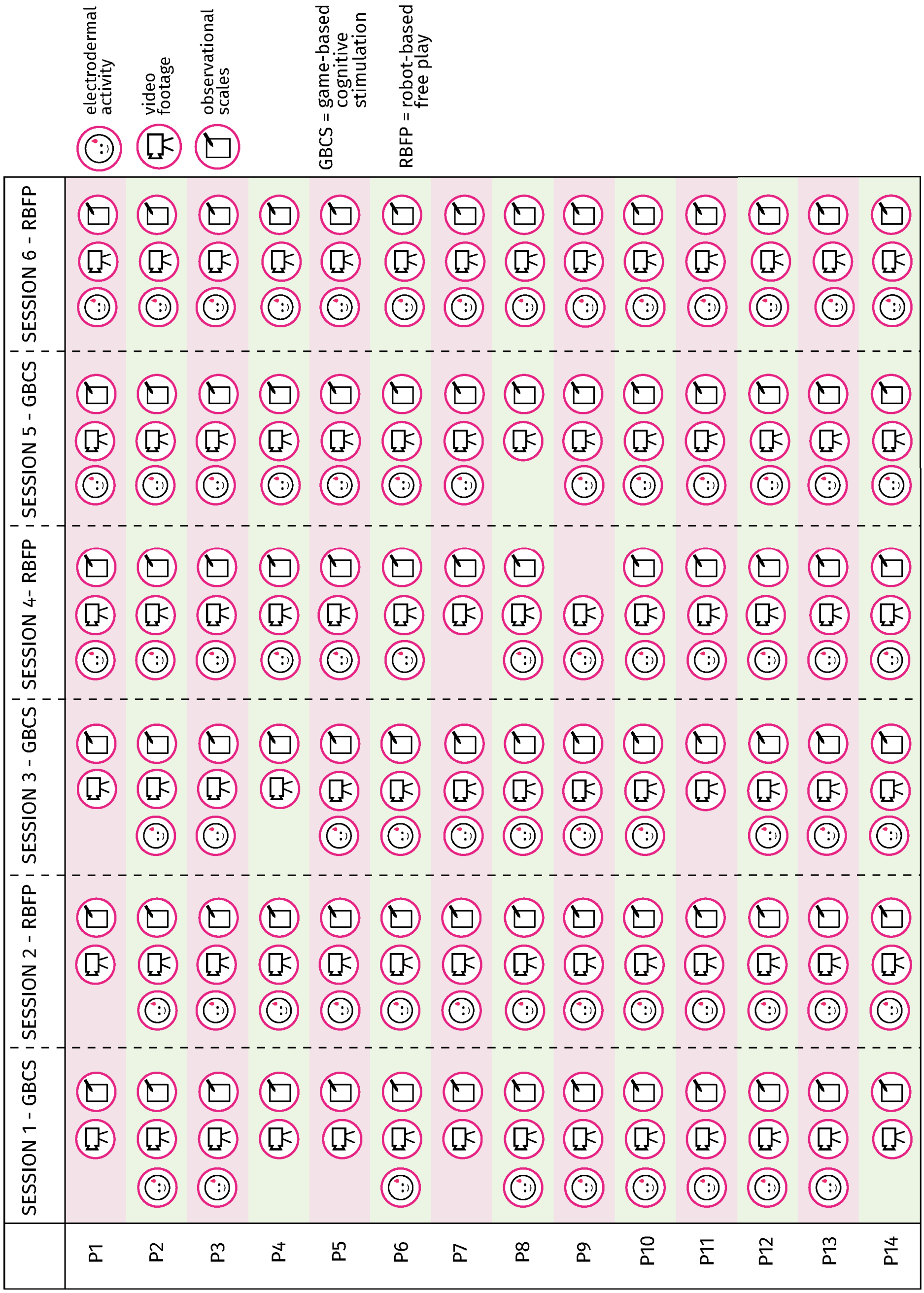


Figure 10. Datasets, overview of missing data (P1–P14= Participant 1 to Participant 14)

5

Measure Development: Electrodermal Activity

5.1 Introduction

In this chapter, we concentrate on the analysis of EDA. We first highlight the research gaps in the state of the art regarding the measurement of the psychophysiology of engagement in dementia. Then, we detail: (i) the rationale that guided us in the removal of improper measurements from the database, (ii) the process through which we synchronized and labeled the signal, (iii) the methods utilized to denoise the signal and extract features from it, and (iv) the testing of the concurrent validity of EDA as a measure of engagement. Last, at the end of the chapter, we discuss the behavior of the EDA signal in the two activities and pinpoint which features of EDA will be introduced in the theory-driven model of engagement.

5.2 Research Gap

The measurement of the psychophysiology of engagement is well documented in the literature. However, when it comes to people with dementia, it is only rarely attempted. This is partially due to the objective difficulty of obtaining sound and reliable physiological sensors' readings from people with dementia, but also to the fact that, in order to gather physiological data in this target group, scientists need to work outside the lab and develop dedicated protocols to collect data in the field.

Most of the studies on psychophysiology in dementia are aimed at evaluating the clinical benefits of robot interaction. For instance, in terms of reduction of stress levels (Wada and Shibata, 2008) or improvement of neuronal activity (Wada et al., 2005). These studies involve costly and invasive procedures, such as EEG and urinalysis. Furthermore, they follow pre-test/post-test designs and are hence more targeted to the assessment of the clinical consequences of engagement rather than to the measurement of its concomitant physiological correlates.

As engagement is the state of heightened attention and enjoyment experienced *during* activities, to capture its essence, we considered far preferable to collect physiology while the activity was taking place, rather than before and after it. Few sensing techniques are apt to do so. These are: (i) PPG sensors – which collect cardiac measures informative of stress states and autonomic activation, (ii) Facial EMG electrodes – which measure the electrical activity of the facial muscles activated during expressions of positive and negative affect, and (iii) EDA response sensors– which gauge electrical changes in the conductance of the skin related to sympathetic arousal. Among them, the most appropriate and less obtrusive measure of engagement is EDA. EDA is more reliable in terms of engagement assessment with respect to cardiac measures, and more straightforward and less intrusive to gauge with respect to EMG.

The biggest methodological issue in collecting and analyzing EDA in dementia is the lack of a sufficient knowledge basis. As the only two studies that measured EDA in people with dementia were targeted to behavioral and functional assessment (Treusch et al., 2015; Alam et al., 2016), we considered EDA studies conducted with children and healthy adults as the closest references for our work and thus used them as a benchmark.

The analysis of EDA was performed in four phases: (i) identification and exclusion of invalid measurements in the database, (ii) synchronization of the EDA signal with the video footage for the identification and labeling of the different phases of the activity, (iii) pre-processing of the signal and feature extraction from the activity windows, and (iv) testing of the concurrent validity of EDA as a measure of engagement.

5.3 Identification and Exclusion of Invalid Measurements

In order to spot invalid sessions in the database, we used a twofold approach. First, we wrote diaries of the experimental sessions. Second, we plotted the EDA signals gathered during activities. With regards to the former approach, for each experimental session, we noted down: (a) the date of the session, (b) the type of activity (game-based cognitive stimulation or robot-based free play), (c) the initials of the participants taking part in the session, (d) the identification number of the wristband that each of them used, (e) the identification number of the session, and (f) the main issues encountered during data collection. For instance, problems in switching on the wristband, errors in pressing the tag button, multiple tags, malfunctions of the wristband, or manipulations of the wristband by participants. This way, we had a preliminary idea of which sessions could be invalid, and for which reasons. With regards to the second approach, we used Matlab[4] to plot the signals collected in each session and visually examine each of them. Artifacts in the EDA signal are easy to recognize, as they are evident alterations of the normal conduct of the signal. For instance, they contain samples with null values (i.e., zero) or sharp square-wave spikes (see Figure 11). Based on this double rationale, we excluded 11 sessions from the initial database. The final database of EDA was composed of 73 sessions of data, 33 of game-based cognitive stimulation and 40 of robot-based free play.

4 https://nl.mathworks.com/products/matlab.html

As a side note, we would like to notify that some of the sessions of EDA were recorded on the dominant wrist (13 sessions out of the 73 that were valid). This was mostly due to problems encountered in measuring the signal on the non-dominant side, for instance, bruises caused by dialysis. EDA might vary between sides with small shifts in amplitude. However, according to the literature, EDA asymmetry is a phenomenon mostly observable in long-term measurement (hours up to days; Schulter and Papousek, 1992; Picard et al., 2016). As in our study the measurement of EDA was carried out for a short period of time, we did not consider the sessions collected from the dominant wrist as invalid.

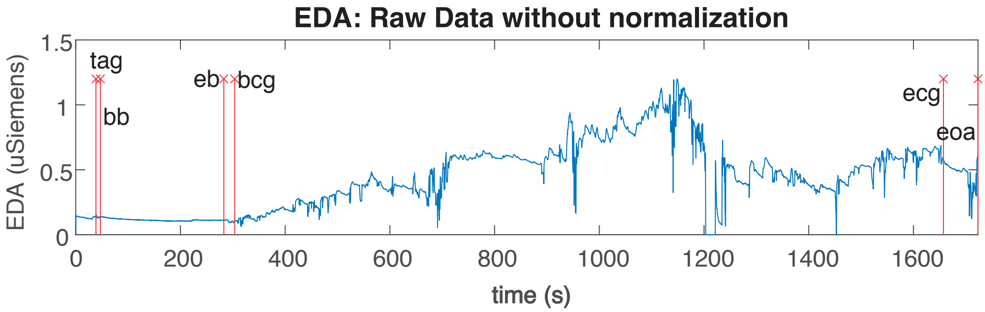


Figure 11. Example of invalid EDA signal. The vertical red lines indicate the different phases of the activity: tag, bb: beginning baseline, eb: end baseline, bcg: beginning cognitive games, ecg: end cognitive games, eoa: end of activity

5.4 Synchronization and Labeling

As a second step in the analysis, the valid EDA signals were synchronized with the video-footage. This was to identify and label the different phases of the activity and extract features from the relevant windows. As described in chapter 4, each experimental session was composed of six phases: (1) preparation, (2) habituation, (3) synchronization, (4) baseline, (5) activity, and (6) end of the activity. The E4 wristband was switched on at the end of the habituation phase and switched off at the end of the activity. Thus, in order to proceed to the extraction of the EDA features, we first needed to align the timestamp of the tag with the corresponding moment in the video. Then, we needed to identify the boundaries of the baseline and activity phases. We did this with the software Kinovea[5]. Using the video as a support, we labeled the moment when the experimenter pushed the tag buttons on the wristbands at the beginning of the session and the moment when s/he switched off the wristband at the end of the session. Moreover, we labeled the beginning and the end of the baseline phase and the beginning and the end of the activity phase. The Kinovea annotation file was then imported in Matlab where the exact frame of the tag moment in

5 https://www.kinovea.org/

the video was paired with the timestamp in the signal (see Figure 12) and the video timeline was synchronized with the signal timeline. We employed the triaxial accelerometer signal to verify that the synchronization was performed in a correct way. With respect to the EDA signal, variations in the accelerometer signal are much easier to connect with events in the videos (e.g., raising arms). The so obtained Matlab synchronization file was then used for features extraction.

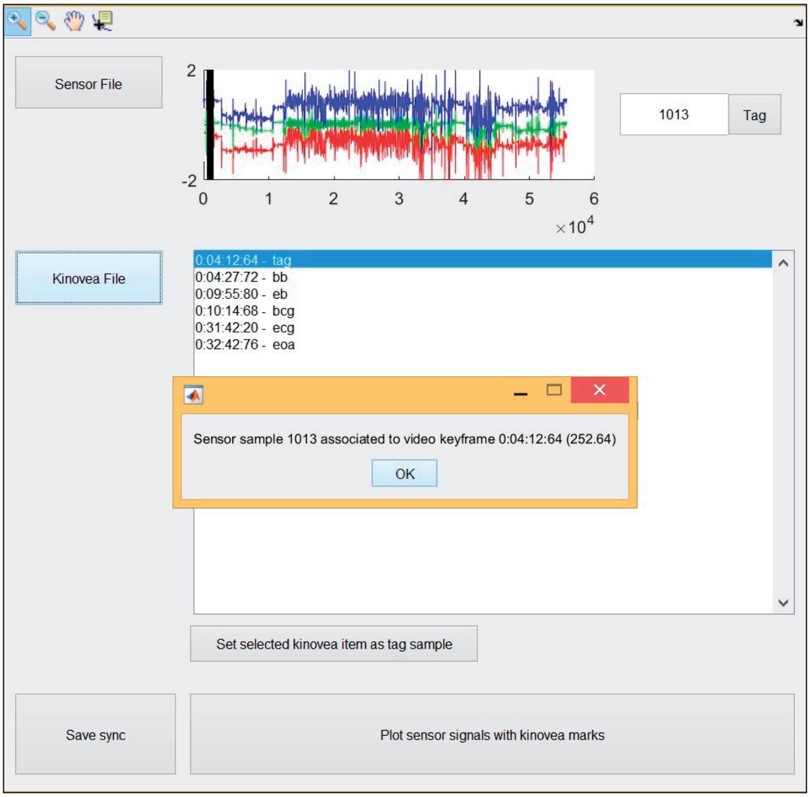


Figure 12. Data synchronization in Matlab

5.5 Pre-processing and Feature Extraction

EDA features were extracted from two windows: the *baseline phase – w(1)–* and the *activity phase – w*. These were respectively described by the following labels in the Matlab synchronization file: (1) beginning and end of the baseline phase (bb= beginning baseline, eb= end baseline) and (2) beginning and end of the activity phase (bcg= beginning cognitive games, ecg= end cognitive games; bri= beginning robot interaction, eri= end robot interaction). Before the extraction of features, the EDA signal was normalized and

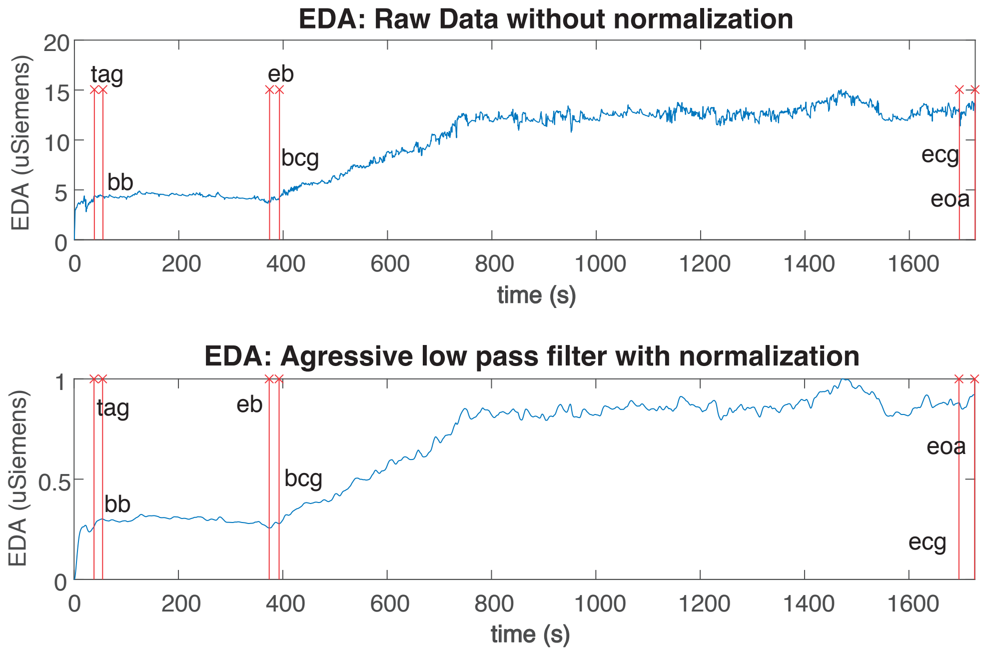


Figure 13. Filtering of EDA signal. In the top plot, the signal without normalization and filtering. In the bottom plot, the signal normalized and denoised

denoised with a second order Butterworth low-pass filter with a cutoff frequency of 0.05 Hz (see Figure 13). The shorter windows between the start of recording and the beginning of the baseline phase (bb, see Figure 13), between the end of the baseline phase (eb) and the beginning of the activity phase (bcg), and between the end of the activity phase (ecg) and the end of the session (eoa) were excluded from the analysis.

The set of features to extract was compiled based on previous literature (Henriques et al., 2013; Leite et al., 2013). In order to take into account the baseline state of the person with dementia in the calculation of the EDA, the values of the features extracted during baseline were subtracted from those of the features extracted during the activity phase. The feature notation in Table 7 was constructed in the following way. The set of samples was recorded in a window of time defined by the beginning and the end of an activity (game-based cognitive stimulation or robot-based free play). The Short Fast Fourier Transform of this sample set was formed by $S_1^W, \dots S_N^W$ through (1):

$$S_h^W = \sum_{n=1}^{N} s_n e^{-i2\pi h \frac{n}{N}} \quad (1)$$

where $h = 1, \dots, N$. S_h^W. is a set of N complex numbers that represents the amplitude and phase of a harmonic. With regards to *Npeaks*, we denoted it as the number of significant local maxima found in *SW*. *NPRW* is defined as *Npeaks*(*SW*) divided by the length of *SW*.

Table 7. Set of EDA features and corresponding equations

FEATURE	EQUATION
SMA EDA[a]	$\sum_{i=1}^{T} s_i^W dt - \sum_{i=1}^{T} s_i^{W(1)} dt$
MEAN EDA[b]	$s^W - \bar{s}^{W(1)}, where\ \bar{s}^W = \frac{1}{N}\sum_{i=1}^{N} s_i^w$
STD EDA[c]	$\sigma s^W - \sigma s^{W(1)}, where\ \sigma s^W = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(s_i^W - \bar{s}^W\right)^2}$
RNG EDA[d]	$Rng(s^W) - Rng\left(s^{W(1)}\right), where\ Rng(s^W) = max(s^W) - min(s^W)$
SUM H EDA[e]	$\sum_{i=1}^{N} S_{\mathrm{i}}^W$
NPR EDA[f]	$NPR^W - NPR^{W(1)}$
KURT EDA[g]	$\delta_s^W - \delta_s^{W(1)}, where\ \delta_z^W = \frac{\mathrm{E}\left[\left(\lvert S_{1,N}^W\rvert_i - \lvert \bar{S}_{1,N}^W\rvert\right)^4\right]}{\left(\mathrm{E}\left[\left(\lvert S_{1,N}^W\rvert_i - \lvert \bar{S}_{1,N}^W\rvert\right)^3\right]\right)^2}$
SKEW EDA[h]	$\gamma_s^W - \gamma_s^{w(1)}, where\ \gamma_s^W = \mathrm{E}\left[\left(\frac{\lvert S_{1,N}^W\rvert_i - \lvert \bar{S}_{1,N}^W\rvert}{\sigma\left(\lvert S_{1,\mathrm{N}}^W\rvert\right)}\right)^3\right]$

a. SMA EDA= difference of signal magnitude area of EDA in ω and ω(1)
b. MEAN EDA= difference of mean of EDA in ω and ω(1)
c. STD EDA= difference of standard deviation of EDA in ω and ω(1)
d. RNG EDA= difference of range of EDA in ω and ω(1)
e. SUM H EDA= difference of summation of harmonics of EDA in ω and ω(1)
f. NPR EDA= difference of number of peaks ratio of EDA in ω and ω(1)
g. KURT EDA= difference of kurtosis of EDA in ω and ω(1)
h. SKEW EDA= difference of skewness of EDA in ω and ω(1)

5.6 Concurrent Validity

Concurrent validity is the extent to which the scores of a new measure are related to the scores of a criterion measure administered at the same time and is tested with a correlation statistic (Mislevy and Rupp, 2010). In order to determine the concurrent validity of EDA, we extracted the EDA features from all the valid sessions in the database. Then, we performed a Spearman rank correlation (one-tailed, listwise exclusion of cases, N_{GBCS}= 33; N_{RBFP}= 40) between the features of EDA and the items of the OME and OERS (see Tables 8 and 9). In this context, it is important to underline that: (1) EDA and the OME and OERS refer to different response systems, respectively the peripheral-physiological and the behavioral/ expressive, (2) the OME and OERS *assess focused attention, rapport, and valence* while EDA measures arousal, (3) *arousal* changes as a result of *focused attention, rapport*, and *valence*, but it is not a direct measure of *focused attention, rapport*, and *valence*, and (4) EDA

Table 8. Concurrent validity of EDA for game-based cognitive stimulation: correlation tested (in bold the significant ones: *≤.05, **<.01, *<.001)**

GAME-BASED COGNITIVE STIMULATION									
CORRELATION OME & OERS/ EDA*		SMA EDA	MEAN EDA	STD EDA	RNG EDA	SUMM H EDA	NPR EDA	KURT EDA	SKEW EDA
Attention	*r*(s)	.007	.080	.008	.016	.104	-.063	-.177	-.161
	p-value	.485	.329	.483	.464	.283	.363	.162	.186
Attitude game	*r*(s)	.025	-.151	-.017	.097	-.032	-.263	-.054	-.059
	p-value	.445	.201	.463	.295	.430	.070	.382	.371
Attitude partner	*r*(s)	.021	-.029	.038	.083	.045	-.086	-.141	-.158
	p-value	.454	.437	.417	.324	.402	.317	.216	.190
Cognitive difficulty	*r*(s)	.277	-.140	-.277	-.210	-.168	.145	-.268	-.287
	p-value	.060	.219	.059	.120	.176	.211	.065	.053
Pleasure	*r*(s)	.159	.079	.221	**.291**	.128	.041	-.062	-.077
	p-value	.189	.331	.109	***.050**	.240	.411	.366	.335
Anger	*r*(s)	.074	.111	.241	.186	.037	.223	-.056	-.074
	p-value	.341	.269	.088	.150	.419	.106	.379	.341
Alertness	*r*(s)	-.080	-.018	-.089	-.036	-.080	-.195	**.302**	**.302**
	p-value	.329	.461	.312	.422	.329	.138	***.044**	***.044**

*The items *anxiety/fear* and *sadness* of the OERS did not yield any result, hence they do not appear in the list

Table 9. Concurrent validity of EDA for robot-based free play: correlations tested (in bold the significant ones: *≤.05, **<.01, *<.001)**

ROBOT-BASED FREE PLAY									
CORRELATION OME & OERS/ EDA*		SMA EDA	MEAN EDA	STD EDA	RNG EDA	SUMM H EDA	NPR EDA	KURT EDA	SKEW EDA
Attention	*r*(s)	.005	-.054	-.108	-.179	-.064	-.071	**-.272**	**-.293**
	p-value	.487	.373	.256	.137	.350	.334	***.047**	***.035**
Attitude game	*r*(s)	**.304**	-.125	-.178	-.117	-.004	.234	**-.272**	**-.281**
	p-value	***.030**	.224	.139	.239	.490	.075	***.047**	***.042**
Attitude partner	*r*(s)	.155	**-.419**	**-.300**	-.265	-.236	.064	-.134	-.161
	p-value	.174	****.004**	***.032**	.052	.074	.350	.207	.163
Pleasure	*r*(s)	**.363**	.153	.039	.051	.177	**.306**	-.214	-.220
	p-value	***.012**	.176	.408	.379	.141	***.029**	.096	.089
Anger	*r*(s)	-.012	**-.303**	-.240	-.190	-.232	.168	-.049	.001
	p-value	.471	***.031**	.071	.124	.078	.153	.384	.497
Anxiety/Fear	*r*(s)	-.014	-.067	.117	.091	.005	.090	**.278**	.236
	p-value	.466	.343	.238	.291	.488	.292	***.044**	.074
Sadness	*r*(s)	.084	**-.341**	-.227	-.136	-.207	.169	-.247	-.187
	p-value	.305	***.017**	.083	.204	.103	.152	.065	.127
Alertness	*r*(s)	.176	.018	-.055	-.061	.123	-.034	.017	-.022
	p-value	.142	.458	.369	.356	.228	.419	.459	.447

*The item *cognitive difficulty* of the OME did not yield any result, hence it does not appear in the list is a continuous measure, while the OME and OERS are ordinal scales. In the game-based cognitive stimulation:

1. The item *pleasure* of the OERS was positively correlated with RNG EDA ($rs(31)= .291$, $p= .050$).

2. The item *alertness* of the OERS was positively correlated with KURT EDA ($rs(31)= .302$, $p= .044$) and SKEW EDA ($rs(31)= .302$, $p= .044$).

In the robot-based free play:

1. The item *attention* of the OME was negatively correlated with KURT EDA ($rs(38)= -.272$, $p= .047$) and SKEW EDA ($rs(38)= -.293$, $p= .035$).

2. The item *attitude toward the game* of the OME was positively correlated with SMA EDA ($rs(38)= .304$, $p= .030$) and negatively correlated with KURT EDA ($rs(38)= -.272$, $p= .047$) and SKEW EDA ($rs(38)= -.281$, $p= .042$).

3. The item *attitude towards the partner* of the OME was negatively correlated with MEAN EDA ($rs(38)= -.419$, $p= .004$) and STD EDA ($rs(38)= -.300$, $p= .032$).

4. The item *pleasure* of the OERS was positively correlated with SMA EDA ($rs(38)= .363$, $p= .012$) and NPR EDA ($rs(38)= .306$, $p= .029$).

We found three additional significant correlations in the robot-based free play. As these involved the items *anger*, *anxiety/fear*, and *sadness*, we excluded them due to the scarce variability of these items throughout the whole dataset.

5.7 Discussion

5.7.1 Research Questions

The analysis of EDA enabled us to reply to the research questions: *which peripheral-physiological measures can we employ to assess engagement in people with dementia? Are these peripheral-physiological measures appropriate to assess engagement?* The analysis of concurrent validity revealed a weak to adequate strength of correlation between the

Chapter 5

features of EDA and the items of the OME and OERS. The number of correlations steeply diverged across activities. In the game-based cognitive stimulation, the number of correlations was rather modest and quite weak in strength, while in the robot-based free play, the correlations were higher in number and more robust. Given the premises we made, we did not consider this difference in the results alarming. In this consideration, we took into account also the fact that the sample size of game-based cognitive stimulation was much lower than the one of robot-based free play and that there was much less variability in the items of the OME and OERS collected during game-based cognitive stimulation with respect to those collected during robot-based free play.

The assessment of EDA also satisfied one of the requirements for a proper measurement framework of engagement in dementia: *Explore the measurement of the physiology of engagement in the field with unobtrusive sensors*. EDA brings about several advantages to the measurement of engagement: (i) it gives a quick and easy to understand overview of the arousal levels during the activity, (ii) it is unobtrusive to measure, and (iii) it deepens the knowledge drawn from behavior with another layer of meaning (see Figure 14).

5.7.2 Dynamics of Psychophysiology

The correlations with the OME and OERS did not just provide information on the concurrent validity of EDA, but also disclosed some important dynamics in the behavior of the signal during activities. For instance, during game-based cognitive stimulation, KURT EDA and SKEW EDA increased with the increase of alertness. Conversely, in robot-based free play, they decreased with the increase of attention. Albeit provided with different shades of meaning, attention and alertness are considered gold standard measures of the same component of engagement: focused attention. What these correlations tell us is that the EDA elicited by focused attention differs across activities in terms of frequency distribution. In the game-based cognitive stimulation, focused attention (i.e., alertness) is characterized by an EDA with a lower frequency (higher skewness) and a more defined peak of the frequency distribution (higher kurtosis). In the robot-based free play, instead, it (i.e., attention) is characterized by an EDA with a higher frequency (lower skewness) and a less defined peak of the frequency distribution (lower kurtosis). These differences in the physiological correlates of focused attention might be due to the different types of attention elicited by the two activities: one strictly cognitive, the other more social and affective.

The changes in EDA seemed to capture also the affective component of engagement, especially in robot-based free play. Indeed, both the tonic (i.e., the slower acting components of the EDA signal) and the phasic components of EDA (i.e., the faster changing peaks of the EDA signal, Braithwaite et al., 2013) increased as a result of positively-valenced affective states. In the game-based cognitive stimulation, the variation mainly involved the tonic component of EDA, while in the robot-based free play, the phasic component seemed to have a more significant role. With regards to rapport, we found significant relationships with the gold standard just in the robot-based free play. In this case, the betterment of the attitude of participants towards their partners was accompanied by a reduction in the tonic EDA. This result is somewhat counterintuitive. However, it can be explained with the fact that, in the robot-based free play, the participants focused on their partners specifically when these were interacting with the robot (see Figure 14). As the arousing element of the activity was the robot, when this was brought afar, the slope of EDA went down.

5.7.3 Selection of EDA Features

As a result of these analyses and of their discussion, we decided to exclude KURT EDA and SKEW EDA from the list of EDA features to be included in the theory-driven model of engagement. Indeed, albeit KURT EDA and SKEW EDA appear to have a high discriminating power – as they can differentiate between different types of attention – they cannot be used to measure engagement across activities, as their results are extremely conflicting. The remaining EDA features were divided into two categories: tonic EDA and phasic EDA (see Table 10).

Table 10. EDA features included in the theory–driven model of engagement

ENGAGEMENT-RELATED PHYSIOLOGY		
AROUSAL		
TONIC EDA	**PHASIC EDA**	**EXCLUDED: FREQ. DISTRIBUTION**
MEAN EDA	SMA EDA	KURT EDA
STD EDA	NPR EDA	SKEW EDA
RNG EDA		
SUM H EDA		

As a side note to these results, we would like to underline that, albeit significantly correlated with most of the items of the OME and OERS, EDA is not a direct measure of the observable engagement components which are measured by the OME and OERS, but a measure of a physiological component – arousal – which is supposed to change as a result of focused attention, valence, and rapport.

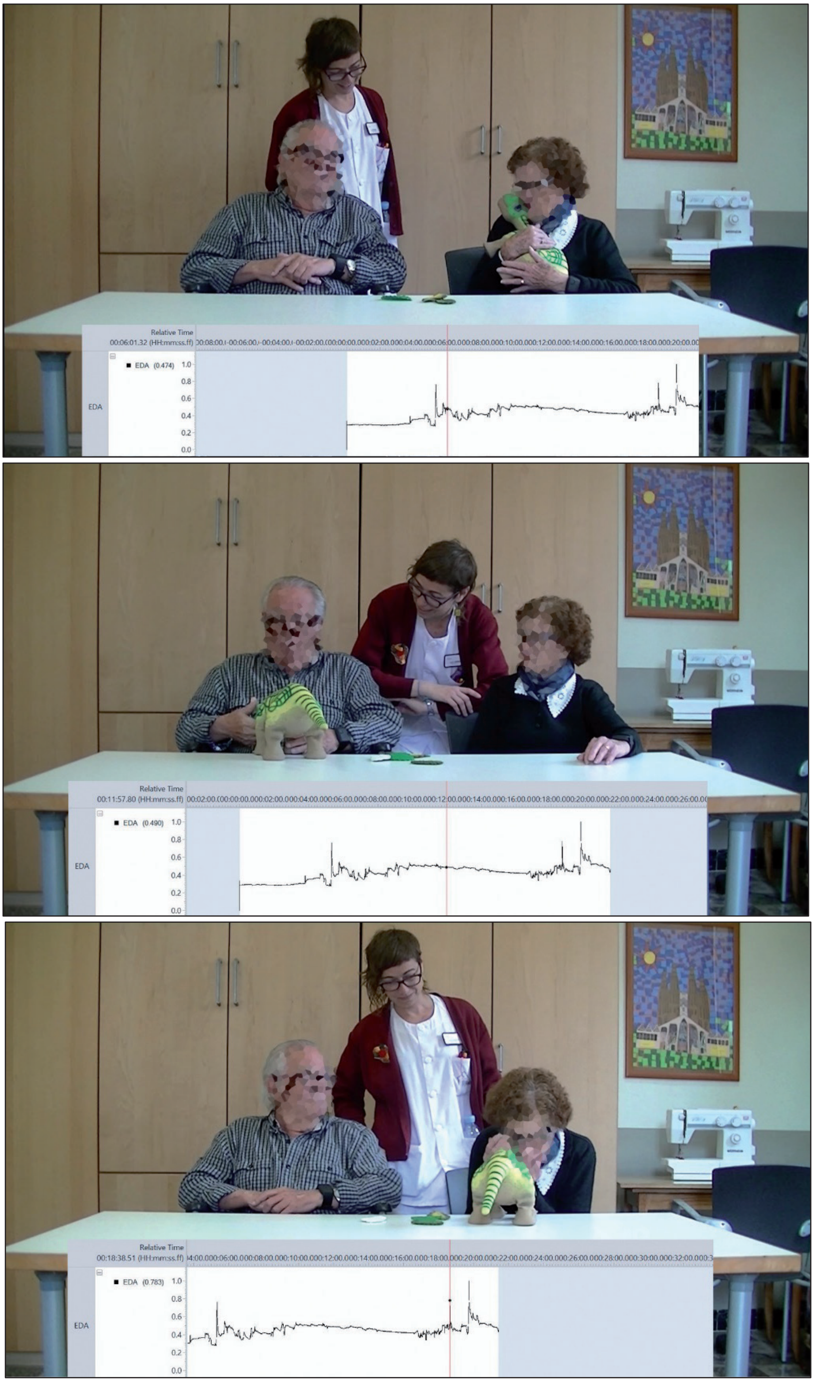


Figure 14. Chronology of an EDA session. Progression of EDA of the participant on the right during a session of robot-based free play. To read from left to right, and from top to bottom

5.8 Conclusions

This chapter focused on the analysis of EDA and on its adequacy as a measure of engagement. It first outlined the process of synchronization and labeling of the EDA signal. Then, it discussed the pre-processing of the signal and detailed the methods of feature extraction. Last, it presented and discussed the results of concurrent validity.

The exploration of EDA as a measure of engagement allowed us to satisfy one of the requirements of this dissertation – explore the measurement of the psychophysiology of engagement with unobtrusive sensors – and deploy a set of engagement-correlated EDA features to include in the theory-driven model of engagement. Also, it revealed that the measurement of engagement on a peripheral-physiological level is as feasible in people with dementia as it is in healthy subjects. In the next chapter, we concentrate on yet another level of measurement – the behavioral/expressive – and develop a coding system to measure engagement-related behavior in people with dementia. The outcomes of the present chapter and of the next one will converge into the definitive version of the theory-driven model of engagement.

6

Measure Development: A Coding System of Engagement-related Behavior

6.1 Introduction

This chapter focuses on the development of the coding system aimed at measuring engagement-related behavior and on the testing of its inter-rater reliability (IRR) and concurrent validity. In the first part of the chapter, we outline the methodological pitfalls of the existing behavioral/expressive assessment techniques of engagement for dementia and identify a mixed approach to address them. In the second part of the chapter, we explain the process of development of the coding system from the observation of the videos in the database to their annotation. In the third part of the chapter, we present the results of the IRR and concurrent validity tests performed on the coding system. Last, at the end of the chapter, we discuss these results and define which behaviors of the coding system will be employed to assess focused attention, rapport, and valence in the theory-driven model of engagement.

6.2 Methodological Issue

From the review of the literature on the behavioral/expressive measures of engagement, it emerged that the measurement of engagement through behavior in dementia, albeit very extensive, is far from providing conclusive results. The available assessment techniques either break down behavior into many small pieces – each given a specific engagement meaning – or assemble behaviors into broad engagement classes measured along categorical scales. Either way, they never describe behavior in its simultaneous multimodality, as a configuration of actions happening at the same time.

One of the requirements to properly measure engagement through behavior in dementia is to develop a measurement system that describes behavior in its parts, but also as a whole-body dynamic. In order to pursue this objective, we adopted two different approaches to the description of behavior – Ethography and Laban-Movement Analysis (LMA; Laban, 1966) – and combined them with the intelligent use of the available software for video annotation.

Ethography refers to the description and assessment of behavior through ethograms. As an approach, it stems from Ethology and aims at measuring behavior through direct observation, rigorous description, and objective analysis. LMA, instead, is a holistic framework that provides a vocabulary to describe and interpret movement (Bartenieff and Lewis, 2002). As an approach, it is more focused on how movement naturally occurs, how it relates to the outer environment, and what it expresses.

LMA is organized into four main categories, each describing a different facet of movement: *body* – the elements of the body structure and their connection in movement (Maletic, 1987), *space* – the mapping of the body structure in relation to the environment, *effort* – the qualities of movement (Bradley, 2008), and *shape* – the changes in the body shape that express attitudes towards the environment (Wile and Cook, 2010). For the sake of this research, we mainly focused on the category shape, this was the most appropriate category to describe engagement-related behavior as it provided a reference framework describing how we as humans shape our bodies to interact with the environment. There are three distinctions in the category shape, also referred to as *modes of shape change*: *shape flow* – changes in the shape of the body in relation to the self, *directional shape* – goal-oriented changes of the body shape in relation to the others and the environment, and *shaping* – molding and carving of the body in interaction with the others and the environment.

We employed: (i) Ethography to describe behavior in its minute parts, (ii) LMA – in the formalization of the modes of shape change – to identify the underlying structure of behavior and create a coding system workable across activities, and (iii) Observer XT to measure behavior in its parts, while at the same time giving an overview of its simultaneous multimodal arrangement.

6.3 Construction of the Coding System

The coding system of engagement-related behavior was built in 18 months by a multidisciplinary research team which included a certified movement analyst (CMA). The development of the coding system consisted of two phases: description and structuring. In the *descriptive* phase, we adopted an ethographic approach similar to that of Jones et al. (2015), Olsen et al. (2016), and Jøranson et al. (2016) and developed two ethograms, one per activity. In the structuring phase, we employed LMA to sort out the complexity of the two ethograms by identifying *shape* commonalities in the behaviors displayed in the two activities and combine them into a unique coding system.

6.3.1 Note on the Video Recordings

To develop the coding system of engagement, the original videos were cut from the beginning to the end of the activity phase. We considered the moment when the facilitator placed the first board game or Pleo on the table in front of the participants as the *beginning of the activity phase* and the moment when s/he removed the last board game or Pleo from the table as the *end of the activity phase*. The database of videos was composed of 42 sessions of play of the duration of 20–25min (~ 17.5 h of video footage). The database of videos was split in two parts. Thirty videos were used to develop the ethograms and structure them in a coding system (15 videos per activity type). Twelve videos (six videos per activity type) were used to test the IRR of the coding system.

6.3.2 Ethographic Approach: Development of the Ethograms

In order to develop the two ethograms, we first watched the 30 videos allocated to the construction of the coding system (30 sessions: 15 of game-based cognitive stimulation, 15 of robot-based free play). Then, we described each video in a separate file by detailing the main events and behaviors in chronological order (Martin and Bateson, 2007). Last, we watched each video at a slow speed, stopped it whenever we identified a relevant micro-behavior, and gave a name and an operational description to each micro-behavior. Dautenhahn and Werry (2002) defined *micro-behaviors* as well-identifiable low-level action-movement-oriented behaviors recognizable by computational systems.

Before proceeding to the structuring phase, we polished the two ethograms by removing those micro-behaviors that had an ambiguous meaning. We removed *manipulators* (e.g., scratching the chin or the scalp), *adjustments* (e.g., adjusting spectacles, watch, bracelets, earrings, clothes; see participant on the left in Figure 15), and *vocalizations* (e.g., sighing and singing). These micro-behaviors occurred several times in the sessions with contrasting meanings. The lack of a univocal meaning made it hard to figure out the contribution of these micro-behaviors to the assessment of engagement and brought us to their exclusion from the ethograms.

Figure 15. Example of adjustment: the participant on the left is adjusting her spectacles after they fell. In this case, the adjustment does not have any ulterior meaning. However, in some cases, it might signal a sense of discomfort of the participant

We also excluded *verbal behavior* from the ethograms. This decision was made with our target group in mind. Although the participants to this study had their language skills still preserved, most people with dementia do not (Thompson, 1987; Klimova and Kuca, 2016). In order for the coding system to be generalizable to other dementia groups (i.e., people with severe dementia), it is of crucial importance to make meaning of engagement without being dependent on language production.

As a last step, we discarded the *head gestures* nodding, negation, and head protruding. Also, we ruled out *co-speech gestures* (i.e., hand and arm movements that accompany spoken language such as saying no with the index finger of the hand). These were hard to relate to a specific element in the activity without resorting to language.

6.3.3 Laban–Inspired Approach: Structuring of the Ethograms

As a first step in the structuring of the two ethograms into a unique coding system, we stated the *body portion involvement*. We did so by making reference to the micro-behaviors in the two ethograms. In LMA, the body portion involvement refers to which body parts are activated during movement (Hackney, 2002). The body parts involved in a movement might be *body areas* (i.e., head, torso, chest and pelvis), *limbs* (i.e., arms, hands, legs, feet), *joints* (e.g., shoulders, elbow, wrists), and *body quadrants* (i.e., right upper, left upper, right lower, left lower). The micro-behaviors in the ethograms involved two body areas – *head* and *torso* – and two limbs – *arms* and *hands*. Given that there were no specific behaviors in the ethograms involving exclusively the arms or the hands, we decided to group the two limbs in a single category. As a result, the body portion involvement of both the ethograms consisted of three body parts: head, torso, and arms/hands. We grouped the micro-behaviors in the two ethograms according to the *body part* – or behavioral modality – they involved.

As a second step of the structuring, we identified those micro-behaviors expressing a *directional shape* of the body parts (i.e., head, torso, and arms/hands) and organized them based on their target in space. Taking the perspective of one of the participants, we identified five *foci* of the activities: the partner, the facilitator, the experimenter, the game (i.e., the board games or Pleo), and none of them.

As a concluding step in the structuring of the ethograms, we studied the remaining micro-behaviors. We noticed that some of them expressed a directional shape with *shaping* support, while others a *shape flow* occurring simultaneously with a directional shape. These micro-behaviors could be seen as traits superimposed on or enriching the directional shape behaviors. They carried an additional item of meaning in terms of engagement. Indeed, they all described the affective attitude of the participant – either positive or negative – towards the foci of the activity. We grouped these micro-behaviors according to their meaning in terms of affect: positive or negative. The affective micro-behaviors involving the head were grouped under the label *gestural support*, those involving the torso were nested under the label *postural support*, and, last, the affective micro-behaviors of the arms/hands under the label *quality of gesture*. The result of the structuring phase is presented in Table 11.

6.4 The Coding System

In this section, we describe how we transposed the structure of the ethograms (*body portion involvement*, *directional shape*, *shape flow*, and *shaping*) back to a coding scheme that could be scored using the software Noldus Observer XT 10.55[6] (see Table 11).

Observer XT gives the possibility to define clusters of behaviors, also called behavior groups. In our case, we defined three behavior groups corresponding to the body parts involved in the activities*: head*, *torso*, and *arms/hands*. The behaviors in each group were the *directional shape* micro-behaviors that shared the same focus in the activity (the partner, the facilitator/experimenter, the game, or none of them).

In order to include in the coding system also the *shape flow* and *shaping* micro-behaviors expressing affect, we resorted to one of the features of Observer XT: the possibility to add modifiers to the behaviors in the behavior groups. *Modifiers* are specifications regarding a behavior that describe it more precisely or limit its scope. In this case, they specified whether the scored behaviors had a positive, neutral, or negative affective nuance (e.g., stroking the robot) or appeared together with other behaviors involving the same body part expressing a positive, neutral, or negative affective nuance (e.g., gaze toward game and smile).

The final coding system obtained through this process – which we called the *Ethographic and Laban-Inspired Coding System of Engagement* (ELICSE) – is thoroughly described in the Tables E1, E2, and E3 in Appendix E and in the screenshot in Figure 16.

6 https://www.noldus.com/human-behavior-research/products/the-observer-xt

Table 11. The coding system (for the operational descriptions, see Appendix E)

HEAD BEHAVIORS (directional)	**MODIFIERS (shape flow and shaping) – gestural support**
GAZE TOWARD PARTNER (GP)	*positive* gestural support (GP_pos)
	no gestural support (GP_no)
	negative gestural support (GP_neg)
GAZE TOWARD FACILITATOR/ EXPERIMENTER (GFE)	*positive* gestural support (GFE_pos)
	no gestural support (GFE_no)
	negative gestural support (GFE_neg)
GAZE TOWARD GAME (GG)	*positive* gestural support (GG_pos)
	no gestural support (GG_no)
	negative gestural support (GG_neg)
NONE OF THE TARGET HEAD MOVEMENTS (NoH)	*positive* gestural support (NoH_pos)
	no gestural support (NoH_no)
	negative gestural support (NoH_neg)
TORSO BEHAVIORS (directional)	**MODIFIERS (shape flow and shaping) – postural support**
LEAN IN PARTNER (LIP)	*positive* postural support (LIP_pos)
	no postural support (LIP_no)
	negative postural support (LIP_neg)
NEAR REACH/LEAN TOWARD THE GAME (NRLTG)	*positive* postural support (NRLTG_pos)
	no postural support (NRLTG_no)
	negative postural support (NRLTG_neg)
NONE OF THE TARGET TORSO MOVEMENTS (NoT)	(none)
ARMS/HANDS BEHAVIORS (directional)	**MODIFIERS (shape flow and shaping) – quality of gesture**
REACH OUT PARTNER (RoP)	*positive* quality of gesture (RoP_pos)
	no quality of gesture (RoP_no)
	negative quality of gesture (RoP_neg)
REACH OUT FACILITATOR/EXPERIMENTER (RoFE)	*positive* quality of gesture (RoFE_pos)
	no quality of gesture (RoFE_no)
	negative quality of gesture (RoFE_neg)
MANIPULATE GAME (MG)	*positive* quality of gesture (MG_pos)
	no quality of gesture (MG_no)
	negative quality of gesture (MG_neg)
POSITIVE SIGNS OF AFFECTION INVOLVING ARMS/HANDS (SOA_pos)	(none)
NEGATIVE SIGNS OF AFFECTION INVOLVING ARMS/HANDS (SOA_neg)	(none)
NONE OF THE TARGET ARMS/HANDS MOVEMENTS (NoAH)	(none)

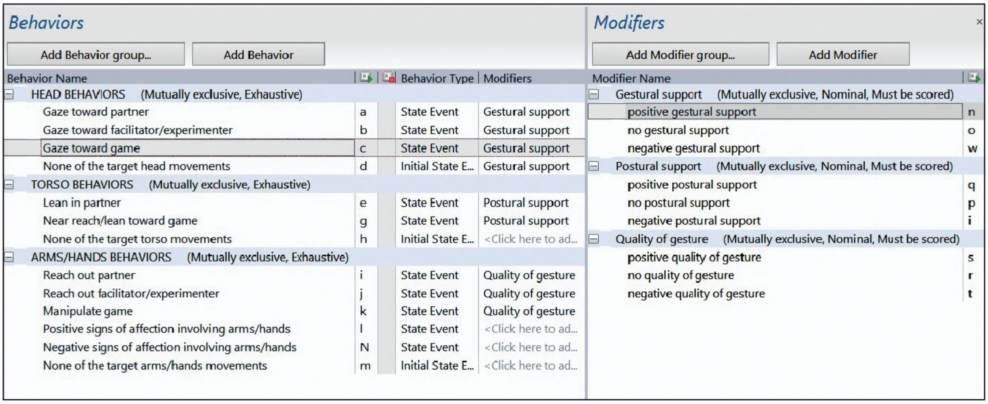


Figure 16. Coding system in Observer XT

6.4.1 Advantages of the Coding System

The behaviors in each behavior group were scored as mutually exclusive with a continuous sampling technique. This entailed that for each participant in each session of activity, we generated six timelines of behavior: a timeline for *head behaviors*, one for *torso behaviors*, and one for *arms/hands behaviors* (see Figure 17), a timeline for *gestural support*, one for *postural support*, and one for *quality of gesture*.

The added value of this coding system resides in its ability to measure behavior in its parts – as the overall duration of single behaviors – but also in its progression – as the succession of different body configurations. Indeed, if we slice the six timelines at a certain point in time, we get a picture of the body configuration of the participant at that precise moment. As an example, in the case of the participant on the left in Figure 18, the body configuration is the following: *gaze toward the partner* with *no gestural support*, *none of the target torso movements*, and *none of the target arms/hands movements*; whereas for the participant on the right, it is: *gaze toward game* with *positive gestural support*, *near reach/lean toward game* with *positive postural support*, and *manipulate game* with *positive quality of gesture*. We hypothesize that these body configurations can be used to infer different levels of engagement intensity.

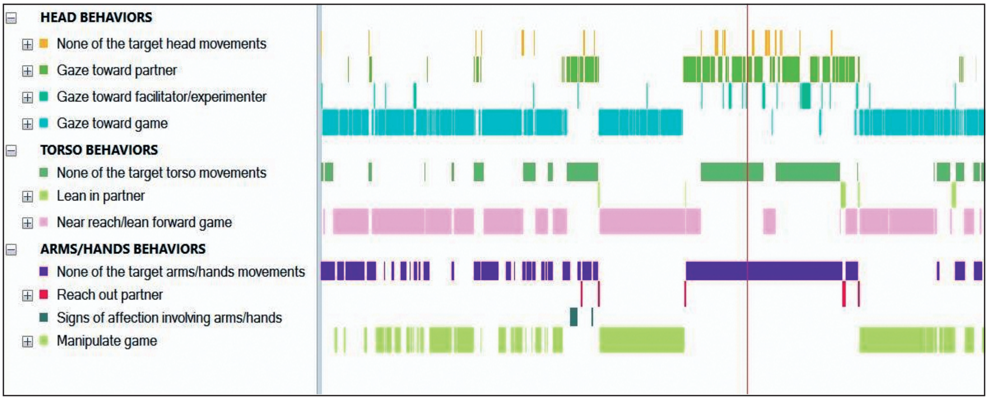


Figure 17. Scoring of ELICSE. The fragments of the three timelines of head, torso, and arms/hands behaviors

Figure 18. Examples of different body configurations. Participant on the left: gaze toward partner (with no gestural support), none of the torso movements, none of the arms/hands movements; participant on the right: gaze toward game (with positive gestural support), near reach/lean toward game (with positive postural support), and manipulate game (with positive quality of gesture)

6.5 Inter-rater Reliability

The IRR of the coding system was performed on twelve videos (29% of the database). The videos were scored by two coders: the researcher involved in the study (GP) and an external independent coder that had not been involved in the study (TvT).

IRR between coders was calculated using the software Observer XT 10.5 with the Cohen's kappa statistic (Cohen, 1960). Observer XT calculates IRR by taking into account both the matching between the behaviors scored by the two coders and their overlap in time. We computed the global Cohen's kappa of the coding system, the Cohen's kappa of the behavior groups, and the Cohen's kappa of single behaviors. With regards to the latter, we included in the calculation of the IRR only behaviors with a mean duration higher that 1% of the session. Indeed, the Cohen's kappa statistic is not accurate with very infrequent behaviors. We report kappa coefficients of behaviors occurring less than 5% of the time. However, we suggest to interpret them with caution (Dael et al., 2012).

Table 12. Inter–rater reliability results

	GAME-BASED COGNITIVE STIMULATION					ROBOT-BASED FREE PLAY				
Behavior	Duration (%)			Reliability		Duration (%)			Reliability	
	Mean	Min	Max	Prop.	Kappa	Mean	Min	Max	Prop.	Kappa
GP	2.19	.24	5.60	.83	.70	13.70	3.16	22.41	.82	.69
GFE	5.31	1.79	12.24	.83	.71	4.48	.42	12.29	.75	.59
GG	88.46	81.86	92.19	.87	.78	73.30	59.26	96.16	.81	.70
NoH	4.04	2.27	6.65	.76	.59	8.52	.27	15.69	.67	.46
HEAD	100.00			.81	.76	100.00			.76	.70
LIP	.03	.00	.23	-	-	.42	.00	3.48	-	-
NRLTG	58.96	10.58	100.00	.87	.76	41.67	6.80	100.00	.81	.71
NoT	41.01	.00	89.43	.87	.75	57.91	.00	93.21	.89	.79
TORSO	100.00			.87	.74	100.00			.82	.73
RoP	.67	.00	2.89	-	-	.41	.00	1.31	-	-
RoFE	.06	.00	.43	-	-	.11	.00	.32	-	-
MG	64.13	49.94	75.97	.77	.59	45.03	6.49	84.33	.80	.70
SOA_pos	.16	.00	1.57	-	-	.18	.00	1.72	-	-
NoAH	34.98	23.94	47.18	.82	.67	54.27	15.09	92.21	.88	.77
ARMS/ HANDS	100.00			.77	.63	100.00			.79	.71
ELISCE	100.00			.80	.78	100.00			.77	.74

(-) The duration of the behavior is too short to allow interpretation. The results of inter-rater reliability for the behavior *negative signs of affection of the torso* (SOA_neg) are not reported as they did not occur in the 12 videos. For the abbreviations in the table, refer to the coding system in Table 11

To evaluate the results of IRR, we referred to the thresholds set by Fleiss (1981) and Bakeman and Gottman (1987). Fleiss suggested that a kappa between .40 and .60 represented a *fair agreement*, between .60 and .75 a *good agreement*, and above .75 an *excellent agreement*.

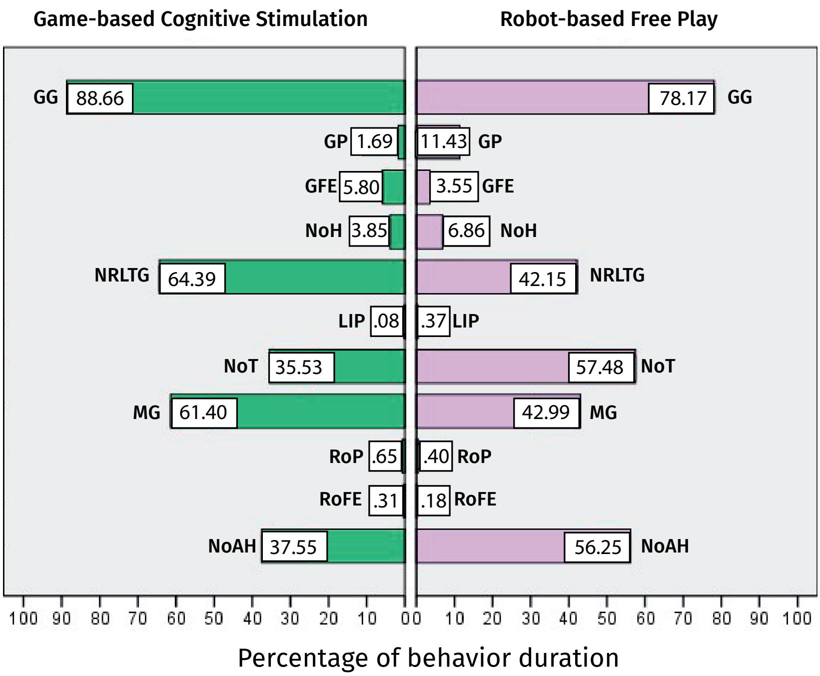
Bakeman and Gottman (1987) considered a kappa coefficient lower than .70 as insufficient, and proposed to interpret it with suspicion.

We report the results of the IRR in Table 12. With regards to the global IRR, this proved to be excellent for the game-based cognitive stimulation (kappa= .78) and very good for the robot-based free play (kappa= .74). As for behavior groups, IRR was excellent for head behaviors in the game-based cognitive stimulation (kappa= .76) and good for head behaviors in the robot-based free play (kappa= .70), very good for torso behaviors in both game-based cognitive stimulation (kappa= .74) and robot-based free play (kappa= .73) and good for arms/hands behaviors in game-based cognitive stimulation (kappa= .63) and robot-based free play (kappa= .71).

With regards to single behaviors, IRR was good to excellent for most of them. However, for some of them it could not be scored due to a frequency issue (< 1%) and for few of them it achieved unsatisfying results. We explained the unsatisfying kappa coefficients of *none of the target head movements* in the game-based cognitive stimulation and of *gaze toward facilitator/experimenter* in the robot-based free play with their low occurrence (< 5%). With regards to the low agreement of *none of the target head movements* during robot-based free play, it could be due to the sharp differences in the frequency of the behavior across participants (from .27% to 15.69%). For what regards *manipulate game* in cognitive games, the moderate agreement could not be justified by appealing to a frequency issue. In this case, the disagreement was due to an incorrect classification of the behavior by the coders.

6.6 Concurrent Validity

In order to ascertain the concurrent validity of the ELICSE, we scored all the videos in the database (42 videos, two participants per video) and calculated the percentage of behavior duration in the observation for each behavior and modifier (for an overview of the percentage of duration of behaviors and modifiers in the two activities, see Figure 19). Then, we performed a Spearman rank correlation (one-tailed, listwise exclusion of cases, N_{GBCS}= 42, N_{RBFP}= 41) between the percentage of behavior duration of each behavior and modifier in the ELICSE and the items of the OME and OERS. This was to ascertain the strength of correlation between the new measure and the criterion measure (Mislevy and Rupp, 2010). The modifiers belonging to each body part (*head*: gestural support, *torso*: postural support, and *arms/hands*: quality of gesture) expressing the same valence were summed to each other (*positive* or *negative gestural support*, *positive* or *negative postural support*, *positive* or *negative quality of gesture*), whereas the behaviors were left in their original form.



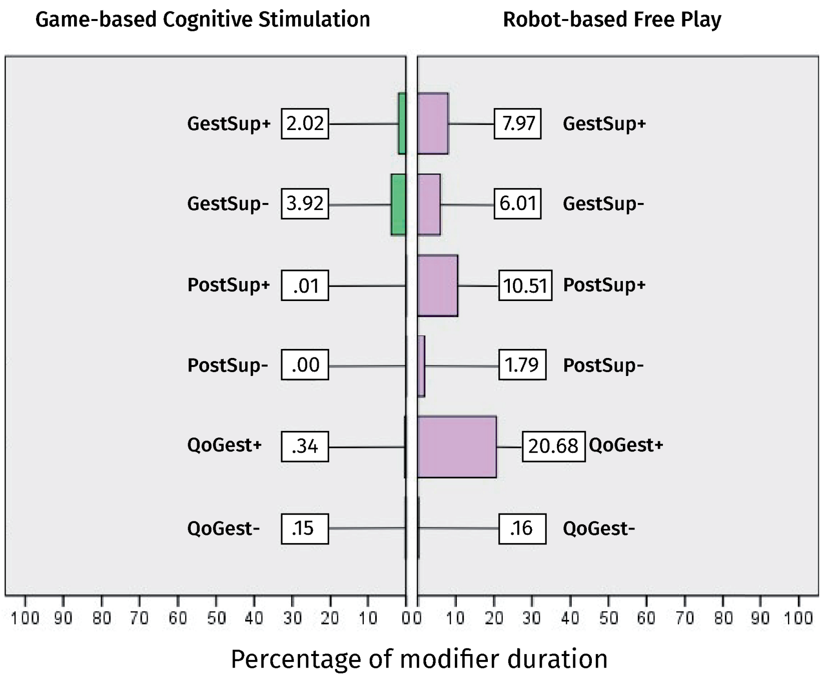


Figure 19. Boxplots of percentages of behavior and modifiers duration. For the abbreviations in the top figure, please refer to Table 11. GestSup+: positive gestural support, GestSup–: negative gestural support, PostSup+: positive postural support, PostSup–: negative postural support, QoGest+: positive quality of gesture, and QoGest–: negative quality of gesture. GestSup+, GestSup, PostSup+, PostSup–, QoGest+, and QoGest– result from the summation of the positive and negative modifiers in each behavior group

We found several significant correlations between the behaviors and modifiers of the ELICSE and the items of the OME and OERS (see Tables 13, 14, 15, and 16). We excluded all the significant correlations involving the items *anger*, *anxiety/fear*, and *sadness* of the OERS. Indeed, as stated at the end of chapter 5, there was not enough variation in these items to justify significant correlations. In the game-based cognitive stimulation (see Tables 13 and 14):

1. The item *attention* of the OME was positively correlated with *manipulate game* (*rs*(40)= .260, *p*= .048) and negatively correlated with *gaze toward facilitator/ experimenter* (*rs*(40)= -.269, *p*= .042), *reach out facilitator/experimenter* (*rs*(40)= -.340, *p*= .014), and *negative gestural support* (*rs*(40)= -.494, *p*< .001).

2. The item *attitude toward game* of the OME was positively correlated with *reach out partner* (*rs*(40)= .278, *p*= .037) and negatively correlated with *reach out facilitator/ experimenter* (*rs*(40)= -.294, *p*= .048) and *negative gestural support* (*rs*(40)= -.456, *p*= .001).

3. The item *attitude toward partner* of the OME was positively correlated with *reach out partner* (*rs*(40)= .345, *p*= .013), *positive gestural support* (*rs*(40)= .344, *p*= .013), and *positive quality of gesture* (*rs*(40)= .277, *p*= .038), while it was negatively correlated with *gaze toward facilitator/experimenter* (*rs*(40)= -.317, *p*= .020), *reach out facilitator/experimenter* (*rs*(40)= -.259, *p*= .049), and *negative gestural support* (*rs*(40)= -.437, *p*= .002).

4. The item *cognitive difficulty* of the OME was positively correlated with *negative gestural support* (*rs*(40)= .376, *p*= .007).

5. The item *pleasure* of the OERS was positively correlated with *gaze toward partner* (*rs*(40)= .481, *p*= .001), *none of the target head movements* (*rs*(40)= .277, *p*= .038)6F[7], *near reach/lean toward the game* (*rs*(40)= .306, *p*= .024), *lean in partner* (*rs*(40)= .501, *p*< .001), *reach out partner* (*rs*(40)= .276, *p*= .039), *positive gestural support* (*rs*(40)= .582, *p*< .001), and *positive quality of gesture* (*rs*(40)= .517, *p*< .001). The item *pleasure* was negatively correlated with *none of the target torso movements* (*rs*(40)= -.309, *p*= .023) and *negative gestural support* (*rs*(40)= -.367, *p*= .008).

6. The item *alertness* of the OERS was positively correlated with *gaze toward game* (*rs*(40)= .275, *p*= .039).

7 This correlation might depend on the fact that participants closed their eyes while laughing or smiling.

In the robot-based free play (see Tables 15 and 16):

1. The item *attention* of the OME was positively correlated with *gaze toward game* (*rs*(39)= .309, *p*= .025) and *manipulate game* (*rs*(39)= .393, *p*= .005) and negatively correlated with *none of the target head movements* (*rs*(39)= -.358, *p*= .011), *none of target arms/hands movements* (*rs*(39)= -.389, *p*= .006), *negative gestural support* (*rs*(39)= -.541, *p*< .001), *negative postural support* (*rs*(39)= -.363, *p*= .010), and *negative quality of gesture* (*rs*(39)= -.362, *p*= .010).

2. The item *attitude toward game* of the OME was positively correlated with *manipulate game* (*rs*(39)= .621, *p*< .001), *positive postural support* (*rs*(39)= .451, *p*= .002), and *positive quality of gesture* (*rs*(39)= .420, *p*= .003) and negatively correlated with *none of the target head movements* (*rs*(39)= -.291, *p*= .033), *none of the target arms/hands movements* (*rs*(39)= -.621, *p*< .001), *negative gestural support* (*rs*(39)= -.577, *p*< .001), *negative postural support* (*rs*(39)= -.314, *p*= .023), and *negative quality of gesture* (*rs*(39)= -.346, *p*= .013).

3. The item *attitude toward partner* of the OME was positively correlated with *manipulate game* (*rs*(39)= .375, *p*= .008) and negatively correlated with *gaze toward facilitator/experimenter* (*rs*(39)= -.412, *p*= .004) and *none of the target arms/hands movements* (*rs*(39)= -.383, *p*= .007).

4. The item *pleasure* of the OERS was positively correlated with *lean in partner* (*rs*(39)= .281, *p*= .038), *manipulate game* (*rs*(39)= .536, *p*< .001), *positive gestural support* (*rs*(39)= .433, *p*= .002), *positive postural support* (*rs*(39)= .595, *p*< .001), and *positive quality of gesture* (*rs*(39)= .570, *p*< .001), while it was negatively correlated with *none of the target arms/hands movements* (*rs*(39)= -.536, *p*< .001), *negative gestural support* (*rs*(39)= -.581, *p*< .001), and *negative postural support* (*rs*(39)= -.329, *p*= .018).

5. The item *alertness* of the OERS was positively correlated with *gaze toward game* (*rs*(39)= .385, *p*= .007) and negatively correlated with *none of the target head movements* (*rs*(39)= -.376, *p*= .008) and *negative postural support* (*rs*(39)= -.309, *p*= .025).

Table 13. Concurrent validity of the behaviors in the ELICSE for game-based cognitive stimulation: correlations tested (in bold the significant ones: *≤.05, **<.01, *<.001)**

GAME-BASED COGNITIVE STIMULATION (BEHAVIORS)												
CORRELATION OME & OERS/ ELICSE*		GG	GP	GFE	NoH	NRLTG	LIP	NoT	MG	RoP	RoFE	NoAH
Attention	*r*(s)	.187	-.058	**-.269**	-.091	.207	.003	-.207	**.260**	.195	**-.340**	-.248
	p-value	.117	.357	***.042**	.283	.094	.493	.094	***.048**	.108	***.014**	.057
Attitude game	*r*(s)	.063	.171	-.251	-.044	.153	.249	-.153	.148	**.278**	**-.294**	-.161
	p-value	.346	.140	.055	.390	.166	.056	.166	.174	***.037**	***.029**	.155
Attitude partner	*r*(s)	.056	.166	**-.317**	.038	.140	.206	-.140	.040	**.345**	**-.259**	-.057
	p-value	.362	.147	***.020**	.406	.188	.095	.188	.401	***.013**	***.049**	.359
Cognitive difficulty	*r*(s)	.129	-.049	-.128	.044	-.145	.047	.136	.109	-.141	-.012	-.098
	p-value	.208	.378	.209	.392	.180	.383	.195	.247	.186	.469	.269
Pleasure	*r*(s)	-.234	**.481**	-.101	**.277**	**.306**	**.501**	**-.309**	.153	**.276**	-.114	-.184
	p-value	.068	*****.001**	.262	***.038**	***.024**	*****<.001**	***.023**	.167	***.039**	.236	.122
Anger	*r*(s)	-.184	.249	.065	**.327**	.037	**.426**	-.037	-.074	.065	.239	.037
	p-value	.121	.056	.342	***.017**	.408	****.002**	.408	.321	.340	.064	.408
Anxiety/Fear	*r*(s)	.006	.110	-.174	.193	.110	**.319***	-.110	.032	.202	.152	-.110
	p-value	.484	.245	.135	.110	.245	***.020**	.245	.420	.100	.168	.245
Alertness	*r*(s)	**.275***	-.220	-.146	-.208	-.141	-.003	.141	-.089	.140	-.089	.098
	p-value	***.039**	.081	.178	.093	.186	.492	.186	.289	.187	.288	.268

*The item *sadness* of the OERS did not yield any result, hence it does not appear in the list

Table 14. Concurrent validity of the modifiers in the ELICSE for game-based cognitive stimulation: correlations tested (in bold the significant ones: *≤.05, **<.01, *<.001)**

GAME-BASED COGNITIVE STIMULATION (MODIFIERS)					
CORRELATION OME & OERS/ ELICSE*		GestSup_pos	GestSup_neg	QoGest_pos	QoGest_neg
Attention	*r*(s)	.189	**-.494**	.028	-.055
	p-value	.115	*****<.001**	.429	.366
Attitude game	*r*(s)	.229	**-.456**	.149	-.108
	p-value	.073	*****.001**	.173	.247
Attitude partner	*r*(s)	**.344**	**-.437**	**.277**	-.078
	p-value	***.013**	****.002**	***.038**	.312
Cognitive difficulty	*r*(s)	-.074	**.376****	.035	-.010
	p-value	.320	**.007**	.413	.474
Pleasure	*r*(s)	**.582**	**-.367**	**.517**	-.048
	p-value	*****<.001**	****.008**	*****<.001**	.380
Anger	*r*(s)	.083	-.015	.089	**.317***
	p-value	.301	.463	.288	**.020**
Anxiety/Fear	*r*(s)	.200	-.150	**.269***	-.051
	p-value	.102	.171	**.043**	.375
Alertness	*r*(s)	-.053	-.003	-.069	-.033
	p-value	.370	.494	.332	.418

*The item *sadness* of the OERS did not yield any result, hence it does not appear in the list. The modifiers *positive postural support* and *negative postural support* are not enlisted as they appeared respectively .01% and .00% of the time during the sessions of game-based cognitive stimulation

Table 15. Concurrent validity of the behaviors in the ELICSE for robot-based free play: correlations tested (in bold the significant ones: *≤.05, **<.01, *<.001)**

ROBOT-BASED FREE PLAY (BEHAVIORS)												
CORRELATION OME & OERS/ ELICSE*		GG	GP	GFE	NoH	NRLTG	LIP	NoT	MG	RoP	RoFE	NoAH
Attention	*r*(s)	**.309**	-.062	-.171	**-.358**	-.082	.202	.078	**.393**	-.176	.022	**-.389**
	p-value	***.025**	.350	.142	***.011**	.306	.103	.314	****.005**	.136	.447	****.006**
Attitude game	*r*(s)	.220	.051	-.159	**-.291**	-.163	.107	.169	**.621**	.005	.036	**-.621**
	p-value	.084	.375	.161	***.033**	.154	.253	.146	*****<.001**	.487	.411	*****<.001**
Attitude partner	*r*(s)	.020	.030	**-.412**	-.043	.045	.227	-.036	**.375**	.185	-.013	**-.383**
	p-value	.450	.427	****.004**	.394	.391	.076	.410	****.008**	.123	.467	****.007**
Pleasure	*r*(s)	.063	.050	.136	-.125	-.250	**.281**	.245	**.536**	.168	.024	**-.536**
	p-value	.349	.378	.198	.217	.057	***.038**	.062	*****<.001**	.147	.441	*****<.001**
Anger	*r*(s)	**-.320**	-.172	.046	**.451**	**.280**	.168	**-.288**	**-.388**	.235	-.153	**.372**
	p-value	***.021**	.140	.387	****.002**	***.038**	.146	***.034**	****.006**	.069	.169	****.008**
Anxiety/Fear	*r*(s)	.043	-.062	-.006	.093	.188	-.033	-.188	-.041	.090	-.209	.041
	p-value	.396	.351	.486	.282	.120	.420	.120	.398	.288	.095	.398
Sadness	*r*(s)	-.215	**-.270**	-.084	**.373**	**.316**	.092	**-.316**	**-.288**	.104	-.059	**.269**
	p-value	.089	***.044**	.301	****.008**	***.022**	.283	***.022**	***.034**	.258	.358	***.045**
Alertness	*r*(s)	**.385**	-.225	-.187	**-.376**	-.192	.019	.198	.258	-.144	.074	-.258
	p-value	****.007**	.079	.121	****.008**	.115	.454	.108	.052	.185	.322	.052

*The item *cognitive difficulty* of the OME did not yield any result, hence it does not appear in the list

Table 16. Concurrent validity of the modifiers in the ELICSE for robot-based free play: correlations tested (in bold the significant ones: *≤.05, **<.01, *<.001)**

ROBOT-BASED FREE PLAY (MODIFIERS)							
CORRELATION OME & OERS/ ELICSE*		GestSup_ pos	GestSup_ neg	PostSup_ pos	PostSup_ neg	QoGest_ pos	QoGest_ neg
Attention	*r*(s)	.233	**-.541**	.195	**-.363**	.211	**-.362**
	p-value	.071	*****<.001**	.111	****.010**	.093	****.010**
Attitude game	*r*(s)	.231	**-.577****	**.451**	**-.314**	**.420**	**-.346**
	p-value	.073	*****<.001**	****.002**	***.023**	****.003**	***.013**
Attitude partner	*r*(s)	-.113	-.122	.152	-.057	.170	-.059
	p-value	.241	.224	.171	.363	.145	.357
Pleasure	*r*(s)	**.433**	**-.581**	**.595**	**-.329**	**.570**	-.224
	p-value	****.002**	*****<.001**	*****<.001**	***.018**	*****<.001**	.080
Anger	*r*(s)	-.171	**.450**	**-.382**	**.794**	**-.387**	**.665**
	p-value	.143	****.002**	****.007**	*****<.001**	****.006**	*****<.001**
Anxiety/Fear	*r*(s)	**-.373**	.035	-.171	.174	**-.298**	.103
	p-value	***.008**	.415	.143	.138	***.029**	.261
Sadness	*r*(s)	-.020	**.356**	**-.308**	**.657**	**-.296**	**.569**
	p-value	.452	***.011**	***.025**	*****<.001**	***.030**	*****<.001**
Alertness	*r*(s)	.018	-.135	.057	**-.309**	.068	-.190
	p-value	.457	.201	.363	***.025**	.337	.117

*The item *cognitive difficulty* of the OME did not yield any result, hence it does not appear in the list

6.7 Discussion

6.7.1 Research Questions

The development of the ELICSE enabled us to respond to the research questions: *Which behaviors externalize engagement in people with dementia? Are these behaviors appropriate to measure engagement?* According to the analyses of the IRR and concurrent validity, the ELICSE is an appropriate measure of engagement. Indeed, on the one hand, it achieved a Cohen's kappa well above the .75 threshold (i.e., mean kappa of both activities). On the other hand, it showed a high correlation tendency with the gold standard scales, the OME and the OERS. As a matter of fact, with the exception of the items anger, anxiety/fear, and sadness – which did not variate enough in the database – the ELICSE correlated with all the items of the OME and OERS.

Also, the ELICSE enabled us to satisfy one of the requirements to get to a proper measurement framework of engagement in dementia, which is to *establish a methodology to measure engagement-related behavior in its units and as a body configuration across activities*. One of the main advantages of the ELICSE is its ability to measure engagement-related behavior in its parts – as the overall duration of single behaviors – but also in its progression – as the succession of different body configurations. The potential of the latter approach has not been investigated yet. We assume that by taking into account body configurations, one can measure different engagement intensities. To pursue this objective, one first needs to associate a score to different body configurations based on the intensity of engagement that they express, then s/he needs to use a time-sampling technique to code the progression of this intensity throughout an activity session.

6.7.2 Behavioral Dynamics of Engagement

The correlations with the OME and OERS did not just provide information on the concurrent validity of the ELICSE, but also enlightened the behavioral dynamics taking place during activities. For instance, during game-based cognitive stimulation, the behaviors directed towards the partner were synonymous of a positive attitude towards the game and were importantly correlated with positive affect, while the behaviors expressing negative affect were connected with cognitive exertion. Apparently, in a situation where the challenges are high (i.e., cognitive difficulty) and the skills are low (mild and moderate dementia), the reciprocal social support (behaviors directed towards the partner) might act as a mediating variable for finding engagement (positive attitude towards the game and positive affect).

With regards to the robot-based free play, engagement was more robot-centric. The behaviors directed towards the game were related to positive affect, while the behaviors directed somewhere else other than the foci of the activity were related to negative affect. In this context, sociality appeared more as a consequence of the activity itself than of its difficulty. It was built in it. Pleo constituted a shared point of attention for the two participants. In fact, in robot-based free play, the presence of behaviors not directed towards the foci was not just a symptom of negative attitude towards the game, but also of negative attitude towards the partner.

In addition to these observations, we would like to pose the attention of one fact: the behaviors directed toward the facilitator and experimenter implied a lack of attention toward the game. The role of the facilitator during activities is a supporting role, s/he helps the person with dementia to overcome the challenges that the activity prospects. As a result, when participants resort to the facilitator, it means that the activity is not progressing as smoothly as it should. On the opposite, when they do not turn to him/her, it means that the activity is proceeding harmoniously.

6.7.3 Adaptability of ELICSE

The ELICSE has been developed with participants with mild and moderate dementia in the context of game-based cognitive stimulation and robot-based free play. However, for its characteristics: (1) it can be modified to suit the needs of people with different levels of dementia (e.g., MCI, severe dementia) and (2) it can adapted to score engagement in most of the co-activities described in chapter 3 (e.g., sensory stimulation).

With regards to point 1, the ELICSE can be adapted to further activities with a three-step procedure. First, the researcher has to define the *foci* of the activity and the *body portion involvement* that the activity entails. Second, s/he has to enumerate the *directional behaviors* performed by participants to address the foci of the activity with each of the involved body parts. Third, s/he has to identify the *gestures* that carry a *positive or negative affective meaning* which are executed by participants on top of (or as part of) the directional behaviors. The first step defines the behavior groups in the ELICSE, the second step the single behaviors, the third step the modifiers.

With regards to point 2, the ELICSE can be adapted to meet the characteristics of different dementia profiles. Indeed, it can be scaled up or down. When measuring engagement in persons with severe dementia, the researcher might want to remove infrequent behaviors, such as arms/hands movements. On the opposite, when assessing engagement in persons with MCI, the researcher might want to add behavioral modalities in order to enrich the assessment of engagement with further details. As the behavior groups in the ELICSE are measured independently from each other, the coding system can be downsized or enlarged with new modalities without altering the balance between single behaviors.

6.7.4 Meaning of Behaviors

In this section, we focus on associating a meaning to the behaviors in the ELICSE based on the components of engagement in our theoretical framework. We do so in three steps: (1) we differentiate between behaviors of engagement and disengagement, (2) we identify behaviors of rapport, focused attention, and valence, and (3) we distinguish between behaviors of positive and negative valence.

With regards to behaviors of engagement and disengagement, we included in the latter category: *none of the target head movements*, *none of the target torso movements*, and *none of the target arms/hands movements*. These behaviors were all targeted to a focus other than the activity. Also, we included in this category *gaze toward facilitator/experimenter*, and *reach out facilitator/experimenter*. Throughout activities, these behaviors were strongly correlated with lack of engagement. The remaining behaviors were considered *behaviors of engagement* and were further divided into behaviors of rapport, focused attention, and valence. We considered: (i) the partner as the significant other in the activity and chose it as the focus of *rapport*, (ii) the game as the limited stimulus field that the participant needs to concentrate on and selected it as the target of *focused attention*, and (iii) the affective characterization of behavior as the hedonic tone of the activity (unpleasant vs pleasant) and chose it to represent *valence*. As a result, rapport was composed by: *gaze toward partner*, *lean in partner*, and *reach out partner*. Focused attention was represented by: *gaze toward game*, *near reach/lean toward game*, and *manipulate game*. Last, valence consisted of: *gestural support*, *postural support*, and *quality of gesture*. As the behaviors of valence expressed different shades of meaning, they were further divided into positive and negative. We included *positive gestural support*, *positive postural support*, and *positive quality of gesture* into *positive valence*; *negative gestural support*, *negative postural support*, and *negative quality of gesture* into *negative valence*. We excluded the neutral modifiers from this classification, as they were already accounted for by the main behaviors. The final result of the sort out process is presented in Table 17.

Table 17. Assignment of behaviors to components of engagement

ENGAGEMENT-RELATED BEHAVIOR				EXCLUDED BEHAVIOR
FOCUSED ATTENTION	**RAPPORT**	**VALENCE**		*None of the target head behaviors*
		POSITIVE	**NEGATIVE**	*None of the target torso behaviors*
Gaze toward game	Gaze toward partner	Positive gestural support	Negative gestural support	*None of the target arms/ hands behaviors*
Near reach/lean toward game	Lean in partner	Positive postural support	Negative postural support	*Gaze towards facilitator/ experimenter*
Manipulate game	Reach out partner	Positive quality of gesture	Negative quality of gesture	*Reach out facilitator/experimenter*

6.8 Conclusions

This chapter described the development of the ELICSE, a coding system that quantifies engagement-related behavior across diverse activities. Through the ELICSE, we achieved another milestone of this dissertation, establish a methodology to measure engagement-related behavior in its units and as a body configuration across activities. Moreover, we identified the behaviors to include in the theory-driven model of engagement to assess focused attention, rapport, and valence. In the next chapter, we present the final version of the theory-driven model of engagement updated according to the results of chapter 5 and 6 and test it using SEM.

7

Construct Modeling: The Theory-driven Model of Engagement

7.1 Introduction

This chapter is aimed at describing and testing the final version of the theory-driven model of engagement and is organized as follows. First, we frame the problem and detail the methodological issues that a model of engagement for people with dementia is supposed to address. Second, we collocate the measures of engagement deployed in chapters 5 and 6 – the features of EDA and the behaviors of the ELICSE – in the model under the respective components of engagement. Third, we test the theory-driven model of engagement with SEM to ascertain its goodness of fit. Last, we discuss the results of the test and apply modifications to the model when necessary.

7.2 Problem Framing and Model Contribution

As highlighted by the literature, engagement can be measured on three different levels, according to three diverse response systems: (i) the experiential/subjective – which deals with the personal self-perceived experience of engagement, (ii) the behavioral/expressive – which addresses the outer manifestation of engagement through behavior, and (iii) the peripheral-physiological – which treats the physiological substrate of engagement.

These three systems are not equally accessible in dementia. The *experiential/subjective* response system, in particular, is the most heavily affected by the disease. Indeed, the progression of dementia makes it always more difficult for a person to access his/her psychological states and consistently report them. Also, subjective reports of engagement are liable to response biases. We observed both of these dynamics during the exploratory study (see Appendix B). Participants struggled to respond to the questions of the SIQ, as they could not distinctly recall all the activities they had participated to and rank them in order of preference. Moreover, they seemed to reply to the questions according to a *social desirability* and *acquiescence bias*, as they tended to endorse any statement made by the facilitator and experimenter, regardless of its content.

Concerning the *behavioral/expressive* response system, disturbances like apathy and depression – which affect conation and have high prevalence rates in dementia – might blunt the expression of engagement at a behavioral level, and restitute an incomplete image of engagement for some people with dementia. As can be observed from the boxplots in Appendix F, also in our sample, participants with dementia-associated apathy and depression displayed a reduction in engagement-related behavior. They showed less attention, a worse attitude towards the game and the partner, and lower pleasure according to the OME and OERS; less behaviors of engagement (i.e., *manipulate the game*) and positive valence (i.e., *positive gestural support, positive quality of gesture*), and more behaviors of disengagement (i.e., *none of the target head movements*, *none of the target torso movements*) and negative valence (i.e., *negative gestural support*) according to the ELICSE.

With regards to the *peripheral-physiological* response system, the measurement of the physiology of engagement in dementia is very rare in the literature since it is riddled with methodological pitfalls. In the big picture of engagement in dementia, the peripheral-physiological system can be used to integrate the behavioral/expressive level of assessment, especially when this is affected by motivational disorders and other behavioral disturbances (e.g., agitation, anxiety). The results of the analysis of EDA are rather promising in this sense. EDA casts a different light on the engagement of the participants with motivational disorders with respect to the one shed by the OME and OERS and the ELICSE. KURT EDA and SKEW EDA – which in game-based cognitive stimulation were positively correlated with alertness and in robot-based free play were negatively correlated with attention and attitude towards the game – were higher in participants with motivational disorders during the former activity and almost even between groups in the latter.

These methodological considerations – which are corroborated by the results of the exploratory study and the further analyses in Appendix F – tell us that self-reports are not always reliable measures of engagement in dementia. They can be adopted during early stages of the disease, but become always more unfeasible the more dementia progresses. For this reason, we excluded them as a gold standard of engagement. Also, these considerations pinpoint that the behavioral/expressive and peripheral-physiological response systems need to be combined to provide a comprehensive assessment of engagement for dementia. This is what we attempt to do in this chapter. In chapter 5, we analyzed EDA and ascertained its concurrent validity as a peripheral-physiological measure of engagement. In Chapter 6, we developed the ELICSE, tested its inter-rater reliability, and determined its concurrent validity as a behavioral/expressive measure of engagement. In this chapter, we finalize the theory-driven model of engagement sketched in chapter 3 by enclosing the features of EDA and the behaviors of the ELICSE under the respective components of engagement and test it with SEM.

Beyond the pure combination of measures belonging to different response systems, a model of engagement serves three purposes: (1) understand whether the deployed behavioral/expressive and peripheral-physiological measures of engagement really measure the components of engagement that they are supposed to measure, (2) set the boundaries between the different components of engagement so that there is no overlap or contradiction between them, and (3) define and formalize the relationships between these components to measure engagement in a more consistent way.

7.3 Components and Measures of Engagement

In chapter 3, we drew the theoretical framework of this dissertation and identified three components of engagement accessible in dementia: (1) focused attention, (2) rapport, and (3) core affect.

1. *Focused attention* was defined as the voluntary focusing of attention on a limited stimulus field (Csikszentmihalyi, 2014).

2. *Rapport* was defined as a meaningful human experience of close and harmonious connection with another that involves common understanding (Tickle-Degnen, 2006).

3. *Core affect* was defined as the neurophysiological state accessible to consciousness as a single simple feeling (Russel, 2003) which is a blend of two dimensions, *valence* (displeasure-pleasure) and *arousal* (deactivation-activation).

In Figure 20, we present the final version of the theory-driven model of engagement. This version is structurally the same as the one in chapter 3. However, it features the measures of the different components of engagement validated in chapters 5 and 6. In the theory- driven model, the state of *engagement* is represented as a latent variable that has a direct effect on another three latent variables. These are the components of engagement according to the literature: focused attention, rapport, and core affect. *Focused attention* is measured through the behaviors of the ELICSE directed toward the game – *gaze toward game* (GG), *near reach/lean toward game* (NRLTG), and *manipulate game* (MG). *Rapport* is measured through the behaviors of the ELICSE directed toward the partner – *gaze toward partner* (GP), *lean in partner* (LIP), and *reach out partner* (RoP). With regards to *core affect*, given its blended nature, it is split into two ulterior latent variables which represent the orthogonal axes of the circumplex model of affect, valence and arousal. *Valence* is quantified via the modifiers in the ELICSE expressing positive and negative valence – *gestural support* (Gest Sup), *postural support* (Post Sup), and *quality of gesture* (QoGest; the modifiers were aggregated into a unique score with the operations outlined in Table 18). *Arousal* is measured with the latent variables *tonic* and *phasic* EDA. *Tonic EDA* is gauged through the features MEAN EDA, STD EDA, and SUM H EDA. *Phasic EDA* is measured with the features SMA EDA and NPR EDA.

As can be noticed, the behaviors directed toward the facilitator/experimenter and those targeted to a focus other than the activity were left aside from the model. The former

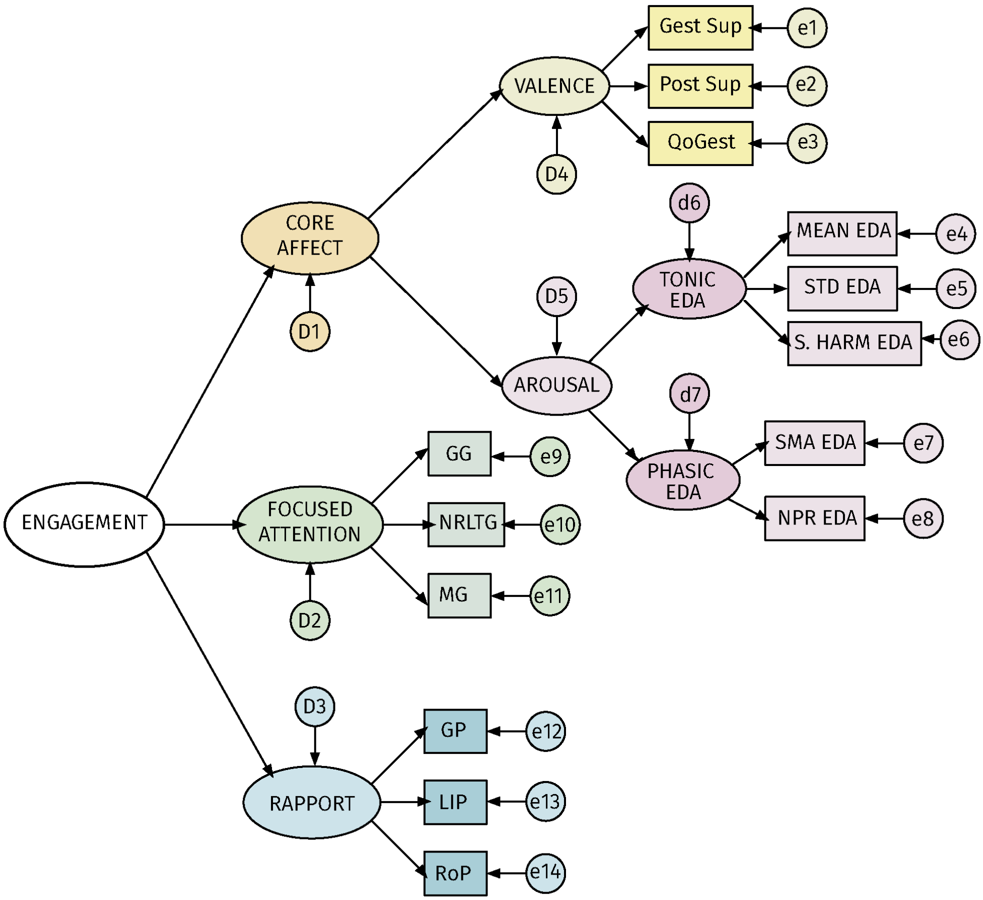


Figure 20. Theory-driven model of engagement

Table 18. Aggregation of modifiers in the ELICSE (for the abbreviations in this table, refer to Table 11)

Variable name	Abbreviation	Data reduction
GESTURAL SUPPORT	Gest Sup	= (GP_pos + GG_pos) – (GP_neg + GG_neg)
POSTURAL SUPPORT	Post Sup	= (LIP_pos + NRLTG_pos) – (LIP_neg + NRLTG_neg)
QUALITY OF GESTURE	QoGest	= (RoP_pos + MG_pos + SOA_pos) – (RoP_neg + MG_neg + SOA_neg)

were negatively correlated with engagement, the latter were considered behaviors of disengagement. Also the features representing the frequency distribution of the EDA signal – KURT EDA and SKEW EDA – were excluded from the model. This was because they provided conflicting results across activities. Last, RNG EDA was ruled out from the model, as it was not a positive definite with STD EDA, thus it was redundant.

According to the theory-driven model of engagement, the more engaged is the person with dementia, the more focused attention, rapport, and core affect increase. This latter component increments for the simultaneous effect of the positivity of valence and of the rise of the tonic and phasic EDA. The theory-driven model of engagement is strictly related to the following definition of engagement drawn from the literature and extended to dementia in chapter 3: engagement is the psychological state of enjoyment and proactive attentiveness experienced by a person with dementia involved in a meaningful activity.

7.4 Model Testing

The theory-driven model of engagement was tested using SEM with the software SPSS AMOS 22.0[8]. SEM is a set of statistical techniques that simultaneously estimates a set of interrelated equations and permits the estimation of direct, indirect, and reciprocal effects within a complex model (Mishler and Rose, 2005; see Appendix G for a SEM glossary). As SEM is a large sample size statistic, we used the data coming from both activities to test the model.

As we had a moderate amount of missing data for EDA (13% of the EDA database; $N_{measured}$= 73, $N_{missing}$= 11), and these data were missing completely at random (MCAR), we used a multiple regression imputation (5 imputations) to assign values to the missing cases (Schumacker and Lomax, 2016). MCAR means that the missingness of the data does not depend on any of the measured variables (Schafer, 1997). As Kang (2013) reports: "*if data are missing by design, because of an equipment failure or because the samples are lost in transit or technically unsatisfactory, such data are regarded as being MCAR*". In our case, the invalid EDA data were all missing because of a failure of the E4 wristband or because the electrodes of the EDA response sensor got detached from the skin surface during data collection.

The first test of the theory-driven model of engagement gave a negative outcome. Indeed, we were not able to obtain the coefficients relative to the goodness of fit, nor the parameters estimates for the latent variables. This was due to an inflated negative error variance on the latent variables *valence* (D4= -303.915) and *rapport* (D3= -.036), which made the solution to the theory-driven model of engagement inadmissible.

8 https://www.ibm.com/analytics/spss-statistics-software

7.4.1 Analysis of the Causes of Negative Error Variance

According to the literature (Garson, 2008), there might be several reasons for a negative error variance in SEM: the presence of outliers, a high multicollinearity between indicators, the presence of negative correlations between variables, and a small sample size.

We investigated the presence of each of these causes in our dataset, as they could provide precious information to modify and refine the model. First, we calculated the Mahalanobis distance and removed the observations farthest from the centroid ones (*p*= .05) – the outliers – from our dataset (N= 75). Then, we performed SEM again. Also this time, the test gave a negative outcome caused by an inflated negative error variance on the latent variable *valence* (D4= -424.109), a variance of zero on the latent variable *rapport*, and a negative error variance on the indicator *gaze toward partner* (e12= -3.423).

For both the original dataset and the dataset deprived of outliers, the sample size could not be the source of the negative error variance. Indeed, it was adequate enough to perform SEM as demonstrated by a Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy higher than the threshold level of .50 ($KMO_{N=84}$= .673; $KMO_{N=75}$= .693; Hair et al., 1998; Tabachnick and Fidell, 2007).

As the sample size and the presence of outliers did not provide any explanation to the negative error variance, we examined the occurrence of multicollinearity. The variance inflation factor (VIF) threshold recommended for SEM is ≤ 3.3 (Petter et al., 2007; Cenfetelli and Bassellier, 2009). In the original dataset:

1. *Near reach/lean toward game* was collinear with *gaze toward game* (VIF= 3.892).

2. *Manipulate game* was collinear with *gaze toward game* (VIF= 3.819).

3. *Lean in partner* was collinear with *gaze toward partner* (VIF= 5.325).

4. *Reach out partner* was collinear with *gaze toward partner* (VIF= 5.262).

5. G*estural support* was close to be collinear with *postural support* (VIF= 3.204) and *quality of gesture* (VIF= 3.155).

In the dataset without outliers, we found the same results, but with higher VIFs:

1. *Near reach/lean toward game* was collinear with *gaze toward game* (VIF= 4.332).
2. *Manipulate game* was collinear with *gaze toward game* (VIF= 4.403).
3. *Lean in partner* was collinear with *gaze toward partner* (VIF= 6.193).
4. *Reach out partner* was collinear with *gaze toward partner* (VIF= 6.279).
5. *Gestural support* was collinear with *postural support* (VIF= 4.802) and *quality of gesture* (VIF= 5.568).

As a result of this analysis, we attributed the negative error variance in the theory-driven model of engagement to the high multicollinearity between indicators. However, we realized that negative correlation might have also had a role in the inadmissibility of the model. Indeed, as *focused attention* and *rapport* concur for the same attentional resources during activities, they could have been negatively correlated. We ran the SEM in Figure 21 to confirm this hypothesis. The model proved to be an excellent fit for the data ($X^2(8, N = 75)= 9.125$, $p= .332$; *RMSEA*= .041; *NFI*= .935; *CFI*= .991; *RFI*= .829; *PNFI*= .356) and the correlation between *focused attention* and *rapport* was significant and negative ($r(73)= -.766$, $p< .001$).

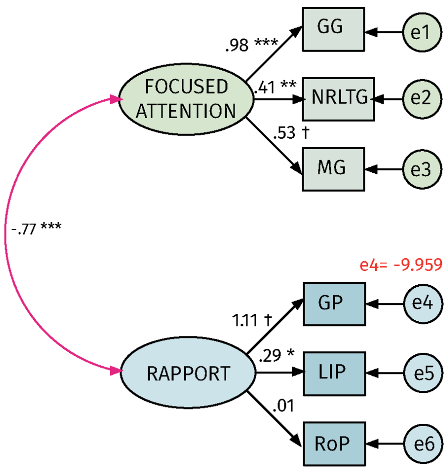


Figure 21. Correlation between focused attention and rapport. † **fixed factor, *< .05, **< .01, ***< .001. In red, the negative error variance on gaze toward partner**

7.4.2 Model Modification

One of the main conclusions that we drew from the misspecification of the theory-driven model of engagement was that focused attention and rapport are not separate components of engagement as postulated by the literature. Instead, they are different forms of focused attention. In fact, rapport can be conceived as the voluntary focus of attention on the limited stimulus field partner. One of the corrections that we made to the theory-driven model of engagement to tackle this issue was that of merging the latent variables *rapport* and *focused attention* and consider both the behaviors directed toward the game and those targeted at the partner as behaviors of focused attention.

Another inaccuracy of the theory-driven model of engagement – which might have caused the high multicollinearity between indicators – was that of excluding behaviors of disengagement (i.e., *none of the target head movements*, *none of the target torso movements*, and *none of the target arms/hands movements*, *gaze toward facilitator/experimenter*, and *reach out facilitator/experimenter*) from the computation of focused attention and rapport. To address this issue, we subtracted the behaviors directed towards the facilitator/experimenter and those that did not target any of the foci of the activity (i.e., *none of the target head movements*, *none of the target torso movements*, and *none of the target arms/hands movements*) from the summation of the behaviors directed towards the game and the partner (see Table 19). As a result, we obtained an engagement score for each of the behavioral modalities in the ELICSE – head, torso, and arms/hands – ranging between -100 and 100, where -100 represented the highest disengagement with the activity and 100 the highest engagement with it. These engagement scores were named *gaze toward activity* (GAct), *lean toward activity* (LTAct), and *reach out activity* (RoAct). In Figure 22, we present the modified version of the theory-driven model of engagement.

Table 19. Aggregation of behaviors in the ELICSE (for the abbreviations in this table refer to Table 11)

Variable name	Abbreviation	Data reduction
GAZE ACTIVITY	GAct	= (GP + GG) - (GFE + NoH)
LEAN TOWARD ACTIVITY	LTAct	= (LIP + NRLTG) – (NoT)
REACH OUT ACTIVITY	RoAct	= (RoP + MG) – (RoFE + NoAH)

7.5 Model Re-testing

We ran the test of the modified version of the theory-driven model of engagement using SEM and the software SPSS AMOS 22.0. The test of the model was inconclusive. Indeed, we found a negative error variance on the latent variable *focused attention* (D2= -27.620), a hugely inflated negative error variance on *valence* (D3= -13975.654), an error variance of zero on *SMA EDA*, and a consistent negative error variance on the observed variable *gaze toward activity* (e9= -118.239).

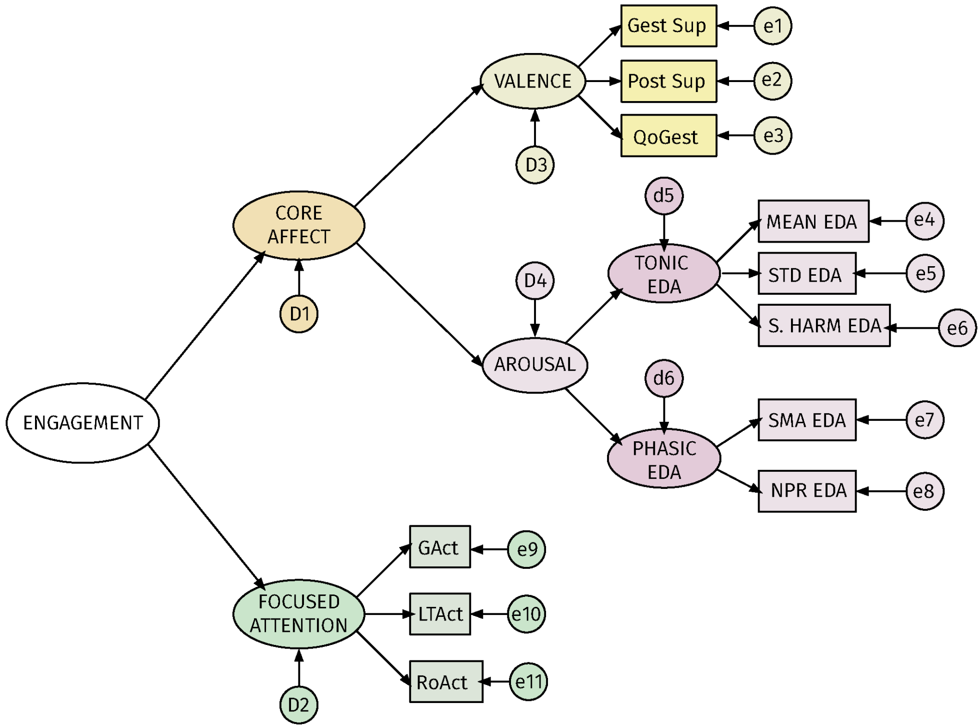


Figure 22. Modified version of theory-driven model of engagement

7.5.1 Analysis of the Causes of Negative Error Variance

As for the previous version of the model, we delved into the causes of the negative error variance in order to refine the model. First, we removed the outliers (N= 77) by identifying the observations further from the centroid ones with the Mahalanobis distance statistic. Then, we re-ran the model. This time we found an inflated negative error variance on the observed variable *quality of gesture* (e3= -525.946), a moderate negative error variance on the latent variable *focused attention* (D2= -82.681), and a mild negative error variance on the latent variable *tonic EDA* (d5= -.031).

The sample size of both datasets (with and without outliers) was adequate enough for SEM ($KMO_{N=84}$= .633; $KMO_{N=77}$= .631) and was hence excluded as the cause of negative error variance. Thereby, we proceeded to examine multicollinearity. With regards to the dataset with outliers, we did not find any multicollinearity between variables. As for the dataset without outliers, we found a high multicollinearity just on one variable, *gestural support*. This was collinear with *postural support* (VIF= 4.909) and *quality of gesture* (VIF= 4.954). As the multicollinearity was much reduced with respect to the previous version of the theory-driven model of engagement and the negative error variance issue remained unsolved even after removing the observed variable *gestural support* from the model, we excluded multicollinearity as a potential cause of the negative error variance.

The negative correlation between *focused attention* and *core affect* could not be calculated. Indeed, the size of the negative error variances in both datasets (with and without outliers) would have greatly influenced the results. Consequently, we could not establish with precision the cause of the misspecification of this second version of the theory-driven model of engagement.

7.5.2 Potential Sources of Misspecifation

In this section of the chapter, we outline a number of potential sources of misspecification of the theory-driven model of engagement, which we meticulously examine in the next chapters of the dissertation.

According to the placement of the negative error variances, we hypothesized *four* sources of misspecification of the theory-driven model of engagement: (i) the existence of hidden regression paths between the indicators of *valence* – *gestural support*, *postural support*, and *quality of gesture*, (ii) the presence of hidden regression paths between the indicators of *focused attention* – *gaze toward activity*, *lean toward activity*, and *reach out activity*, (iii) the bivariate distribution of *valence* and *arousal*, and (iv) the instability of the correlation between *focused attention* and *valence.*

With regards to the first two sources of misspecification, they imply that the indicators *gestural support*, *postural support*, and *quality of gesture* and the indicators *gaze toward activity*, *lean toward activity*, and *reach out activity* are linked to each other with regression paths. Hence, they do not have the same measurement value in terms of valence and focused attention.

Concerning the third source of misspecification, it entails that arousal and valence do not entertain a monotonic relationship and therefore cannot be nested under the same latent variable. According to the circumplex model of affect, arousal grows both when valence is positive and when it is negative. One of the issues with the theory-driven model of engagement is that it modeled engagement considering that valence could only range from neutral to positive and thus assumed that arousal and valence could linearly correlate. As this is a bias shared by most of the behavioral and physiological measures of engagement and in general by the literature on engagement, for its theory-driven nature, the tested model could not be immune to it.

With regards to the last hypothesized source of misspecification, it assumes that *focused attention* and *valence* might not always progress in the same direction as postulated by the theory-driven model of engagement. One can be extremely attentive to the activity, while expressing negative affect. Likewise, one can be disengaged, while displaying positive affect.

7.6 Discussion

To conclude, the testing of the theory-driven model of engagement highlighted a number of pitfalls in the literature and contributed to denounce a widespread erroneous vision on engagement, that of a construct imbued with exclusively positive traits. The testing of the model revealed that inattention and negative affect are as crucial and essential for the assessment of engagement as their positive equivalents (i.e., attention and positive affect). The statistical examination of the model enabled us to redefine the boundaries between the components of engagement set by the literature. For instance, it brought us to merge two components of engagement that in the literature often appear as separate: rapport and focused attention. Also, it served to gain insights to deploy a model of engagement more consistent with the data. Indeed, it brought us to hypothesize a series of relationships between components and measures of engagement whose data-driven verification can lead to a better fitting model.

As a consequence of the misspecification of the theory-driven model of engagement, we can state that the definition of engagement as a psychological state of enjoyment and proactive attentiveness put forward by the literature is not confirmed by the data.

7.7 Conclusions

This chapter presented the definitive version of the theory-driven model of engagement and described the testing of its goodness of fit. The theory-driven model of engagement was developed according to a thorough literature review and featured measures of engagement provided with good concurrent validity. However, it proved to be inconclusive across two consecutive tests of its goodness of fit, one carried out with the original model, the other performed with its modified version.

The original theory-driven model of engagement was inadmissible for a number of statistical reasons, namely high multicollinearity and negative correlation. The modified version of the model, instead, did not present any underlying statistical problem, and thus was assumed to be incorrect in its relational structure. We identified four potential sources of its misspecification: (1) the existence of hidden regression paths between the indicators of valence, (2) the presence of hidden regression paths between the indicators of focused attention, (3) the bivariate distribution of valence and arousal, and (4) the instability of the correlation between focused attention and valence.

These potential sources of misspecification will be addressed in the upcoming chapters and their resolution will lead to the ENGAGE-DEM, the model of engagement for people with dementia that constitutes the main achievement of this dissertation. In chapter 8, we investigate whether the indicators of focused attention and valence are organized in a hierarchical way and thus connected by regression paths. In chapter 9, we introduce another measure of engagement – quantity of movement – which is partially inspired by the results of chapter 8. In chapter 10, we detail the ENGAGE-DEM and test its goodness of fit with SEM. The issues regarding the relationships between valence and arousal, and between focused attention and valence are addressed in chapter 10 and inform the construction of the final model.

8

Construct Modeling: The Hierarchical Organization of Engagement-related Behavior

8.1 Introduction

In the previous chapter, we performed two consecutive tests of the theory-driven model of engagement. Both of these tests disclosed that the model was not a good structural representation of the engagement of people with dementia. We identified four potential sources of misspecification for it. Two of them concerned the behavioral/expressive level of measurement. They postulated the presence of hidden regression paths connecting the indicators of focused attention (i.e., *gaze toward activity*, *lean toward activity*, and *reach out activity*) and those of valence (i.e., *gestural support*, *postural support*, and *quality of gesture*). As the indicators of focused attention and valence belonged each to a different behavioral modality, we suspected that these regression paths described the propagation of behavior across body parts.

In this chapter, we first identify patterns in the way behavior spread across body parts. Then, we transform these patterns into hypothetical regression paths connecting the indicators of focused attention and valence and test them with SEM.

8.2 Patterns of Body Part Organization

In order to examine how behavior progressively spread across body parts, we made reference to the category body of LMA and particularly to the formalization of body part organization. The category *body* describes the structure of the body, which parts of the body are moving, which parts are linked to others, and which parts are influenced by others (Maletic, 1987). The formalization of *body part organization* specifically defines how body parts are connected in movement (Hackney, 2002). The body part organization can be *successive* (adjacent body parts move one after the other), *sequential* (non-adjacent body parts move one after the other), and *simultaneous* (all active body parts move together at the same time). In this context, we are not interested in identifying the temporal succession of behavior, but in discovering constant relationships between body parts moving successively, sequentially, and simultaneously.

8.2.1 Patterns of Body Part Organization: Focused Attention

Across activities, we observed one main pattern of body part organization in the behaviors of focused attention: *successive* – *space hold* (with various gestures) – *successive*. This pattern appeared both when participants addressed the focus *game* and when they addressed the focus *partner*. In the pattern, the threefold elements had different scopes: (i) the first successive movement was aimed at reaching the artifact or the partner in the activity, (ii) the space hold (with various gestures) was targeted at manipulating the artifact or physically interacting with the partner during the activity (e.g., pat the partner), and (iii) the second successive movement was aimed at withdrawing from the activity by moving away from the artifact or the partner.

The successive organization mostly appeared after the game was placed on the table by the facilitator (see Figure 23). In this situation, the movement of the participant towards the game was initiated by the head (*gaze toward game*) and sequenced into arms/hands (*manipulate game*) via the torso (*near reach/lean toward game*). After this successive movement, the head and the torso of the participant remained locked in the same position (i.e., space hold), while the arms/hands kept manipulating the game (with gestures; see participant on the right in Figure 24A and participant on the right in Figure 24B). When the game was removed from the table by the facilitator, the successive movement of the participant towards the seat was initiated by the head (*none of the target head movements*) and then progressed via the torso (*none of the target torso movements*) into the arms/hands (*none of the target arms/hands movements*).

With regards to the appearance of the same pattern in relation to the partner (see participant on the left in Figure 24B), the movement was again: (i) initiated by the head (*gaze toward partner*) and then sequenced into arms/hands (*reach out partner*) via the torso (*lean in partner*), (ii) partially locked in space (space hold of head and torso, with various gestures of the arms/hands), and (iii) concluded with a successive movement of the participant towards the seat initiated by the head (*none of the target head movements*) and transmitted to the arms/hands (*none of the target arms/hands movements*) by the torso (*none of the target torso movements*).

We observed *three variations* of the pattern *successive – space hold* (with various gestures) – *successive*. The first occurred when participants addressed the game or the partner exclusively with the head, held the head in the same position in space (see participant on the left in Figure 24C, and participant on the right in Figure 25), and then moved the head away from the focus. The second appeared when they addressed the game or the partner with a successive movement of the head and the torso, which was then held in the same position in space, and subsequently dissolved. The third took place when participants addressed the game or the partner with a *sequential* movement of the head and arms/hands, which was then held in the same position in space without further activation of the torso (see participant on the right in Figure 25), and finally sequentially dismantled. In this latter case, it would be more correct to rename the pattern *sequential – space hold* (with various gestures) – *sequential*.

Figure 23. Main pattern of body part organization of focused attention visible from the timeline of Observer XT. In the three pictures, the movement of the participant on the right towards the game is initiated by the head (gaze toward game, leftmost picture), and sequenced into arms/hands (manipulate game, rightmost picture) via the torso (near reach/lean toward game, central picture). GG= gaze toward game; NRLTG= near reach/lean toward game; MG= manipulate game

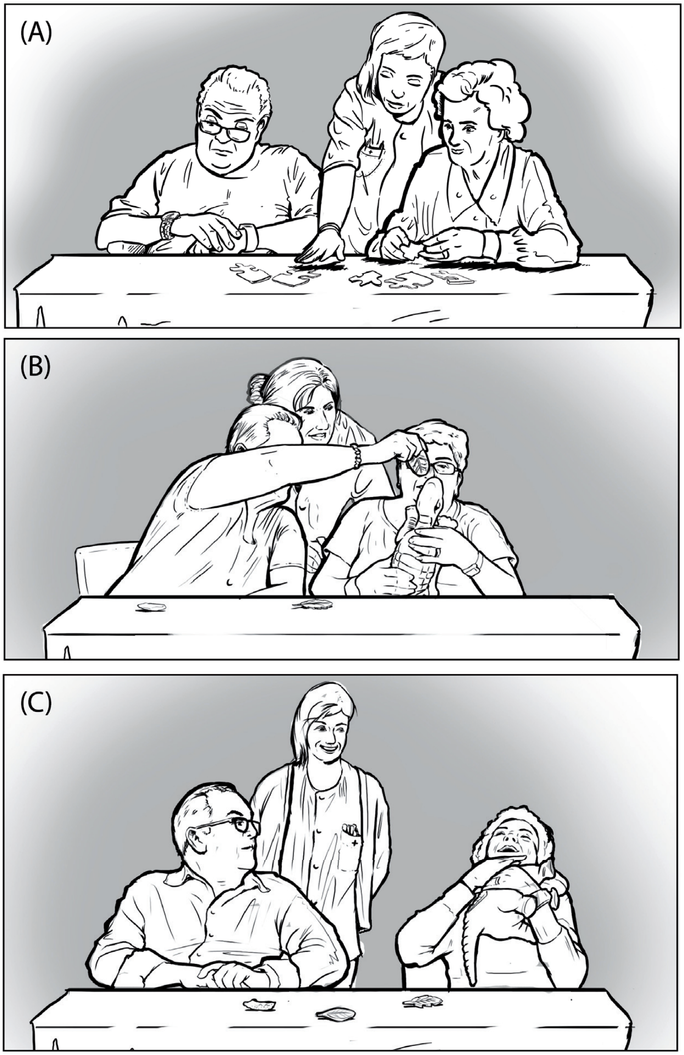


Figure 24. Patterns of body part organization of focused attention and valence

From the described patterns of body part organization in the indicators of focused attention, we postulated that: (1) the head had a leading role in focused attention, (2) the movement of the head directed towards the game and the partner could progress into arms/hands via the torso, (3) the movement of the head toward the game and the partner could progress into the torso without the further involvement of arms/hands, and (3) the movement of the head toward the game and the partner could be sequenced into arms/ hands even without the involvement of the torso.

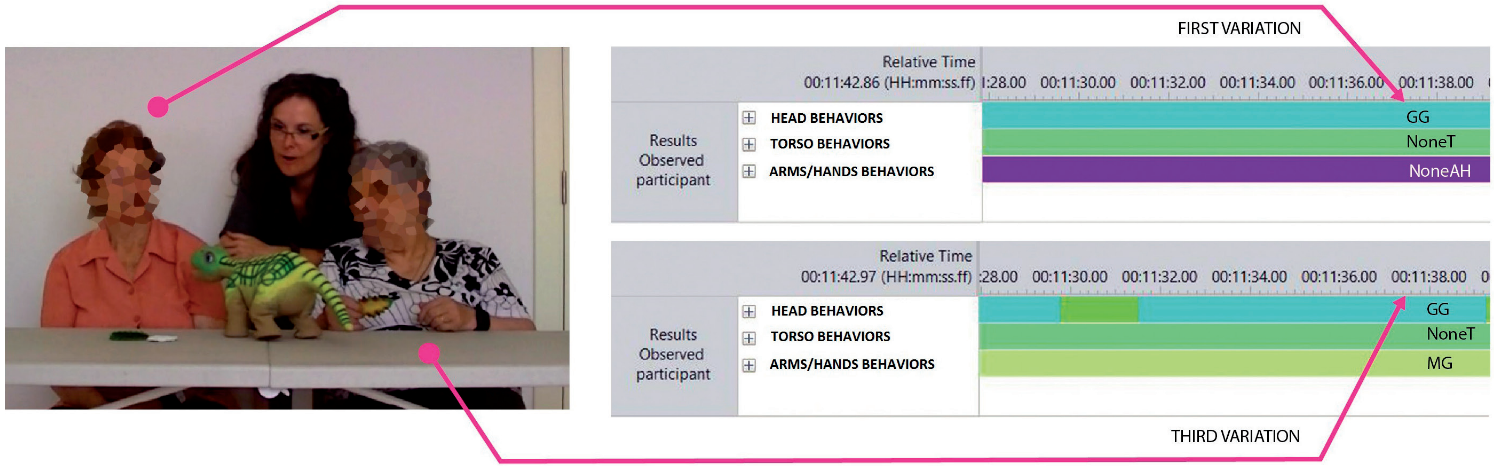


Figure 25. Variations to the main pattern of body part organization of focused attention visible from the timeline of Observer XT. The participant on the left addresses the game exclusively with the head (GG= gaze toward game), while the participant on the right addresses the game both with the head (GG) and with the arms/hands (MG= manipulate game)

8.2.2 Patterns of Body Part Organization: Valence

Across activities, we observed one main pattern of body part organization in the behaviors of valence: *successive – space hold* (with or without various gestures). This could involve all body parts or just some of them. The amount of body parts involved in the expression of valence greatly depended on the focus of the attention of the participant. When this was an inanimate object, such as a piece of the jigsaw puzzle, valence was mostly expressed through the gestures of one single body part, the head (i.e., *gestural support*). On the contrary, when the focus of the attention was a social interactor like the partner or a social robot like Pleo, the expression of valence could get to involve all body parts, the head, the torso, and the arms/hands (i.e., *gestural support*, *postural support*, *quality of gesture*).

The first element of the pattern - the *successive* movement – appeared for example when the participant directed the head towards Pleo (*gaze toward game*), smiled at it (*positive gestural support*), initiated the approach towards Pleo with the chest (*near reach/lean toward game*), embraced the robot with both arms/hands (*positive postural support*), lifted the robot to bring it close to the torso (*manipulate game*), and hugged it (*positive quality of gesture*). Also, it appeared when the participant directed the head towards the partner (*gaze toward partner*), smiled at him/her (*positive gestural support*), leaned sideways towards the partner (*lean in partner*), and embraced him/her with both arms/hands (*positive postural support*, *reach out partner* with *positive quality of gesture*).

The second element of the pattern – the *space hold* (with or without various gestures) – occurred once the successive movement was completed. Sometimes, it consisted of a *space hold* of the three body parts in the affective behavior (e.g., hug the robot; see the participant on the right in Figure 24C), other times of a space hold of two body parts with gestures (e.g., hug the robot while cradling it, hug the partner while patting his/her back).

As anticipated, the expression of valence did not always involve all body parts. Indeed, we identified two alternatives to the *successive – space hold* (with or without various gestures) pattern. The first appeared when participants expressed valence just with head gestures (i.e., *gestural support*, e.g., smile, laughter). The second appeared when they expressed valence with head gestures and arms/hands gestures (i.e., *gestural support* and *quality of gesture*, e.g., stroke the robot, or hit the robot).

From the described patterns of body part organization in the behaviors of valence, we assumed that: (1) the head had a leading role in the expression of valence, (2) the expression of valence could be initiated by *gestural support* and progressively intensified by *postural support* and *quality of gesture*, and (3) the expression of valence could be initiated by *gestural support* and intensified by *quality of gesture* alone without the involvement of *postural support*. These latter assumptions regarded situations where the focus of the attention of the participant was the partner or the robot, but not the board games.

8.3 Hypothetical Regression Paths

In summary, the analysis of the patterns of body part organization of focused attention and valence brought us to the following theoretical assumptions:

1. The *head* has a leading role in focused attention.
2. The *head* might initiate the movement of the *torso* towards the game or partner.
3. The movement of the *torso* towards the game or the partner can be then sequenced into *arms/hands*.
4. The *head* alone might initiate the movement of the *arms/hands* towards the game or the partner.
5. The *gestural support* might be intensified by the *postural support*.
6. The *postural support* could be further intensified by the *quality of gesture*.
7. The *gestural support* alone might be intensified by the *quality of gesture*.

In order to test our theoretical assumptions, we transformed them into seven hypothetical relations between the indicators of focused attention and valence (see Appendix G for a SEM glossary):

H1. The indicator *gaze toward activity* is an exogenous variable (i.e., a variable whose value is not dependent on the values of other variables in the model).

H2. The indicator *gaze toward activity* has a direct effect on the indicator *lean toward activity*.

H3. The indicator *lean toward activity* has a direct effect on the indicator *reach out activity*.

H4. The indicator *gaze toward activity* has a direct effect on the indicator *reach out activity*.

H5. The indicator *gestural support* has a direct effect on the indicator *postural support*.

H6. The indicator *postural support* has a direct effect on the indicator *quality of gesture*.

H7. The indicator *gestural support* has a direct effect on the indicator *quality of gesture*.

On top of these seven hypothetical relations, we added three additional ones. These were aimed at disclosing relationships between indicators of focused attention and valence pertaining to the same body part (i.e., head, torso, arms/hands).

H8. The indicator *gaze toward activity* has a direct effect on the indicator *gestural support*.

H9. The indicator *lean toward activity* has a direct effect on the indicator *postural support*.

H10. The indicator *reach out activity* has a direct effect on the indicator *quality of gesture*.

8.4 Results of SEM

The model in Figure 26 depicts all the hypothetical relationships between variables (H2-H7: blue arrows, H8-H10: red arrows). We tested this model using SEM with the software SPSS Amos 22.0. The dataset used to test this model was the same used to test the theory-driven model of engagement. The observed variables in blue – *gaze toward activity*, *lean toward activity*, and *reach out activity* – are the indicators of focused attention. The observed variables in violet – *gestural support*, *postural support*, and *quality of gesture* – are the indicators of valence.

We ran SEM twice using the data from both activities (N= 84). The first time, we calculated the Mahalanobis distance and identified the farthest observations from the centroid ones. The second time, we fitted the model excluding the outlier observations (N= 7). The

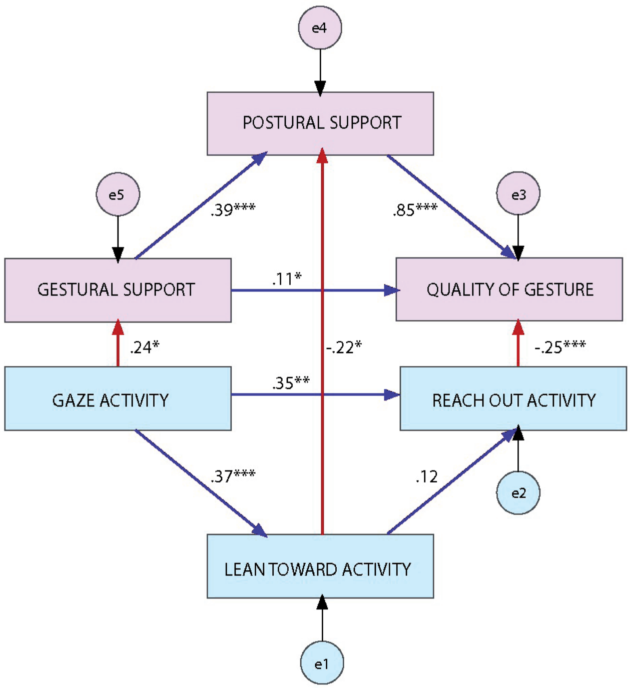


Figure 26. Regression paths between indicators of focused attention and valence. Significance level of path estimates: † fixed factor, *< .05, **< .01, *< .001**

model proved to be an excellent fit for the data (X^2(6, *N*= 77)= 5.866, *p*= .436; *RMSEA*= .000; *NFI*= .970; *CFI*= 1.000; *RFI*= .896; *PNFI*= .277) and almost all the hypothesized relations (H1-H10) between observed variables were significant (see Table 20 and Figure 26). H1 was confirmed by the goodness of fit of the model. H2-H10 (except H3) were confirmed both by the goodness of fit of the model and by the significance of the path estimates. The only postulated relation between variables that was not significant was the one between *lean toward activity* and *reach out activity* (H3). We ran two regression analyses to figure out whether this result depended on the behaviors directed towards the game or on those directed towards the partner. The results disclosed that *near reach/lean toward game* had a significant effect on *manipulate game* (β= .246, *t*(76)= 2.201, *p*< .05) and *lean in partner* had a significant effect on *reach out partner* (β= .231, *t*(76)= 2.057, *p*< .05). Compared to regression analysis, SEM calculates also an error term for the variables. Thus, the lack of a significant result for this relation depended on the size of the error term of the two variables and not on the lack of relationship between them.

With regards to the relationships between focused attention and valence, *gaze toward activity* and *gestural support* were connected by a positive regression, while *lean toward activity* and *postural support* and *reach out activity* and *quality of gesture* were connected by negative regression paths. This result might be a consequence of the fact that, while *gestural support* is superimposed on *gaze toward activity*, *postural support* and *quality of gesture* are in fact subsets of *lean toward activity* and *reach out activity*. They are descriptions of how *lean toward activity* and *reach out activity* are performed.

Table 20. Path estimates of model

	Hypothesized path	Estimate	S.E.	C.R.	*p-value*	Hypothesis supported yes/no
H2	GAct → LTAct	.372	.680	3.499	***< .001	yes
H3	LTAct → RoAct	.122	.060	1.079	> .05	no
H4	GAct → RoAct	.349	.386	3.099	**< .01	yes
H5	GAct → Gest Sup	.240	.099	2.159	*< .05	yes
H6	RoAct → QoGest	-.255	.017	-5.998	***< .001	yes
H7	Gest Sup → Post Sup	.390	.106	3.750	***< .001	yes
H8	Post Sup → QoGest	.845	.068	18.563	***< .001	yes
H9	Gest Sup→ QoGest	.112	.069	2.444	*< .05	yes
H10	LTAct → Post Sup	-.218	.015	-2.102	*< .05	yes

Significance level: *< .05, **< .01, ***< .001.

8.5 Discussion

We called the model presented in this chapter, the Evidence-based MODel of Engagement-related Behavior (EMODEB). The EMODEB is a model that describes the spread of behavior across body parts and details the relationships between the indicators of focused attention and valence of the theory-driven model of engagement. The EMODEB confirmed what postulated in chapter 7, that the indicators of *focused attention* and *valence* in the theory-driven model of engagement were connected by regression paths. Moreover, it revealed that *head behaviors* (i.e., *gaze toward activity* and *gestural support*) and *arms/hands behaviors* (i.e., *reach out activity* and *quality of gesture*) are more crucial to measure engagement than *torso behaviors* (i.e., *lean toward activity* and *postural support*). Indeed, while the former constituted respectively the starting and the accomplishment of focused attention and valence, the latter just energized the passage of movement from the head to the arms/hands[9].

8.5.1 Consequences for a Model of Engagement

In the economy of a good modeling of the construct of engagement for dementia, one of the principal consequences of the EMODEB is that head behaviors, torso behaviors, and arms/

9 The demonstration of this assertion is that the movement of the head can be sequenced into arms/hands without the involvement of the torso.

hands behaviors do not have the same value in terms of focused attention and valence. Indeed, they are organized in a hierarchical way.

In order to account for the hierarchical structuring of the indicators of focused attention and valence, we computed the weighted averages of *gaze toward activity* (weight= .50), *lean toward activity* (weight= .10), and *reach out activity* (weight= .40), and of *gestural support* (weight= .50), *postural support* (weight= .10), and *quality of gesture* (weight= .40) and used them to quantify the two components of engagement in further versions of the model. The weights were chosen based on the EMODEB according to a three-point rationale: (i) as the behaviors of the *head* and *arms/hands* were the building blocks of focused attention and valence, they were assigned similar weights, (ii) as *head* behaviors were the initiators of focused attention and valence, they were assigned slightly higher weights than those assigned to arms/hands behaviors, and (iii) as *torso* behaviors had a marginal role in focused attention and valence, they were assigned much smaller weights than those assigned to head and arms/hands behaviors.

8.5.2 Consequences for the Measurement of Behavior

In the context of behavior measurement, the EMODEB complements the knowledge drawn in chapter 6 from the scoring of the behaviors in the ELICSE. Indeed, it has one important corollary: it enables us to make inferences regarding the meaning of different body configurations in terms of engagement intensity. According to the hierarchical ordering of behaviors in the EMODEB, we can score focused attention and valence along two categorical scales whose points are operationalized through body configurations. With regards to *focused attention*, we consider the person with dementia:

1. *Not attentive* when all three body parts – head, torso, and arms/hands – are directed toward a focus other than the activity (*none of the target head movements*, *none of the target torso movements*, and *none of the target arms/hands movements*).

2. *Passively attentive*, when the head is the only body part directed toward the activity, while the torso and the arms/hands are directed somewhere else (*gaze toward activity*, *none of the target torso movements*, and *none of the target arms/hands movements*).

3. *Ready to participate*, when the head and the torso, but not the arms/hands are directed toward the activity (*gaze toward activity*, *lean toward activity*, and *none of the arms/hands movements*).

4. *Moderately participating*, when the head and the arms/hands, but not the torso are addressing the activity (*gaze toward activity*, *none of the torso movements*, and *reach out activity*).

5. *Fully participating,* when all three body parts – head, torso, and arms/hands – are directed toward the activity (*gaze toward activity*, *lean toward activity*, and *reach out activity*).

With regards to valence, we consider the person with dementia as expressing:

1. *Strong negative valence*, when all three body parts – head, torso, and arms/hands – are involved in the expression of negative valence (*negative gestural support*, *negative postural support*, and *negative quality of gesture*).

2. *Moderate negative valence*, when the head and the arms/hands, but not the torso, are involved in the expression of negative valence (*negative gestural support*, *no postural support,* and *negative quality of gesture*).

3. *Mild negative valence*, when the head, but not the torso and the arms/hands, are involved in the expression of negative valence (*negative gestural support*, *no postural support*, and *no quality of gesture*).

4. *Neutral valence*, when none of the body parts is involved in the expression of negative or positive valence (*no gestural support*, *no postural support*, and *no quality of gesture*).

5. *Mild positive valence*, when the head, but not the torso and the arms/hands, are involved in the expression of positive valence (*positive gestural support*, *no postural support*, and *no quality of gesture*).

6. *Moderate positive valence*, when the head and the arms/hands, but not the torso, are involved in the expression of positive valence (*positive gestural support*, *no postural support*, and *positive quality of gesture*).

7. *Strong positive valence*, when all three body parts – head, torso, and arms/hands – are involved in the expression of positive valence (*positive gestural support*, *positive postural support*, and *positive quality of gesture*).

The scale of focused attention shares some similarities with the MPES (Judge et al., 2000) and the item attention of the OME (Cohen-Mansfield et al., 2009). However, it refers to specific body configurations in order to operationalize the different levels of engagement, and features also the behaviors of the torso as meaningful in engagement terms. The scale of valence shares some similarities with the item attitude of the OME. However, it does not rely on the subjective meaning of behavior to judge valence, but clearly maps each point of the scale to a specific body configuration.

These scales clearly need to go through a process of validation. However, once validated, they might allow researchers to observe the temporal progression of focused attention and valence throughout an activity session and code engagement in its progression, as a time-series data. For the time being and in the context of this thesis, all the constructs related to engagement are static and calculated globally for the whole session.

8.6 Conclusions

This chapter described the process that brought to light the existence of regression paths connecting the indicators of focused attention and valence of the theory-driven model of engagement. It outlined: (i) the identification of patterns in the way behavior spread across body parts, (ii) the translation of these patterns into regression paths connecting the indicators of focused attention and affect, and (iii) the testing of the regression paths through SEM.

The main result of the chapter is the EMODEB, a model that describes the natural flow of engagement-related behavior across body parts in people with dementia. According to the EMODEB, the indicators of focused attention and valence are organized hierarchically and are not provided with the same measurement value. Thus, in a model of engagement, they need to be assigned weights and be compiled through weighted averages. One of the incidental contributions of the EMODEB is that it enabled us to describe the intensity of focused attention and valence in terms of body configurations and hence paved the way for the measurement of the temporal progression of engagement-related behavior in dementia.

In the next chapter, we present the development another measure of engagement, *quantity of movement*. This measure was partially inspired by the literature on the diagnosis of motivational disorders and partially inspired by the results of this chapter.

9

Measure Development: Quantity of Movement

9.1 Introduction

In this chapter, we describe the development of another measure of engagement for people with dementia, quantity of movement. This measure was not suggested by the literature on engagement. Instead, it was partially inspired by the hierarchical ordering of behaviors in the EMODEB and partially suggested by the knowledge drawn from studies regarding the diagnosis of apathy in dementia. In this chapter, we first explain the rationale behind the utilization of quantity of movement as a measure of engagement. Then, we detail the process of elimination of the invalid sessions of measurement and of feature extraction. Last, we present the results of concurrent validity and discuss them in detail.

9.2 Measurement Rationale

The introduction of quantity of movement as a measure of engagement had a twofold inspiration. First, it was suggested by a series of diagnostic studies regarding apathy in dementia. Second, it was a result of the EMODEB, specifically of the ranking of the behaviors of focused attention and valence belonging to different body parts (i.e., head, torso, arms/hands).

With regards to diagnostic studies, a number of recent investigations uncovered that apathy can be more easily diagnosed by measuring the quantity of movement that people with dementia produce over long periods of time (hours, days, weeks). This could be pursued with a wrist-worn actigraph (i.e., an accelerometer) placed on the non-dominant wrist. Apparently, people with dementia-associated apathy have lower mean motor activity with respect to people affected by dementia alone (Volkers et al., 2003; David et al., 2010; David et al., 2012; Kuhlmei et al., 2013). As the inextricable inverse relationship between apathy and engagement in dementia is well-known to the literature (Robert et al., 2010; Dechamps et al., 2010; Mulin et al., 2011; Cohen-Mansfield et al., 2011), we assumed that quantity of movement on the wrist could increase as a result of engagement in the same way as it decreased as a consequence of apathy.

With regards to the EMODEB, one of the main corollaries of the model was that arms/hands behaviors constituted the accomplishment of the movement performed by the person with dementia to reach the game or the partner and that the highest intensity of focused attention and valence was represented by body configurations involving the arms/hands and not just the head and the torso. As a consequence of this outcome, we assumed that we could measure the amount of participation of the person with dementia in the activity by assessing the quantity of movement on the wrist.

Drawing inspiration from these considerations, we exploited the database of accelerometer data collected with the E4 wristband during the experimental sessions and extracted from the signal features related to quantity of movement. Then, we attested the concurrent validity of quantity of movement as a measure of engagement by correlating the features extracted from the accelerometer signal with the gold standard observational measures, the OME and the OERS.

9.3 Identification and Exclusion of Invalid Measurements

In the database of accelerometer signals, just one session was missing and the others were all valid in terms of measurement. However, due to problems encountered in placing the wristband on the non-dominant wrist of participants (e.g., bruises due to dialysis), some accelerometer data were collected from the dominant wrist (15 sessions overall), and thus could not be used for further analyses. From the diary that we kept during data collection (see Chapter 5), we were able to retrieve the sessions collected on the dominant wrist and discard them. As a result, the database of accelerometer data employed for analysis was composed of 68 sessions of data: 32 of game-based cognitive stimulation and 32 of robot-based free play.

9.4 Feature Extraction

The accelerometer features of quantity of movement were extracted from one window: the *activity phase* (duration: ~20-25 minutes). This was the phase of the experimental session where participants completed the three board games or played with the pet robot Pleo. In order to extract features from the window, we inputted in Matlab the same synchronization files used for EDA.

We did not perform any pre-processing of the accelerometer signal before feature extraction. This was for two reasons: (1) the features to which we were interested concerned the amount of variation in the signal and not the specific qualities of its behavior and (2) the participants did not displace nor they moved abruptly and excessively during activities, thus the amount of noise in the signal was rather low.

With regards to the selection of accelerometer features, we could not rely on previous literature. Indeed, Volkers et al. (2003), David et al. (2010), David et al. (2012), and Kuhlmei et al. (2013) did not extract features from the raw accelerometer signal, but relied on the counts of supra-threshold movements on the wrist provided by an actigraph. We assumed that the most adequate accelerometer feature of quantity of movement could be the signal magnitude area of the acceleration. This gauges the amount of variation in the accelerometer signal within a certain window. We extracted two features from the accelerometer signal with the software Matlab: (i) the signal magnitude area of the module of the three axes (SMA Acc_M) following Equation (a) and (2) the summation of the signal magnitude areas of the three axes (SMA Acc_S) as defined in Equation (b). The former is more related to the general quantity of movement, the latter is more related to the variability of movements.

$$SMA_M = \sum_{i=1}^{T} \left| \sqrt{x_i^2 + y_i^2 + z_i^2} \right| dt \qquad \text{(a)}$$

$$SMA_S = \sum_{i=1}^{T} |x_i| dt + \sum_{i=1}^{T} |y_i| dt + \sum_{i=1}^{T} |z_i| dt \qquad \text{(b)}$$

X*i*, Y*i* , Z*i* are the acceleration on the X, Y and Z axes in the *i* sample. *T* is the length of the window measured in number of samples.

9.5 Concurrent Validity

To verify the concurrent validity of quantity of movement as a measure of engagement, we extracted the accelerometer features from all the valid sessions in the database. Then, we ran a Spearman rank correlation (one-tailed, listwise exclusion of cases, N_{GBCS}= 34, N_{RBFP}= 34) between SMA Acc_M and SMA Acc_S and the items of the OME and OERS (see Table 21). In the game-based cognitive stimulation:

1. The item *attention* of the OME was significantly positively correlated with SMA Acc_M (*rs*(32)= .407, *p*= .008) and SMA Acc_S (*rs*(32)= .350, *p*= .021).

2. The item *attitude toward game* of the OME was significantly positively correlated with SMA Acc_M (*rs*(32)= .431, *p*= .005) and SMA Acc_S (*rs*(32)= .366, *p*= .017).

3. The item *attitude toward partner* of the OME was significantly positively correlated with SMA Acc_M (*rs*(32)= .350, *p*= .021) and SMA Acc_S (*rs*(32)= .293, *p*= .046).

4. The item *cognitive difficulty* of the OME was significantly negatively correlated with SMA Acc_M (*rs*(32)= -.451, *p*= .004) and SMA Acc_S (*rs*(32)= -.372, *p*= .015).

With regards to robot-based free play:

1. The item *attitude toward game* of the OME was significantly positively correlated with SMA Acc_M (*rs*(32)= .476, *p*= .002) and SMA Acc_S (*rs*(32)= .472, *p*= .002).

2. The item *pleasure* of the OERS was significantly positively correlated with SMA Acc_M (*rs*(32)= .415, *p*= .007) and SMA Acc_S (*rs*(32)= .426, *p*= .006).

We found two additional significant correlations in the robot-based free play. However, as they involved the items *anger* and *sadness* of the OERS, we excluded them due to the scarce variability of these items throughout the whole dataset.

Table 21. Concurrent validity of quantity of movement for game–based cognitive stimulation and robot–based free play: correlations tested (in bold the significant ones: *≤.05, **<.01, *<.001)**

GAME-BASED COGNITIVE STIMULATION				ROBOT-BASED FREE PLAY			
CORRELATION OME & OERS/ QUANTITY OF MOVEMENT*[1]		SMA Acc_M	SMA Acc_S	CORRELATION OME & OERS/ QUANTITY OF MOVEMENT*[2]		SMA Acc_M	SMA Acc_S
Attention	*r*(s)	**.407**	**.350**	Attention	*r*(s)	.164	.154
	p-value	****.008**	***.021**		*p*-value	.176	.192
Attitude game	*r*(s)	**.431**	**.366**	Attitude game	*r*(s)	**.476**	**.472**
	p-value	****.005**	***.017**		*p*-value	****.002**	****.002**
Attitude partner	*r*(s)	**.350**	**.293**	Attitude partner	*r*(s)	.036	.013
	p-value	***.021**	***.046**		*p*-value	.419	.470
Cognitive difficulty	*r*(s)	**-.451**	**-.372**	Pleasure	*r*(s)	**.415**	**.426**
	p-value	****.004**	***.015**		*p*-value	****.007**	****.006**
Pleasure	*r*(s)	.184	.210	Anger	*r*(s)	**-.492**	**-.474**
	p-value	.148	.117		*p*-value	****.002**	****.002**
Anger	*r*(s)	.019	.032	Anxiety/Fear	*r*(s)	-.183	-.182
	p-value	.457	.429		*p*-value	.151	.152
Anxiety/Fear	*r*(s)	.044	.071	Sadness	*r*(s)	**-.396**	**-.362**
	p-value	.402	.345		*p*-value	****.010**	***.018**
Alertness	*r*(s)	-.022	-.056	Alertness	*r*(s)	.167	.157
	p-value	.450	.377		*p*-value	.172	.187

*[1] The item *sadness* of the OERS did not yield any significant result, hence it does not appear in the list

*[2] The item *cognitive difficulty* of the OME did not yield any significant result, hence it does not appear in the list

9.6 Discussion

The analysis of concurrent validity revealed a good correlation tendency between the features of quantity of movement and the items of the OME and OERS. Quantity of movement was highly correlated with all the items of the OME in game-based cognitive stimulation. Moreover, it was closely related to the items of the OME and OERS depicting positive valence in robot-based free play.

As suggested at the beginning of the chapter, quantity of movement increased as a result of engagement as much as it decreased as a consequence of apathy. Indeed, in our sample, participants with motivational disorders (i.e., apathy and depression) showed much lower quantity of movement during activities than participants without such disorders (see Figure 27 and Appendix F for the results of the analysis), a dynamic that was not equally visible when participants were clustered according to dementia severity (mild vs moderate).

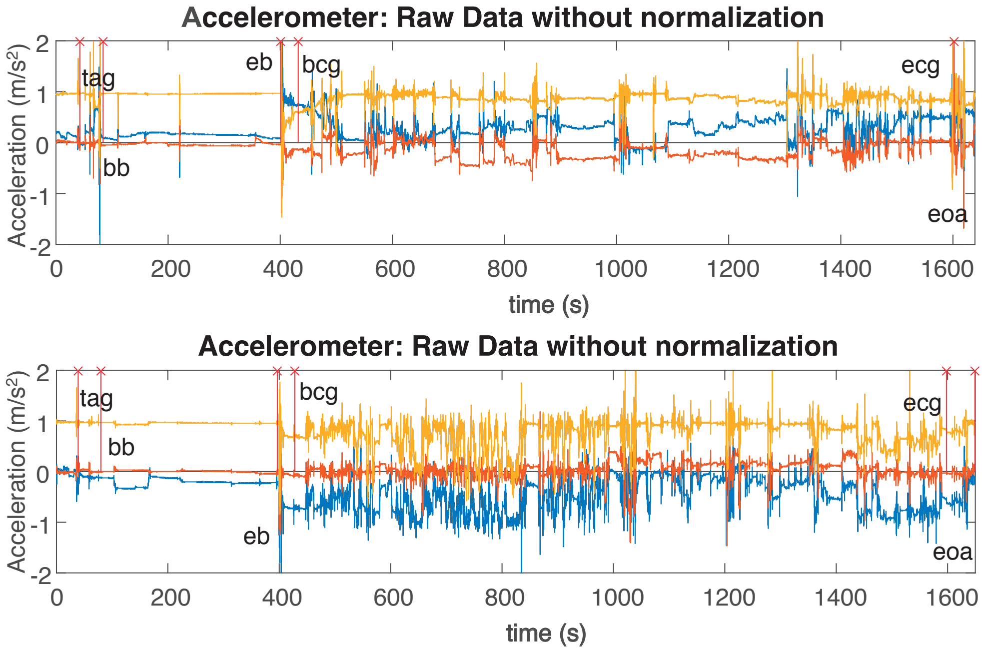


Figure 27. Plots of effects of motivational disorders on quantity of movement. In the top plot, the accelerometer signal of a participant with apathy, in the bottom plot, the accelerometer signal of a participant without motivational disorders. The two participants were involved in the same activity session

The correlations with the OME and OERS did not just provide information on the concurrent validity of quantity of movement, but also confirmed some of the dynamics observed and discussed in the previous chapters, for instance, that in the game-based cognitive stimulation, the interaction with the partner entailed a high degree of engagement in the activity, while the perceived cognitive difficulty of the game was a deterrent to engagement. Also, they confirmed that, in robot-based free play, when people were engaged, they were so in a positive affective manner.

As quantity of movement measures the amount of movement of the arms/hands during activities, we considered it a good measure of attention. In the final model of engagement presented in the next chapter, the two features that we extracted from the accelerometer signal are used as indicators of the latent variable *participation* together with the weighted average of focused attention obtained by compiling *gaze toward activity*, *lean toward activity*, and *reach out activity*.

9.7 Conclusions

In this chapter, we presented the last measure of engagement deployed in the context of this dissertation, quantity of movement. As a measure, quantity of movement gauges the amount of arms/hands behavior of people with dementia during activities with a wrist-worn triaxial accelerometer. In the chapter, we described the rationale that guided us in the conception of the measure, the process of feature extraction, and the results of concurrent validity. As quantity of movement achieved a good concurrent validity, it will be introduced in the final model of engagement that we describe and test in the next chapter of the thesis.

10

Construct Modeling: The Final Model of Engagement

10.1 Introduction

In this chapter, we present and test the final model of engagement for dementia. First, we describe the process through which we modified the theory-driven model of engagement to accommodate the results of chapter 8 and 9 and obtained the final model of engagement. Second, we define the components of engagement in the model, their measures, and the relationships that they entertain. Last, we test the final model of engagement with SEM, and discuss its results and implications for the measurement of engagement in dementia.

10.2 Stock of the Situation

In chapter 7, we tested the theory-driven model of engagement and uncovered that the model abridging the knowledge drawn from the literature was not a good structural representation of engagement in dementia. Consequently, we hypothesized the existence of four sources of misspecification whose testing could bring to a model of engagement for dementia more consistent with the data. The hypothetical sources of misspecification of the theory-driven model of engagement were: (i) the presence of hidden regression paths between the indicators of *focused attention* – *gaze toward activity*, *lean toward activity*, and *reach out activity*, (ii) the existence of hidden regression paths between the indicators of *valence* – *gestural support*, *postural support*, and *quality of gesture*, (iii) the bivariate distribution of *valence* and *arousal*, and (iv) the instability of the correlation between *focused attention* and *valence.*

In chapter 8, two of these sources of misspecification were confirmed, namely those regarding the behavioral/expressive level of measurement. Indeed, we brought to light the existence of regression paths connecting the indicators of *focused attention* and *valence* of the theory-driven model of engagement. These paths described the way engagement-related behavior spread across body parts, and were condensed in a dedicated model, the EMODEB. According to the EMODEB, the indicators of focused attention and valence were organized hierarchically and hence did not have the same measurement value. In the final model of engagement that we present in this chapter, *focused attention* was computed as the weighted average of *gaze toward activity*, *lean toward activity*, and *reach out activity*, while *valence* was calculated as the weighted average of *gestural support*, *postural support*, and *quality of gesture.* The entity of the weights was determined based on the ranking of the indicators in the EMODEB as detailed in chapter 8.

In chapter 9, we presented a further measure of engagement, quantity of movement. This was inspired by the EMODEB and further corroborated by a number studies on the diagnosis of apathy in dementia. Quantity of movement was measured by extracting two features from the triaxial accelerometer signal, SMA Acc_M and SMA Acc_S. As these features assessed the amount of arms/hands behavior of participants, they were considered particularly suited to quantify the degree of participation in the activity. In the final model of engagement, we employed SMA Acc_M and SMA Acc_S to measure the latent variable *participation* together with the now observed variable *focused attention*.

With regards to the last two potential sources of misspecification, as they regarded the relationships between the components of engagement, we utilized them to inform the structuring of the final model of engagement. The postulated bivariate distribution of *valence* and *arousal* entailed that arousal could grow not just as a result of positive valence, but also as a consequence of negative valence. In the final model of engagement, we removed the latent variable *core affect* and connected arousal and valence with a covariance path. This way the relationship between *arousal* and *valence* was left free to turn negative, positive, or to be non-significant.

For what concerns the instability of the correlation between *focused attention* and *valence*, it implied that focused attention could be correlated as well as uncorrelated with valence. Indeed, during activities, we observed situations where participants were strongly focused on the activity, but displayed behaviors of negative valence, situations where participants were paying attention to the activity, but showed neutral valence, and also situations where the allocation of attentional resources to the activity was the trigger of positive valence. In the final model of engagement, the latent variable *participation* and the now observed variable *valence* were connected with a covariance path. This way they were left free to correlate or not.

10.3 Final Model of Engagement

In summary, with respect to the theory-driven model of engagement, in the final model of engagement:

1. *Focused attention* was computed as the weighted average of *gaze toward activity*, *lean toward activity*, and *reach out activity* and thus ceased to be a latent variable.

2. *Valence* was calculated as the weighted average of *gestural support*, *postural support*, and *quality of gesture* and thus ceased to be a latent variable.

3. SMA Acc_M and SMA Acc_S were added to the model and employed to measure the latent variable *participation* together with the now observed variable *focused attention.*

4. The latent variable *core affect* was removed and *arousal* and *valence* were connected with a covariance path.

5. The latent variable *participation* and the now observed variable *valence* were connected with a covariance path.

Also, but this was not suggested by the analyses:

6. The latent variable *participation* was connected to the latent variable *arousal* by a covariance path.

This last modification of the theory-driven model of engagement was suggested by the fact that, according to the literature, arousal does not grow just as a result of episodes of excitement and anxiety, but also as a consequence of attentional allocation (Andreassi, 2010). As participation is the voluntary focus of attention on a limited stimulus field that is proactively given, we assumed that it could correlate with arousal, and thus added the covariance path at point 6.

In conclusion, the only variable that remained untouched in the passage from the theory-driven model of engagement to the final model of engagement was arousal. This was still measured with the latent variables *tonic* and *phasic* EDA, which were respectively assessed through the features MEAN EDA, STD EDA, and SUM H EDA and through the features SMA EDA and NPR EDA. The final model of engagement that we have described is presented in Figure 28.

10.4 Hypothesized Relations between Components

According to the changes made to the theory-driven model, engagement is not anymore a latent variable from whom the different components of engagement derive, but a correlational concept arising from the relationships between three components: participation, valence, and arousal. *Participation* refers to the act of taking part in the activity by proactively manipulating the game and interacting with the partner(s) and ranges from *passive* to *full participation*. *Arousal* is the degree of sympathetic activation of the participant during the activity and ranges from *low* to *high arousal*. Last, *valence* is the hedonic tone expressed by the participant during the activity and ranges from *negative* to *positive valence*. Following the arguments presented in the previous sections, we hypothesized that in engagement:

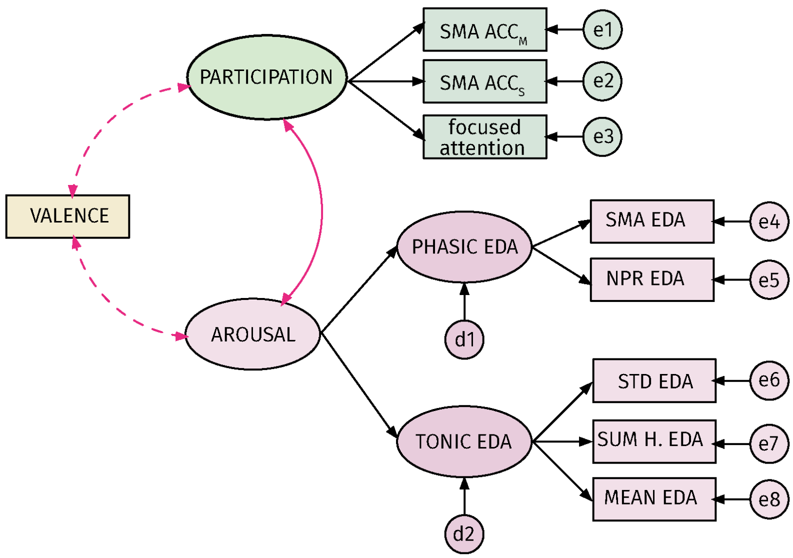


Figure 28. Final model of engagement. The dotted curvilinear paths refer the correlations that can vary across activities, the continuous curvilinear path refers to the correlation that is stable across activities

a. The relationship between *arousal* and *participation* remains stable. It is always positive.

b. The relationship between *arousal* and *valence* is mutable. For instance, it can be *positive* in activities where valence is mostly positive, it can be *negative* in activities where valence is mostly negative, and it can be *non-significant* in activities where valence is mostly neutral, in activities where valence spans from negative to positive, or in activities where arousal or valence do not vary consistently.

c. The relationship between *participation* and *valence* is mutable. For instance, it can be *positive* in activities where participation is mostly positive, it can be *negative* in activities where participation is mostly negative, and it can be *non-significant* in activities where participation is mostly neutral, in activities where participation spans from negative to positive, or in activities where participation or valence do not vary consistently.

In the model in Figure 28, the dotted curvilinear paths express the correlations that can vary across activities, while the continuous curvilinear paths refer to the correlation that are stable across activities.

10.5 Model Testing

Before running the final model of engagement with SEM, we used regression imputation to assign values to the missing cases of SMA Acc_M and SMA Acc_S (Schumacker and Lomax, 2016). This was possible as the database of accelerometer data contained a moderate number of missing cases (19%) and the data were missing completely at random (MCAR). Also, we ran some preliminary analyses. First, we calculated the sampling adequacy of the dataset (KMO= .662). Then, we ran an exploratory factorial analysis (EFA) using a principal component method of extraction and a varimax method of rotation to confirm that the indicators of the components of engagement were exactly those hypothesized by the final model. The EFA showed satisfying factor loadings for all observed variables (> .400; Wayne et al., 1982). As expected, we found three factors (See Table 22). *Factor 1* (i.e., participation) included the observed variables SMA Acc_M, SMA Acc_S, and *focused attention. Factor 2* (i.e., tonic EDA) included the observed variables STD EDA, SUM H. EDA, and MEAN EDA. Last, *factor 3* (i.e., phasic EDA) included the observed variables SMA EDA and NPR EDA. As supposed, *valence* was not grouped under any of these factors.

Table 22. Factor loadings and Cronbach's α

FEATURE/SCORE	FACTOR LOADINGS*		
	1	**2**	**3**
SMA Acc_M	**.959**	.103	.141
SMA Acc_S	**.959**	.147	.104
B. ENGAGEMENT	**.422**	.073	-.364
Cronbach's α based on standardized item= .725			
STD EDA	-.050	**.882**	.190
SUM H. EDA	.061	**.858**	.310
MEAN EDA	.145	**.885**	.068
Cronbach's α based on standardized item= .894			
SMA EDA	.094	.119	**.893**
NPR EDA	.149	. 047	**.881**
Cronbach's α based on standardized item= .815			
VALENCE	.044	.376	.213

In bold, factor loadings >.400

As a last step, we calculated the Cronbachs's α coefficients for each of the factors highlighted by the EFA. The results were very good. Indeed, all factors achieved an alpha higher than .70, which is the cutoff score for reliability (Santos, 1999). *Factor 1* had a Cronbach's α of .725, *factor 2* had a Cronbach's α of .894, and *factor 3* had a Cronbach's α of .815 (see Table 22).

The α coefficients reported here were calculated based on standardized items. We opted for this method of calculation to compensate for the mixed nature of the observed variables in the model (i.e., interval variables and ratio variables; Santos, 1999).

We ran the final model of engagement using SEM with the software SPSS AMOS 22.0. The test of the model was performed twice using the data from both activities (N= 84). The first time, we ran the model with outliers, the second time without outliers. The model proved to be an excellent fit for the data in both cases (test with outliers: X^2(24, *N*= 84)= 30.793, *p*= .160; *RMSEA*= .058; *NFI*= .937; *CFI*= .985; *RFI*= .906; *PNFI*= .625; test without outliers: X^2(24, *N*= 78)= 30.809, *p*= .159; *RMSEA*= .058; *NFI*= .935; *CFI*= .984; *RFI*= .878; *PNFI*= .499) and all the regression paths leading to the observed variables were significant (see Figure 29 and Table 23).

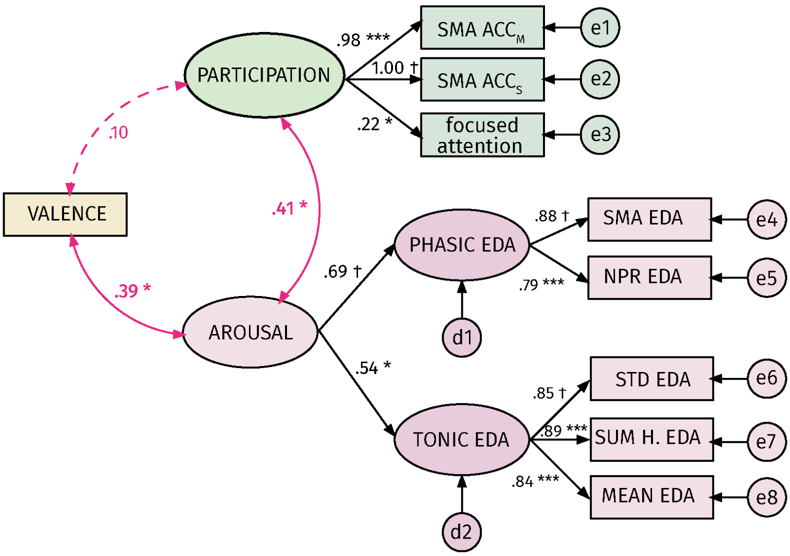


Figure 29. Final model of engagement with path estimates (without outliers). Significance level: † fixed factor, *< .05, **< .01, *< .001**

Chapter 10

Table 23. Path estimates of final model of engagement

Regression Paths	**Estimate**	**S.E.**	**C.R.**	***p*-value**
PARTICIPATION → SMA ACC_M	.976	.008	39.06	***< .001
PARTICIPATION → SMA ACC_S	†1.000	/	/	/
PARTICIPATION → FOCUSED ATTENTION	.225	56.74	2.026	*< .05
AROUSAL → PHASIC EDA	†.694	/	/	/
AROUSAL → TONIC EDA	.536	94.90	2.184	*< .05
PHASIC EDA → SMA EDA	†.875	/	/	/
PHASIC EDA → NPR EDA	.787	1174.130	4.149	***< .001
TONIC EDA → STD EDA	†.846	/	/	/
TONIC EDA → SUM H. EDA	.886	.158	9.112	***< .001
TONIC EDA → MEAN EDA	.840	.255	8.674	***< .001
Covariance Paths	**Estimate**	**S.E.**	**C.R.**	***p*-value**
AROUSAL ←→ PARTICIPATION	.405	.000	2.238	*< .05
PARTICIPATION ←→ VALENCE	.099	.010	.866	> .05
VALENCE ←→ AROUSAL	.388	.000	2.126	*< .05

† fixed factor, *< .05, **< .01, ***< .001

10.5.1 Model Testing with Single Activities

The final model of engagement was tested also for the single activities. We employed the data without outliers to perform the analyses (N_{GBCS}= 39; N_{RBFP}= 39). The model proved to be an excellent fit both for game-based cognitive stimulation (X^2(25, *N*= 39)= 24.607, *p*= .485; *RMSEA*= .000; *NFI*= .903; *CFI*= 1.000; *RFI*= .825; *PNFI*= .502; see Figure 30) and for robot-based free play (X^2(23, *N*= 39)= 29.056, *p*= .178; *RMSEA*= .080; *NFI*= .898; *CFI*= .975; *RFI*= .801; *PNFI*= .459; see Figure 31). However, in robot-based free play, we needed to add a correlation path between the error term of *focused attention* and the error term of *valence* to stabilize the model.

With regards to path estimation, most of the regression weights were significant in robot-based free play. Conversely, some of them were not significant in game-based cognitive stimulation and might have influenced the lack of significance of some of the covariance paths. As the goodness of fit was nearly perfect for this latter activity, we attributed the issue to the low sampling adequacy of the dataset (KMO_{GBCS}= .563; KMO_{RBFP}= .641), and based our conclusions on the model on the size of the estimates and not on their significance.

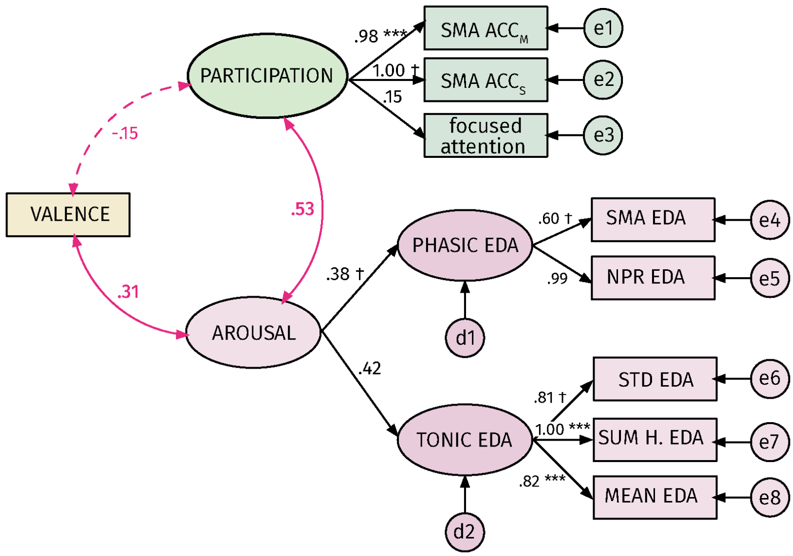


Figure 30. Path estimates for game-based cognitive stimulation. Significance level: † fixed factor, *< .05, **< .01, *< .001**

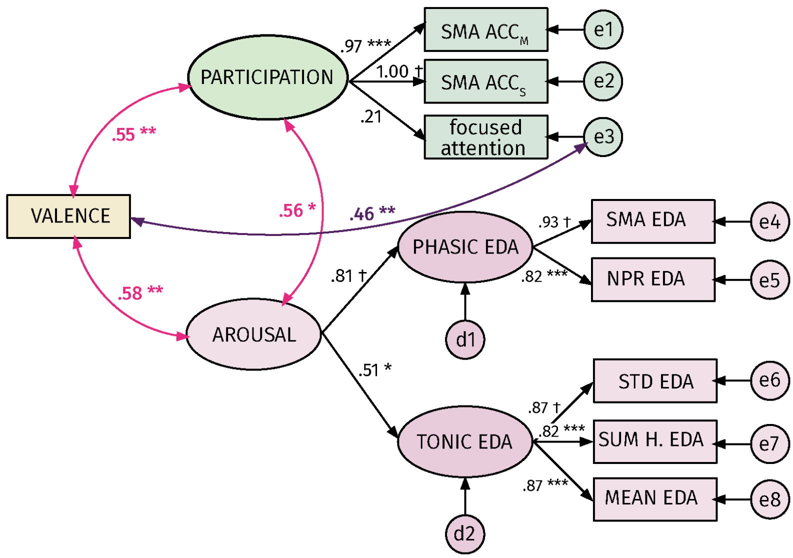


Figure 31. Path estimates for robot-based free play. Significance level: † fixed factor, *< .05, **< .01, *< .001**

10.6 Discussion

We called the final model of engagement described in this chapter the ENGAGE-DEM. The ENGAGE-DEM constitutes a framework to measure engagement in dementia in all its components without the necessity to resort to the experiential/subjective response system. The test of the model confirmed that the components of engagement are three: *participation*, *arousal*, and *valence* and that these components can be measured with the behavioral/expressive and peripheral-physiological measurement techniques presented in the dissertation: EDA, the ELICSE (in the hierarchical structure defined by the EMODEB), and quantity of movement. Moreover, the testing of the ENGAGE-DEM confirmed that the relationships entertained by the different components of engagement were those hypothesized at the beginning of the chapter. Indeed:

1. The relationship between *arousal* and *participation* was stable. It was positive when the model was tested with the complete dataset ($r(76)= .405$, $p< .05$), when it was tested with the dataset of robot-based free play ($r(76)= .559$, $p< .05$), and when it was tested with the dataset of game-based cognitive stimulation ($r(76)= .531$, $p> .05$; non-significance presumably due to the low sampling adequacy).

2. The relationship between *arousal* and *valence* changed across activities. It was positive when the model was tested with the complete dataset ($r(76)= .388$, $p< .05$), when it was tested with the dataset of robot-based free play ($r(76)= .585$, $p< .01$), and when it was tested with the dataset of game-based cognitive stimulation ($r(76)= .307$, $p> .05$; non-significance presumably due to the low sampling adequacy). However, it varied in strength. It became always stronger (see covariance estimates in Figures 29, 30, and 31) with the increase of positive valence (positive valence: robot-based free play > complete dataset > game-based cognitive stimulation).

3. The relationship between *participation* and *valence* changed across activities. It was very low when the model was tested with the complete dataset ($r(76)= .099$, $p> .05$), slightly negative when the model was tested with the dataset of game-based cognitive stimulation ($r(76)= -.149$, $p> .05$), and positive when the model was tested with the dataset of robot-based free play ($r(76)= .550$, $p< .005$).

With regards to the relationship between *participation* and *valence*, in robot-based free play, participation was positive to such an extent that, by replacing focused attention with valence, the ENGAGE-DEM became a better fit for the data ($X^2(17, N= 39)= 19.791$, $p= .285$; $RMSEA= .063$; $NFI= .923$; $CFI= .987$; $RFI= .837$; $PNFI= .436$; see Figure 32). This is equivalent to saying that, when participation is positive, valence can take the place of focused attention as most of the focused attention is positively given.

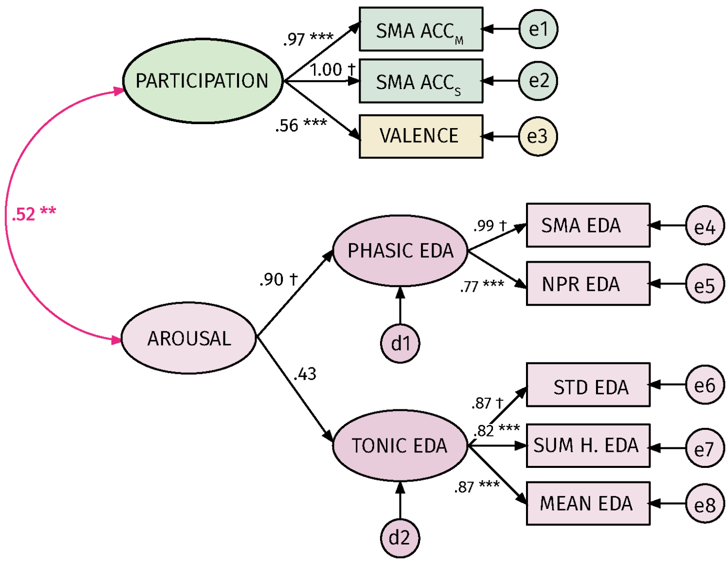


Figure 32. Modification of final model of engagement for robot-based free play. Significance level: † fixed factor, *< .05, **< .01, *< .001**

To conclude, the ENGAGE-DEM enabled us to satisfy two requirements for a proper measurement framework of engagement in dementia: (i) *identify the components of engagement and their boundaries* and (ii) *profile the expression of engagement by determining the relationships among its different components and measures*. Moreover, its testing allowed us to reply to the following research questions: *what are the relationships between the different components of engagement?*

10.6.1 Reflections on Engagement

The last research question that we need to respond to is: *which conclusions can be drawn from the relationships between the components of engagement regarding the functioning of engagement in dementia?* Taking inspiration from this question, we detail a number of considerations on engagement that the literature overlooks.

Engagement is not just positive. In spite of the fact that most of the literature defines engagement as a construct imbued with exclusively positive traits, the lack of fit of the theory-driven model and the goodness of fit of the ENGAGE-DEM substantiate the fact that engagement can be positive, but also neutral and negative. Positive engagement is a form of engagement, the most intense and, certainly, the most desirable in clinical terms. However, it is not engagement itself. In this sense, we would like to promote a different and more consistent definition of engagement for dementia and thus meet one of the first requirements of a proper measurement framework of engagement: *deploy a definition of engagement capturing its true connotation*. Engagement is the degree of proactive participation in an activity that can take different hedonic tones and achieve different levels of energy mobilization.

Disengagement is a form of engagement that does not have the activity as a focus. Despite the literature describes disengagement as lack of engagement, the body of work carried out for this dissertation restituted a different perspective. Indeed, across activities, we noticed that what we considered disengagement was in most cases a form of engagement directed toward a focus other than the activity: one's own body, the objects in the environment (e.g., paintings, a sewing machine, a bird outside the window), the persons in the environment (e.g., the experimenter). In light of this, the component *participation* of the ENGAGE-DEM assumes a very important role. Indeed, it serves to establish that the increase or decrease in *valence* and *arousal* is due to the engagement in the activity and not to the engagement with elements other than the activity (i.e., disengagement).

Different activities have different engagement potentials. Not all the activities for people with dementia are susceptible to reach the highest degree of engagement (*full participation-high arousal- positive valence*). For instance, in our experimental study, game-based cognitive stimulation could reach a neutral to somewhat positive engagement, whereas robot-based free play could achieve highly positive engagement. In research, the engagement potential needs to be taken into account when selecting activities as experimental conditions, as it defines which activities are really comparable in engagement terms.

Different people with dementia have different engagement potentials. Not all people with dementia are able to reach the highest degree of engagement. People with dementia differ in many respects, for instance in terms of mobility, dementia severity, sensorial impairment, and behavioral disturbances. These differences affect the level of engagement that a person with dementia can reach already at baseline. In this sense, it is extremely important to profile people with dementia according to their characteristics and establish their engagement baselines in order to truly understand whether their engagement increases over time or across experimental conditions.

10.7 Conclusions

In this chapter, we presented and successfully tested the final model of engagement for dementia, the ENGAGE-DEM. This model was the result of a progressive refinement of the theory-driven model of engagement and represents the main milestone of this dissertation. The ENGAGE-DEM constitutes a framework to measure engagement in dementia with objective measures pertaining to the behavioral/expressive and peripheral-physiological response systems. It details the components of engagement, how these are measured with the tools deployed throughout the thesis, and which relationships they entertain. The ENGAGE-DEM achieved an excellent goodness of fit both for the complete dataset and for the datasets of the single activities and, to the best of our knowledge, is the first formalization of the functioning of engagement in dementia.

The ENGAGE-DEM model could contribute to several domains of knowledge. It could benefit the field of nursing research since it could prompt a better understanding of the person with dementia and enable a more informed choice of meaningful activities. It could be an aid for designers aiming to test the user experience of playful technologies for people with dementia. Last, it could be used to enable socially interactive robots and interactive playful technologies to detect the engagement state of the person with dementia online and react accordingly.

11

Conclusions

11.1 Introduction

This chapter is aimed at summarizing the results and limitations of this dissertation. In the first part of the chapter, we detail the answers that we gave to the research questions posed in the Introduction. In the second part of the chapter, we reflect on the limitations of this research and on how prospective research should address them. Moreover, we envision one of the scientific scenarios that this research might generate and discuss how to face its ethical implications.

11.2 Research Questions

This thesis had a twofold objective. On the one hand, it was aimed at exploring and validating novel techniques to measure engagement in dementia at a behavioral/expressive and peripheral-physiological level. On the other hand, it was aimed at building a model of engagement for dementia detailing the main components of engagement, their measures, and their relationships. In this section, we summarize the answers to each of the research questions enlisted in the introduction, which have been discussed more extensively in the single chapters.

11.2.1 RQ1. How can engagement in dementia be measured?

RQ1a. *Can self-reports be employed as a gold standard of engagement in dementia? If not, which validated measures of engagement can be used as a gold standard?*

During the exploratory study, we employed the SIQ to collect the opinions of participants regarding the three board games that we proposed them: (i) jigsaw and shape puzzles, (iii) a match with the tiles of domino, and (iv) a categorization game. We posed the questions in the SIQ to participants both between activities and at the end of the sessions. What we noticed was that participants struggled in recalling the activities to which they had participated and in ranking their different experiences. A noteworthy fact in this sense was that when, at the end of the session, participants were asked which activity they liked the most among the three proposed, they always chose the latest one, which was always different due to randomization. On top of this, the responses of participants seemed affected by a *social desirability* and an *acquiescence bias*. Indeed, participants tended to be overly positive in their responses and to endorse any statement made by the facilitator and the experimenter, regardless of its content. As a result of this evidence, we did not consider self-reports reliable enough to be used as a gold standard for concurrent validity.

As a side remark, we would like to underline that the difficulty to self-report was observed in a sample of participants with mild and moderate dementia and therefore might not regard people diagnosed with very mild cognitive decline or MCI – the scores 2 and 3 of the Reisberg GDS. Also, these results do not mean to undermine the opinion of the person with dementia. We firmly believe that consulting the person with dementia regarding his/her activity preferences is of outmost importance. However, we retain that this should be done in a more qualitative and less quantitative way. As in the context of this research we needed a reliable and statistically valid quantitative measure of engagement as a gold standard, we decided to discard self-reports and resort to observational rating scales rated by experts – the OME and the OERS – as reference measures.

RQ1b. *Which peripheral-physiological measures can be employed to assess engagement in people with dementia? Are these peripheral-physiological measures appropriate to assess engagement?*

The review of the literature performed in chapter 2 unveiled the existence of several peripheral-physiological measures of engagement. However, some of them were prone to provide results difficult to interpret in engagement terms and difficult to judge in terms of reliability in people with dementia, for instance HR and HRV. Others entailed methods of data collection way too intrusive for people with dementia, this is the case of EEG and facial EMG. Still others were simply not suited for field data collection, namely fNIRS, urinalysis, and hormone analysis. We identified in EDA the most reliable and less intrusive peripheral-physiological measure of engagement. EDA measures the change in the skin conductance that derives from the activation of the SNS. The analysis of concurrent validity performed against the gold standard (i.e., OME and OERS) confirmed the appropriateness of EDA as a measure of engagement and brought us to include it in the list of measures to adopt in a model of engagement for dementia.

RQ1c. *Which behaviors externalize engagement in people with dementia? Are these behaviors appropriate to measure engagement?*

We retrieved several behavioral measures of engagement in the literature. However, we did not consider them methodologically adequate to measure engagement in its complexity and entirety. On the one side, the available ethograms produced measurements of engagement segmented into many small pieces of information difficult to trace back to overall engagement states. On the other side, coding schemes grasped only some of the characteristics of engagement-related behavior, as they were fine-tuned to reply to specific research questions. In order to address the issues highlighted in the literature, we developed a dedicated coding system to measure engagement-related behavior across activities, the ELICSE. The ELICSE was built following a mixed approach, ethographic and Laban-Inspired. Ethography served to deploy ethograms describing behavior in its more minute units, LMA to identify the underlying structure of behavior and create a coding system of engagement workable across activities. The ELICSE achieved an excellent IRR and a remarkable concurrent validity, hence the behaviors it featured were considered appropriate to measure engagement in dementia and were included in the model of engagement.

With respect to state of the art measurement techniques, the ELICSE has three crucial advantages: (1) it can be adapted to suit most of the activities provided with the characteristics described in chapter 3, (2) it can be modified to meet the characteristics of people with different levels of dementia (e.g., MCI, severe dementia), and (3) it can be used to measure behavior in its parts – as the overall duration of single behaviors – but also in its progression – as the succession of different body configurations.

With regards to this last point, in the process of construction of the model of engagement, we uncovered that the behaviors in the ELICSE were organized hierarchically and connected by regression paths depicting the way engagement-related behavior spread across body parts. The hierarchical organization of behavior was condensed in a model – the EMODEB – and can be used to infer different levels of engagement intensity, each operationalized with a different body configuration.

Within this dissertation, we developed a second measure of engagement that can be considered behavioral/expressive, quantity of movement. Quantity of movement is the amount of movement on the non-dominant wrist produced by people with dementia during activities and gauged with a triaxial accelerometer. As a measure, quantity of movement was not suggested by the literature on engagement. Instead, it was inferred from observed behavior and further endorsed by a number of articles regarding the use of actigraphy to diagnose apathy in dementia. Quantity of movement reached an optimal concurrent validity and thus was included in the final model of engagement.

As a side note, we would like to emphasize that also the observational rating scales OME and OERS fall within the behavioral/expressive measures of engagement. However, they differ from ethograms and coding schemes in that they do not directly measure behavior, but rather use behavior to operationalize engagement. For this reason, and also for their ordinal level of measurement, they were utilized as a ground truth, but not as measures of the components of engagement in the model.

11.2.2 RQ2. What are the dynamics of engagement in dementia?

RQ2a. *What are the relationships between the different components of engagement? Which conclusions can be drawn from these relationships regarding the functioning of engagement in dementia?*

The model of engagement developed in chapter 10 – the ENGAGE-DEM – constitutes the main milestone of this doctoral research. The ENGAGE-DEM is the result of a process of refinement of a model of engagement drawn from the literature – the theory-driven model of engagement – which proved to be misspecified. The ENGAGE-DEM features three components of engagement – *participation*, *arousal*, and *valence* – which are connected by correlation paths and assessed through the features of EDA, the features extracted from the accelerometer signal to gauge quantity of movement, and the aggregated scores of the behaviors in the ELICSE. According to the ENGAGE-DEM, the relationship between *arousal* and *participation* is stable and always positive, while the relationships between *arousal* and *valence* and between *participation* and *valence* change across activities and assume positive, negative, and non-significant values depending on the predominant affective state expressed by participants. The testing of the ENGAGE-DEM with SEM confirmed that the model was an excellent fit both for the complete dataset and for the datasets of the single activities. Moreover, it revealed that the relationships between the components of engagement varied as we had hypothesized.

As a result of the ENGAGE-DEM, we can state the following:

1. The engagement of the person with dementia is defined by his/her level of *participation* (ranging from *passive* to *full*), his/her level of *arousal* (ranging from *low* to *high*), and his/her level of *valence* (ranging from *negative* to *positive*).

2. *Arousal* and *participation* are always positively correlated, while *valence* might be correlated as well as uncorrelated with arousal and participation.

3. The state of engagement of the person with dementia can take on different *forms*, and range from *passive participation/low arousal/negative valence* to *full participation/high arousal/positive valence.*

11.3 Limitations and Future Work

The research outlined in this dissertation has been mainly limited by the sample size and by its geographical uniformity. Future work should attempt to include a higher number of participants coming from diverse cultural and geographic backgrounds. Given the amount of structuring that the collection of multimodal data in dementia entails, one of the possible solutions to enlarge the sample size and include a wider spectrum of nationalities in the samples is for researchers to work in transnational networks and collect data using common procedures. If future research manages to increase the sample size of the studies on engagement in dementia, it could also attempt to test the reliability of cardiac measures as peripheral-physiological correlates of engagement. This can be achieved by profiling participants according to their type of cardiac disease and cognitive deterioration.

Another limitation of this work resides in the narrow range of activities used to test the model of engagement. Forthcoming research should test the goodness of fit of the ENGAGE-DEM in activities other than game-based cognitive stimulation and robot-based free play, in activities involving robots other than Pleo, and in activities that envisage a larger number of participants. Additional work should also be devoted to understand whether the ENGAGE-DEM and the measures in it could be adapted to accommodate the characteristics of people in the earlier (e.g., MCI) and in the more advanced stages of dementia (i.e., moderately severe to severe dementia).

Besides these methodological issues, future research should focus on validating the scales of *focused attention* and *valence* presented in chapter 8 and use them to code the progression of engagement-related behavior over the activity with a time-sampling technique. This is of crucial importance as it could enable to gauge behavior as a time-series data and thus measure the components of engagement and their relationships in their temporal evolution.

As a last note, we would like to bring the attention of the reader to the fact that the ENGAGE-DEM might constitute a reference framework for the automatic recognition of engagement in people with dementia. Indeed, the model features peripheral-physiological measures that are already gauged with unobtrusive sensors and behavioral/expressive measures that *could be* gauged with unobtrusive sensors (e.g., gaze, postures). The usage of sensing technologies to detect affective states in dementia is a completely new scientific field without an ethical regulation and without specific privacy boundaries. Prospective research should interrogate itself on the ethical implications of the application of affective computing to dementia, and draft a legal frame encompassing issues such as data protection, data portability, third uses of data, and legal liability (Fosch Villaronga, 2016). This way, the automatic recognition of engagement would be targeted to the enhancement of QoL without being detrimental to the rights of the person with dementia.

Appendices

Appendix A. The Ethnographic Study

Appendix B. The Exploratory Study

Appendix C. Script for Facilitators

Appendix D. OME and OERS

Appendix E. The Coding System (ELICSE)

Appendix F. Further Analyses

Appendix G. SEM Glossary

Appendix A. The Ethnographic Study

This appendix describes the objectives, methods, and results of the first study carried out in the context of this doctoral research, the ethnographic study.

Objectives. The ethnographic study was meant to reply to the questions: *how do activities for people with dementia work in a real-life environment? How can they be constrained for data collection?* In this respect, the study had four objectives: (1) determine the exact *profile* of participants to be included in further studies (i.e., dementia severity, behavioral disturbances), (2) establish the participants' *group size* and *composition* for the activities to be performed in subsequent studies, (3) identify the *specific activities* on which the model of engagement should focus, and (4) unveil how these activities are usually *structured* in the nursing home.

Participants. The participants that we observed over the course of the study were most of the residents of the nursing homes Redós de Sant Josep i Sant Pere (113 residents) and Casa d'Empara (78 residents) situated in the province of Barcelona. During observations, we specifically focused on nine of these residents. These had an age comprised between 79 and 92 years (M_{age}= 85), a diagnosis of MCI, mild, and moderate dementia (scores of 3, 4, and 5 at the Reisberg GDS), and different types of behavioral and psychological disturbances (i.e., apathy, depression, both of them, or none of them).

Design. The study was an ethnographic study conducted with a participant stance. For one month, we took part in the activities of the nursing homes together with the residents and interacted with them in their real-life environment. The activities to which we participated were the following: cognitive games, gymnastic, painting workshop, basketball, bocce, cooking workshop, memory workshop, reading/writing workshop, and music activity (i.e., singing). These activities were repeated weekly, thus we observed them more than once.

Measures. The study had a strongly qualitative character. We mainly collected notes and semi-structured interviews with stakeholders (i.e., residents, nurses, geriatrician, social educator, psychologist, and occupational therapist).

Figure A1. Activities of residents. On the left, the memory workshop; on the right, the gymnastic session (from elredos.cat)

Results. The observed activities could be grouped in three types: (1) those encompassing *physical effort*, for instance gymnastic, basketball, and bocce, (2) those encompassing the use of *tangible artifacts*, for instance cognitive games, the painting workshop, the cooking workshop, and the reading/writing workshop, and, last, (3) those encompassing the use of *abstract artifacts* (i.e., thought and music), for instance the memory workshop and the music activity. These activities were mostly carried out in large groups. Sometimes, they gathered 15-30 residents (see Figure A1) that participated to the same activity altogether (e.g. memory workshop, gymnastic). Other times, they gathered 50 residents clustered in small groups of four to five persons that completed the activities on their own (e.g., cooking workshop, cognitive games). When activities entailed group participation, they consisted of three phases: a warm-up, the activity itself, and a conclusion. For instance, in the memory workshop, the activity was structured as follows: (i) *warm-up phase*: temporal and spatial orientation, (ii) *activity*: the specific cognitive and mnemonic exercises, and (iii) *conclusion*: listening and dancing to music.

The composition of the groups was quite heterogeneous. Participants in the activities had varying degrees of deterioration – usually spanning from MCI to severe dementia – and were affected by different behavioral disturbances. In this respect, during group activities, we noticed that: (i) participants with less cognitive deterioration tended to take the lead, and outclass other participants, (ii) participants with MCI had a much higher ability to respond to activities and were considerably faster in reacting to stimulation with respect to participants with mild and moderate dementia, (iii) participants with mild and moderate dementia did not show steep differences in their responses, and (iv) participants with mild and moderate dementia had a much higher ability to respond to activities and were considerably faster in reacting to stimulation with respect to participants with severe dementia. Indeed, in some cases, residents with severe dementia fell asleep during activities or entered catatonic states. In addition to this, we observed that: (iv) regardless of dementia severity, participants with motivational disorders (i.e., apathy and depression) were more discontinuous in their responses to activities and had more blunted emotional reactions.

Discussion. As a result of the ethnographic study, we decided to focus on activities encompassing the use of tangible artifacts and entailing social interaction. Indeed, this category of activities could be extended to include technology-based activities. Moreover, we decided to adopt the observed structuring in three phases in the subsequent studies. With regards to group size, we resolved to reduce the size of groups to the minimum (i.e., a couple) to be still able to observe social dynamics, but at the same time have a control on lurking variables. Also, we chose to make group composition more homogenous by working with a sample of participants representative of dementia – people with mild and moderate dementia – but not at the early nor at the last stages of the disease. Last, we decided not to exclude participants with motivational disorders from the study, but to control the effect of apathy and depression on engagement.

Appendix B. The Exploratory Study

this appendix summarizes the objectives, methods, and results of the exploratory study that was carried out before conducting the experimental study described in chapter 4.

Objectives. The exploratory study was meant to reply to the question: *how can we bring the spontaneity of a real-life activity into an experimental set-up while at the same time measuring engagement in a sound manner?* In this regard, the study had three objectives: (1) establish the *setting* of data collection, (2) determine the *procedure* of data collection, and (3) investigate the reliability of *self-reports* as a gold standard for the assessment of engagement. In order to pursue these objectives, we followed the guidelines resulting from the ethnographic study: (i) we constrained the setting and the group size of activities, (ii) we included in the study only participants with mild and moderate dementia, (iii) we employed one of the activities encompassing the use of tangible artifacts, and (iv) we structured it in three phases.

Participants. The participants taking part in the exploratory study were eight residents of the nursing home Redós de Sant Josep i Sant Pere situated in the province of Barcelona. They were aged between 78 and 92 years (M_{age}= 81) and had a diagnosis of dementia ranging from mild to moderate (scores of 4 and 5 at the Reisberg GDS; scores between 10 and 23 at the MEC).

Design. The exploratory study followed a repeated measures design. The selected participants were randomly coupled and took part in three activity sessions together. During these sessions, they were asked to play three board games: (1) puzzles (i.e., jigsaw and shape puzzles), (2) a match with the tiles of domino, and (3) a categorization game. The order of presentation of the three board games was randomized across sessions with a Latin squares technique and each activity was conducted by a facilitator (i.e., psychologist or social educator of the care facility) at the presence of an experimenter.

Measures. The study had an eminently methodological objective, thus it did not entail the use of measures other than self-reports. Also in the case of self-reports, the goal was not to assess which board game elicited more engagement in the person with dementia, but to judge whether self-reports could reliably grasp the experience of engagement of the person with dementia. To do so, we adopted the *Subjective Impression Questionnaire* (SIQ;

Wada et al., 2005) and adapted its questions to the context of board games. The questions were the following: (i) *Was the game interesting/not interesting?* (ii) *Did you like/dislike the game?* (iii) *Was playing the game fun/boring?* (iv) *Would you want to join the activity again?* The questions were asked by the facilitator during breaks in the activity. At the end of each session, participants were also asked which of the three games they preferred, and would be more interested in replicating. After sessions, the responses were transcribed to assign them a score on a five-point Likert scale.

The sessions of the exploratory study were video recorded.[10] The videos were thoroughly observed and used to evaluate the appropriateness of the scenario of data collection. Also, they were used to familiarize more closely with the behavioral/expressive responses of the participants to the activities.

Setting. The exploratory study took place in a small room usually destined to activities. During sessions, participants sat on the same side of the table, and two hand-held video cameras were positioned one in front of them, the other on their side. The experimenter sat in front of participants, while the facilitator was left free to displace, seat, or stand.

Procedure. Each session had three phases: (1) a *warm-up phase*, where the facilitator asked few riddles to the participants to get them accustomed to the activity set-up, (2) an *activity phase*, where participants played the three board games in randomized order, and (3) a *wrap-up phase*, where participants were asked to express which game they had preferred. During the activity phase, there was a break after each game was completed. During these breaks, participants were asked the questions of the SIQ relative to the just completed game. In view of the collection of physiological data, these breaks could serve to bring EDA back to baseline before the start of a new game.

Results. The SIQ was difficult to collect in its entirety from both participants. Also, in most cases, responses were affected by a *social desirability* and *acquiescence* bias. Participants were overly positive in their reports and agreed with most of the statements of the facilitator and experimenter. Finally, we noticed that when, at the end of the session, participants were asked which game they liked the most and would have liked to replicate, they nearly always chose the latest one, which was always different due to randomization. We attributed this phenomenon to memory impairment. As a consequence of this scenario, we could not attribute scores to participants' responses on a five-point Likert scale as suggested by Wada et al. (2005).

10 In order to collect videos, we obtained an informed consent both from participants and from their legal guardians.

With regards to the setting of the study, it proved to work well. The cozy room created a relaxed environment. Video cameras were well-positioned and enabled to collect quality data for behavioral analysis. The only inconvenient regarded the positioning of facilitators. They were sometimes off-camera, other times positioned in-between participants, still other times on the side of one of the participants. This latter positioning seemed to be perceived by the farthest participant as a preference statement in favor of the closest participant. In terms of group size and composition, we noticed that, albeit participants were arranged in couples, we could still observe a lot of social dynamics. Also, in spite of the different cognitive load of the activity on participants with mild and with moderate dementia, we could not observe consistent changes in engagement between them.

With respect to the procedure of data collection, the subdivision of the sessions in the exploratory study was poor. The warm-up phase was very hectic, thus not extendible to an experimental study aimed at collecting physiological data. Also, we realized that the breaks between games could not be used to bring arousal back to baseline in further studies. Indeed, during them, participants made lengthy autobiographical digressions susceptible to trigger peripheral activation.

Discussion. As a result of the exploratory study, in the experimental study: (1) *self-reports* were discarded as means to collect a gold standard of engagement, as they were not considered reliable, (2) the *setting* of data collection was confirmed, but facilitators were asked to stand between participants during activities, (3) the *group size and composition* were considered adequate to promote social interaction, while at the same time controlling variables, and (4) the *warm-up phase* was turned into an habituation phase, while the breaks between games were removed.

As a follow-up of the exploratory study, we performed some trials to figure out how to collect the *baseline* of EDA. We attempted to collect the EDA baseline while participants were inactive, while they were conversing, and while they were listening to stories. This last task appeared to be the one eliciting the most neutral activation and was thus employed in the experimental study.

Appendix C. Script for Facilitators

in this appendix, we present the scripts that facilitators followed during the sessions of the experimental study. First, we provide the script for game-based cognitive stimulation, then, the one for robot-based free play.

C1. Script for Game-based Cognitive Stimulation

1. **Rest/conversation (5 minutes)**: once in the activity room, the participants seat on the same side of the table. Then, the facilitator and the experimenter start a conversation with them.

2. **Introduction of activity (2 minutes)**: after this brief conversation, the facilitator explains the activities to the participants, shows them the video cameras and reminds them about the study to which they are participating[11].

3. **Wristband positioning (3 minutes)**: once the activity is introduced and participants agree to participate, the experimenter helps them to wear the wristbands (see Figure C1) and verifies that both wristbands properly function.

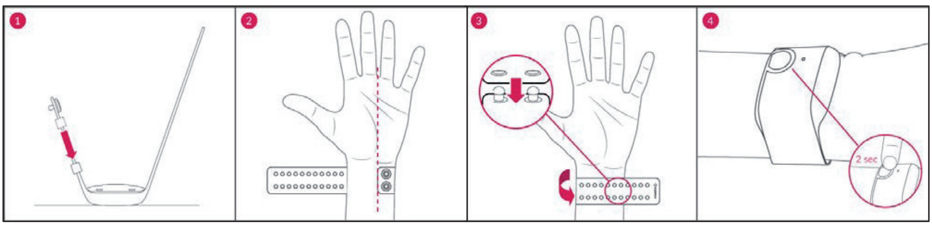


Figure C1. Positioning of E4 wristband

4. **Switch on wristband (1 minute)**: as soon as the wristbands are properly positioned, the experimenter switches them on (see Figure C1).

11 Note that participants were asked the informed consent well before the start of the data collection. At each session, we just wanted to be sure that they recalled the reason why they were participating in the study and agreed to proceed.

5. **Synchronization (1 minute)**: then, s/he synchronizes them with the video-footage by simultaneously pressing the tag buttons on the wristbands of both participants in front of the video cameras for one second.

6. **Baseline collection (5 minutes)**: to facilitate the baseline collection, the facilitator reads a fairytale to the participants. The reading lasts five minutes. The experimenter notifies the facilitator when the five minutes have elapsed.

7. **Jigsaw Puzzle**[12]: After the baseline collection, the facilitator places the pieces of the first jigsaw puzzle on the table in the middle of the two participants. The participants complete three jigsaw puzzles presented in a progressive order of difficulty. The jigsaw puzzles are arranged in the right order by the experimenter before the session.

8. **Shape Puzzle**: once the jigsaw puzzles are completed, the facilitator proceeds to present the participants with the shape puzzles. S/he places the pieces of the shape puzzle on the table in front of the participants and the wooden board in which the pieces should be wedged in in the middle of participants. The participants complete three shape puzzles presented in a progressive order of difficulty. The shape puzzles are organized in the appropriate order by the experimenter before the session.

9. **Dominoes**: after the shape puzzles are completed, the facilitator presents the new game, the domino. S/he explains the rules of the domino match to the participants, places the tiles of the domino in a stack, and asks the participants to select seven of them. As soon as the participants choose their tiles, the match starts. Participants play just one match of domino.

   - *It is very importa*nt that the facilitator:
   - Promotes a relaxed environment so that participants do not feel under test
   - Checks that participants have equal access to the activity
   - Repeats the rules of the game if participants are confused

12 Please note that the order of activity presented in this appendix is the one of the first session of game-based cognitive stimulation. Thus, it is not valid for all three sessions.

- Helps participants in case of difficulty

- Places him/herself in the middle *of the participants during the activity*

10. **Removal of the Wristband (5 minutes)**: Once the board games are completed, the facilitator ends the activity. As the activity is concluded, the experimenter switches off the wristbands and removes them from the wrists of participants. Then, s/he switches off the video cameras.

11. **Observational rating scales**: at this point, the facilitator helps the participants to go back to their units and, when back, fills out the observational rating scales (OME and OERS).

C.2 Script for Robot-based Free Play

1. **Rest/conversation (5 minutes)**: once in the activity room, the participants seat on the same side of the table. Then, the facilitator and the experimenter start a conversation with them.

2. **Introduction of activity (2 minutes)**: after this brief conversation, the facilitator explains the activities to the participants, shows them the video cameras and reminds them about the study to which they are participating[13].

3. **Wristband positioning (3 minutes)**: once the activity is introduced and participants agree to participate, the experimenter helps them to wear the wristbands (see Figure C2) and verifies that the wristbands properly function.

4. **Switch on wristband (1 minute)**: as soon as the wristbands are properly positioned, the experimenter switches them on (see Figure C2).

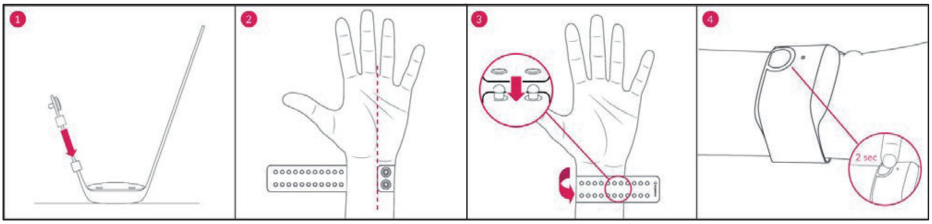


Figure C2. Positioning of E4 wristband

5. **Synchronization (1 minute)**: then, s/he synchronizes them with the video-footage by simultaneously pressing the tag buttons on the wristbands of both participants in front of the video cameras.

6. **Baseline collection (5 minutes)**: to facilitate the baseline collection, the facilitator reads a fairytale to the participants. The reading lasts five minutes. The experimenter notifies the facilitator when the five minutes have elapsed.

7. **Play with Pleo**: the facilitator switches on Pleo and places it in the middle of the table. S/he observes the reactions of participants, what they say and what they do. S/he could prompt a series of interactions with the robot:

13 Note that participants were asked the informed consent well before the start of the data collection. At each session, we just wanted to be sure that they recalled the reason why they were participating in the study and agreed to proceed.

- *Call Pleo:* the facilitator asks participants to call Pleo by name. The participants try to convince Pleo to approach them by calling its name. Pleo perceives sounds and orientates towards them, thus it will move its head in the direction of the participant that calls it.

- *Make Pleo sleep*: the facilitator asks participants to make Pleo sleep as they would do with a baby, for instance by cradling it. Pleo perceives when it is picked up and starts sleeping in few seconds.

- *Feed Pleo*: the facilitator asks participants to feed Pleo because it is hungry. Participants are provided with the leaves to feed the robot. The facilitator suggests them to place the leaf in front of the snout of the robot so that it can smell it. Once Pleo perceives the leaf, it opens its mouth.

- *Heal Pleo*: when Pleo cries, the facilitator asks participants if they think he might be sick, then s/he provides them with the stone that can heal it. When Pleo perceives the stone, it moves its head away as it does not like it.

The activity with Pleo lasts twenty minutes. The experimenter notifies the facilitator once the twenty minutes elapse.

It is very important that the facilitator:

- *Promotes a relaxed environment*

- *Checks that participants have equal access to the robot*

- *Places him/herself in the middle of the participants during the activity*

8. **Removal of the Wristband (5 minutes)**: Once 20 minutes have passed, the experimenter notifies the facilitator to end the activity. Once the activity is concluded, s/he switches off the wristbands and removes them from the wrists of participants. Last, s/he switches off the video cameras.

9. **Observational rating scales**: at this point, the facilitator helps the participants to go back to their units and, when back, fills out the observational rating scales (OME and OERS).

Appendix D. OME and OERS

The Observational Measurement of Engagement – OME

Attention

4 Very Attentive
3 Attentive
2 Somewhat Attentive
1 Not Attentive

Attitude toward game

7 Very Positive
6 Positive
5 Somewhat Positive
4 Neutral
3 Somewhat Negative
2 Negative
1 Very Negative

Cognitive Difficulty

5 Very Difficult
4 Difficult
3 Moderately Difficult
2 Slightly Difficult
1 Not at all Difficult
0 Could not Determine

Attitude toward partner

7 Very Positive
6 Positive
5 Somewhat Positive
4 Neutral
3 Somewhat Negative
2 Negative
1 Very Negative

Attention. Amount of attention that the resident is paying to the object during the engagement (manipulating/holding/content of talking about object is all attention) toward the stimulus. Following staff instructions without any change in affect is still attention. Attention can be physical (i.e., stroking Pleo even if looking away) or visual (i.e., staring at Pleo while it moves even if not interacting with it).

Attitude. Amount of excitement/expressiveness toward the stimulus (smiling, frowning, energy, excitement in voice). If the resident is involved (manipulating stimulus), but has no visible affect then still mark "somewhat positive"; if resident is not interested at all or looking at the object, but never actively participates (not holding or manipulating), then mark "neutral".

Observed Emotion Rating Scale – OERS

Please, rate the extent of each affective state during the activity. Some possible signs of each emotion are listed. If you see no sign of a particular feeling, rate "Never"

		7	1	2	3	4	5
		Not in view	Never	Less than 16 sec.	16-59 sec.	1-5 min.	More than 5 min.
PLEASURE Signs: Laughing, Signing, Smiling, Kissing, Stroking or Gently touching other, Reaching out warmly to other.							
ANGER Signs: Physical Aggression, Yelling, Cursing, Berating, Shaking fist, Drawing eyebrows together, Clenching teeth, Pursing lips, Narrowing eyes, Making distancing gestures.							
ANXIETY/FEAR Signs: Shrieking, Repetitive calling out, Restlessness, Wincing/Grimacing, Repeated or Agitated movements, Line between eyebrows, Lines across forehead, Hand wringing, Tremor, Leg jiggling, Rapid breathing, Eyes wide, Tight facial muscles.							
SADNESS Signs: Crying, Frowning, Eyes drooping, Moaning, Sighing, Head in hands, Eyes/ head turned down and face expressionless (only counts as sadness if paired with another sign).							
GENERAL ALERTNESS Signs: Participating in a task, Maintaining eye-contact, Eyes following object or person, Looking around room, Responding by moving or saying something, turning body or moving towards object or person.							

Appendix E. The Coding System (ELICSE)

Table E1. Operational descriptions of head behaviors

HEAD BEHAVIORS	MODIFIERS – gestural support
Gaze toward partner The observed participant* directs the head and the eyes or just the eyes toward the partner* **CODE: GP**	***positive* gestural support (code: GP_pos)** *with* facial expressions or gestures that display positive affect: smile, laugh, stick the tongue out, send/give kisses ***no* gestural support (code: GP_no)** without facial expressions or gestures that display positive affect ***negative* gestural support (code: GP_neg)** *with* facial expressions or gestures that display negative affect: anger, sadness, disgust, fear, frowning, boredom (yawn), pain
Gaze toward facilitator/experimenter The observed participant directs the head and the eyes or just the eyes toward the facilitator or the experimenter **CODE: GFE**	***positive* gestural support (code: GFE_pos)** *with* facial expressions or gestures that display positive affect: smile, laugh, stick the tongue out, send/give kisses ***no* gestural support (code: GFE_no)** without facial expressions or gestures that display positive affect ***negative* gestural support (code: GFE_neg)** *with* facial expressions or gestures that display negative affect: anger, sadness, disgust, fear, frowning, boredom (yawn), pain
Gaze toward game The observed participant directs the head and the eyes or just the eyes toward the game* **CODE: GG**	***positive* gestural support (code: GG_pos)** *with* facial expressions or gestures that display positive affect: smile, laugh, stick the tongue out, send/give kisses, nuzzle the robot, blow of the surface of the robot ***no* gestural support (code: GG_no)** without facial expressions or gestures that display positive affect ***negative* gestural support (code: GG_neg)** *with* facial expressions or gestures that display negative affect: anger, sadness, disgust, fear, frowning, boredom (yawn), pain
None of the target head movements The observed participants does not perform any of the target head behaviors (see above) **CODE: NoH**	***positive* gestural support (code: NoH_pos)** *with* facial expressions or gestures that display positive affect: smile, laugh ***no* gestural support (code: NoH_no)** without facial expressions or gestures that display positive affect ***negative* gestural support (code: NoH_neg)** *with* facial expressions or gestures that display negative affect: anger, sadness, disgust, fear, frowning, boredom (yawn), pain, avoidance, closed eyes

*1 Observed participant: the observed participant is the one whose behavior we are scoring

*2 Partner: the partner is the participant who is taking part in the activity with the observed participant, but whose behavior we are not currently scoring

*3 Game: the game is the physical set of tools that the participants manipulate during the activity. For instance, in dominoes, the tiles of the domino, in puzzles, the pieces of the puzzle, in the robot interaction, the robot itself, but also its accessories (e.g., leaves, stones)

Table E2. Operational descriptions torso behaviors

TORSO BEHAVIORS	MODIFIERS – postural support
Lean in partner The observed participant leans in the partner by rotating the torso and advancing it in the direction of the partner or by spreading it sideways in the direction of the partner **CODE: LIP**	***positive* postural support (code: LIP_pos)** *with* movements of the torso that display positive affect: hug the partner ***no* postural support (code: LIP_no)** without movements of the torso that display positive affect ***negative* postural support (code: LIP_neg)** *with* movements of the torso that display negative affect: pull away the partner when s/he is approaching
Near reach/lean toward the game The observed participant tilts the torso toward the game or holds the game close to its torso **CODE: NRLTG**	*positive* postural support (code: NRLTG_pos) *with* movements of the torso that display positive affect: hug the robot, cradle the robot, make the robot sleep on the shoulder ***no* postural support (code: NRLTG_no)** without movements of the torso that display positive affect ***negative* postural support (code: NRLTG_ neg)** *with* movements of the torso that display negative affect: reject the game by pulling it away
None of the target torso movements The observed participants does not perform any of the target torso behaviors (see above) **CODE: NoT**	**(none)**

Table E3. Operational descriptions of arms/hands behaviors

ARMS/HANDS BEHAVIORS	MODIFIERS – quality of gesture
Reach out partner The observed participant touches or indicates the partner, passes the game to or receives the game from the partner, manipulates the game while this is held by the partner **CODE: RoP**	***positive*** **quality of gesture (code: RoP_pos)** *The* observed participant strokes, pats or hugs the partner. The observed participant waves or strokes the robot, puts the finger in the mouth or takes the chin of the robot, strokes the robot with one of its accessories (i.e., leaves, stones) when the robot is held by the partner ***no*** **quality of gesture (RoP_no)** (See reach out partner) ***negative*** **quality of gesture (code: RoP_Pos)** *The* observed participant rejects or hits the partner. The observed participant rejects the game, hits or puts a finger in the eyes of the robot when the robot is held by the partner
Reach out facilitator/experimenter The observed participant touches or indicates the facilitator/experimenter, passes the game to or receives the game from the facilitator, manipulates the game while this is held by the facilitator **CODE: RoFE**	***positive*** **quality of gesture (code: RoFE_pos)** *The* observed participant strokes, pats or hugs the facilitator. The observed participant waves or strokes the robot, puts the finger in the mouth or takes the chin of the robot, strokes the robot with one of its accessories (i.e., leaves, stones) when the robot is held by the facilitator ***no*** **quality of gesture (RoFE_no)** (See reach out facilitator/experimenter) ***negative*** **quality of gesture (code: RoFE_neg)** *The* observed participant rejects or hits the facilitator. The observed participant rejects the game, hits or puts a finger in the eyes of the robot when the robot is held by the facilitator
Manipulate game The observed participant holds, touches, interacts (e.g., puts the pieces of the puzzle in place, feeds the robot), or indicates the game **CODE: MG**	***positive*** **quality of gesture (code: MG_pos)** *The* observed participant waves, hugs, cradles or strokes the robot, puts the finger in the mouth or takes the chin of the robot, strokes the robot with one of its accessories (i.e., leaves, stones). The observed participant touches strongly the pieces of the shape or jigsaw puzzle with his/her fist or hand after combining them or flips the tile of the domino on its center ***no*** **quality of gesture (code: MG_no)** (See manipulate the game) ***negative*** **quality of gesture (code: MG_neg)** *The* observed participant rejects, hits or throws away the game or makes the robot fall, turns it upside down, and puts a finger its eyes
Positive signs of affection involving arms/hands The observed participant claps his/her hands, dances, mimics a torero, rhythmically bangs his/her fists on the table **CODE: SOA_pos**	**(none)**

ARMS/HANDS BEHAVIORS	MODIFIERS – quality of gesture
Negative signs of affection involving arms/hands *The* observed participant hides its face in the hands, covers its face with the hands, or performs insulting gestures **CODE: SOA_neg**	**(none)**
None of the target arms/hands movements *The* observed participants does not perform any of the target arms/hands behaviors (see above) **CODE: NoAH**	**(none)**

Appendix F. Further Analyses

In this appendix, we detail the results of the analyses performed to unclose differences in engagement occurring between participants with and without motivational disorders (i.e., apathy and depression). As the discussion of the results is presented at the beginning of chapter 7, here we limit ourselves to enumerate the significant results.

Significant Results

We performed a number of mixed factorial ANOVAs in SPSS 22.0 (listwise exclusion of cases) to disclose whether there were significant differences between participants with and without motivational disorders along the different measures of engagement – the OME and OERS, the ELICSE, quantity of movement, and EDA. Across analyses, we employed *activity type* (game-based cognitive stimulation vs robot-based free play) as within-subject factor, and the presence of *motivational disorders* (presence vs absence) as between-subject factor. In the lists, we report all the significant results, in the figures, just the results including a main effect of motivational disorders on engagement. Also, we do not report main effects of activity type on engagement, as game-based cognitive stimulation and robot-based free play are not considered comparable in engagement terms. For the OME and OERS (N_{motdis}= 18; $N_{w/omotdis}$= 24), results revealed:

1. An interaction effect of motivational disorders and activity type on the item *attention* of the OME ($F(1,40)= 11.672$, $p= .001$, $\eta^2= .226$).

2. A main effect of motivational disorders on the item *attention* of the OME ($F(1,40)= 28.442$, $p< .001$, $\eta^2= .416$).

3. An interaction effect of motivational disorders and activity type on the item *attitude toward game* of the OME ($F(1,40)= 4.603$, $p< .05$, $\eta^2= .103$).

4. A main effect of motivational disorders on the item *attitude toward game* of the OME ($F(1,40)= 26.963$, $p< .001$, $\eta^2= .403$).

5. A main effect of motivational disorders on the item *attitude toward partner* of the OME ($F(1,40)= 13.261$, $p= .001$, $\eta^2= .249$).

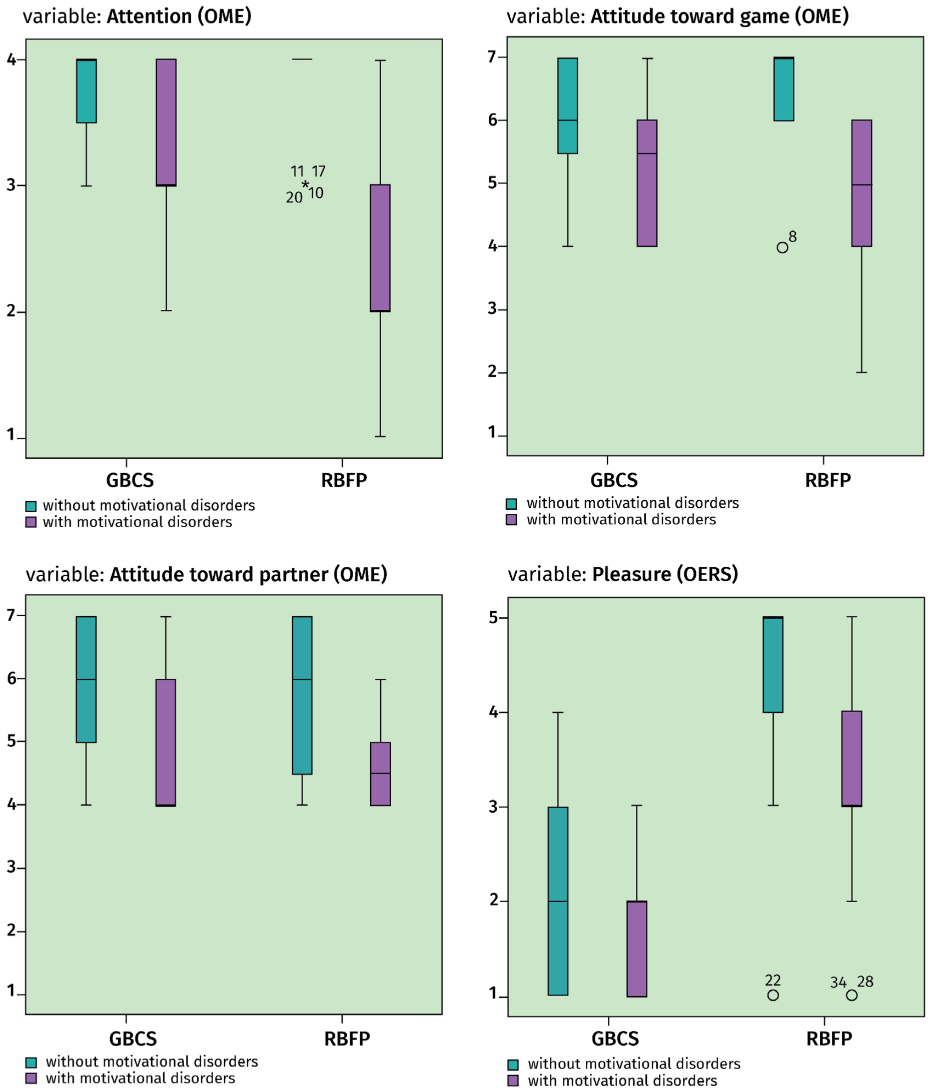


Figure F1. Effects of motivational disorders on items of OME and OERS (significance< .05). GBCS: game-based cognitive stimulation; RBFP: robot-based free play

6. A main effect of motivational disorders on the item *pleasure* of the OERS ($F(1,40)= 14.548$, $p < .001$, $\eta^2 = .267$).

Across items, engagement was always lower in participants with motivational disorders with respect to participants without such disorders (see Figure F1). Also, participants with apathy and depression preferred game-based cognitive stimulation to robot-based free play and did not differ significantly on pleasure across activities.

For the ELICSE (N_{motdis}= 18; $N_{w/omotdis}$= 24), results revealed:

1. An interaction effect of motivational disorders and activity type on *near reach/lean toward game* ($F(1,40)= 5.973$, $p< .05$, $\eta^2= .130$).

2. An interaction effect of motivational disorders and activity type on *none of the target torso behaviors* ($F(1,40)= 5.765$, $p< .05$, $\eta^2= .126$).

3. An interaction effect of motivational disorders and activity type on *manipulate game* ($F(1,40)= 7.502$, $p< .01$, $\eta^2= .158$).

4. A main effect of motivational disorders on *manipulate game* ($F(1,40)= 14.970$, $p< .001$, $\eta^2= .272$).

5. An interaction effect of motivational disorders and activity type on *none of the target arms/hands behaviors* ($F(1,40)= 7.678$, $p< .01$, $\eta^2= .161$).

6. A main effect of motivational disorders on *none of the target arms/hands behaviors* ($F(1,40)= 15.693$, $p< .001$, $\eta^2= .282$).

7. An interaction effect of motivational disorders and activity type on *positive gestural support* ($F(1,40)= 5.236$, $p< .05$, $\eta^2= .116$).

8. A main effect of motivational disorders on *positive gestural support* ($F(1,40)= 5.116$, $p< .05$, $\eta^2= .113$).

9. A main effect of motivational disorders on *negative gestural support* ($F(1,40)= 16.503$, $p< .001$, $\eta^2= .292$).

10. An interaction effect of motivational disorders and activity type on *positive quality of gesture* ($F(1,40)= 4.279$, $p< .05$, $\eta^2= .097$).

11. A main effect of motivational disorders on *positive quality of gesture* ($F(1,40)= 4.488$, $p< .05$, $\eta^2= .101$).

12. An interaction effect of motivational disorders and activity type on *negative quality of gesture* ($F(1,40)= 4.741$, $p< .05$, $\eta^2= .106$).

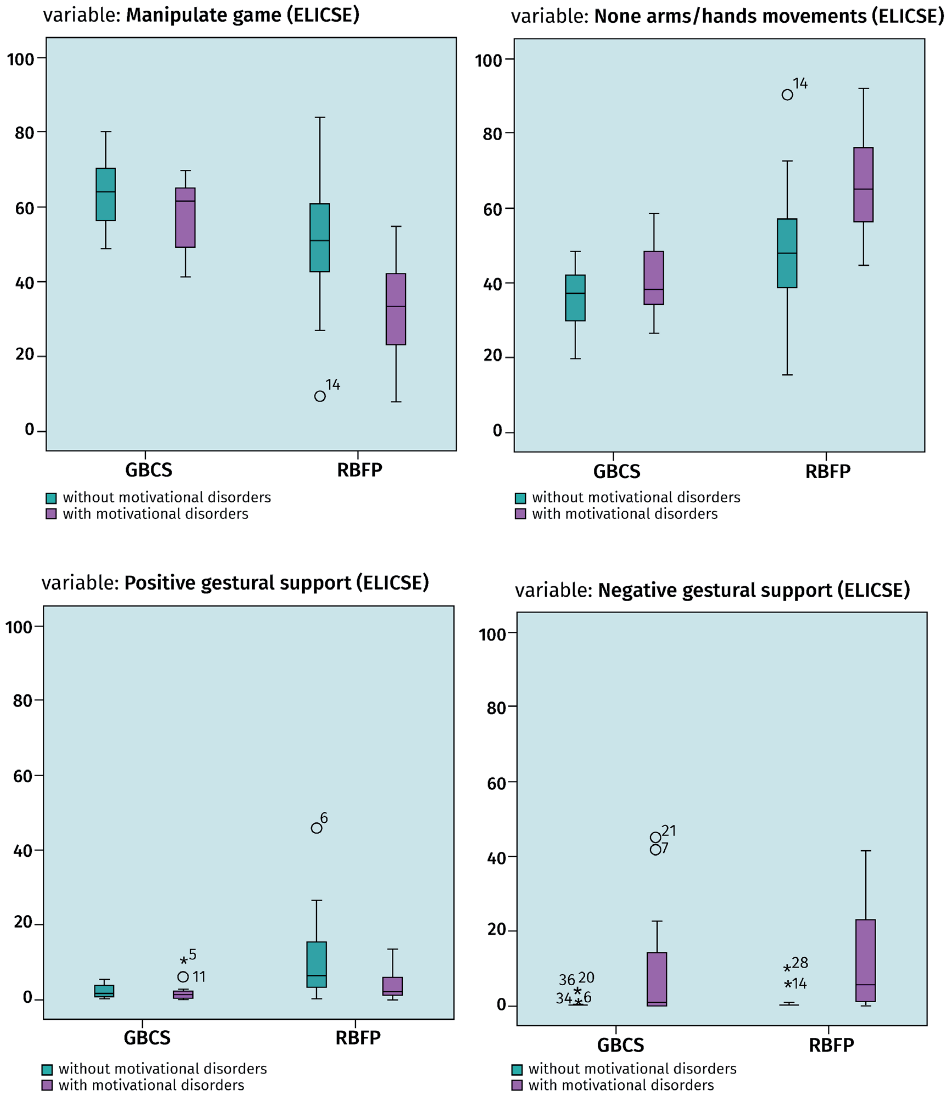


Figure F2. Effects of motivational disorders on behaviors of the ELICSE (significance< .05). GBCS: game-based cognitive stimulation; RBFP: robot-based free play

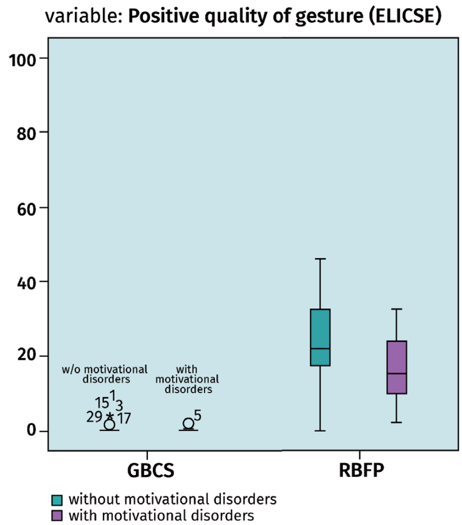


Figure F2 (continuation). Effects of motivational disorders on behaviors of the ELICSE (significance< .05). GBCS: game-based cognitive stimulation; RBFP: robot-based free play

Behaviors of engagement (*gaze toward game*, *gaze toward partner*, *near reach/lean toward game*, *lean in partner*, *manipulate game*, and *reach out partner*) and positive valence (*positive gestural support*, *positive postural support*, and *positive quality of gesture*) were lower in participants with apathy and depression with respect to those not affected with these disorders, while behaviors of disengagement (*gaze toward facilitator/experimenter*, *none of the target head movements*, *none of the target torso movements*, *reach out facilitator/experimenter*, *none of the target arms/hands movements*) and negative valence (*negative gestural support*, *negative postural support*, *negative quality of gesture*) were higher in the former group with respect to the latter (see Figure F2). When an interaction effect was involved, in most cases it was robot-based free play to elicit worse results in participants with depression and apathy (the only exceptions to this paradigm are *positive gestural support* and *positive quality of gesture*). The results regarding *postural support*, both positive and negative, were excluded from the analyses as *postural support* occurred just once during game-based cognitive stimulation.

For quantity of movement (N_{motdis}= 17; $N_{w/omotdis}$= 17), results revealed:

1. A main effect of motivational disorders on the feature SMA Acc_M (F(1,32)= 12.373, p= .001, η^2= .279).

2. A main effect of motivational disorders on the feature SMA Acc_S (F(1,32)=14.359, p= .001, η^2= .310).

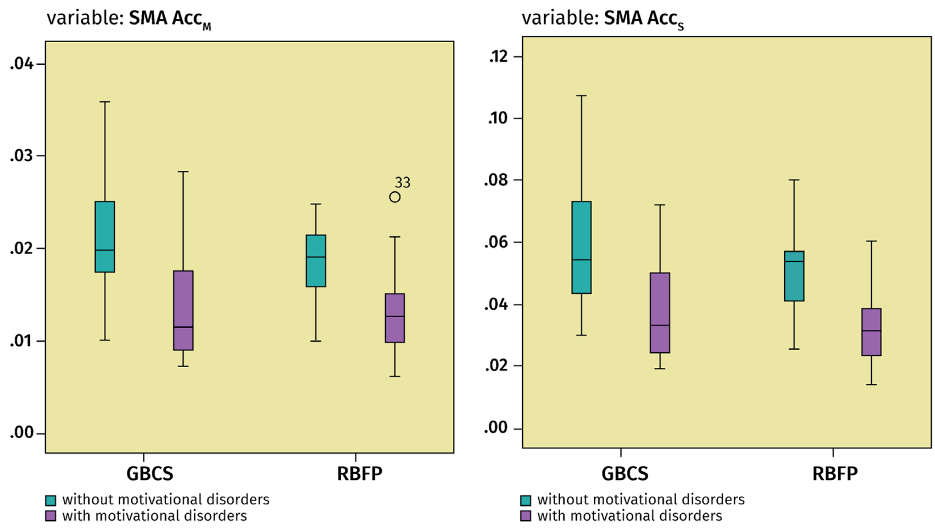


Figure F3. Effects of motivational disorders on quantity of movement (significance< .05). GBCS: game-based cognitive stimulation; RBFP: robot-based free play

Quantity of movement on the wrist was lower in participants with motivational disorders compared to participants without motivational disorders (see Figure F3).

For EDA (N_{motdis}= 12; $N_{w/omotdis}$= 20), results revealed:

1. A main effect of motivational disorders on the feature KURT EDA ($F(1,30)= 6.792$, $p< .05$, $\eta^2= .185$).

2. A main effect of motivational disorders on the feature SKEW EDA ($F(1,30)= 6.972$, $p< .05$, $\eta^2= .189$).

In an opposite trend, KURT EDA and SKEW EDA (see Figure F4) – which in game-based cognitive stimulation were positively correlated with alertness, and in robot-based free play were negatively correlated with attention and attitude towards the game – were higher in participants with motivational disorders during the former activity, and almost even between groups in the latter.

The same analyses were repeated using *dementia severity* (mild vs moderate) as between-subject factor and brought about only *two* significant results: (1) an interaction effect of activity type and dementia severity on the item *attitude toward partner* of the OME ($F(1,30)=$ 10.332, $p< .005$, $\eta^2= .205$), with participants with moderate dementia displaying a better

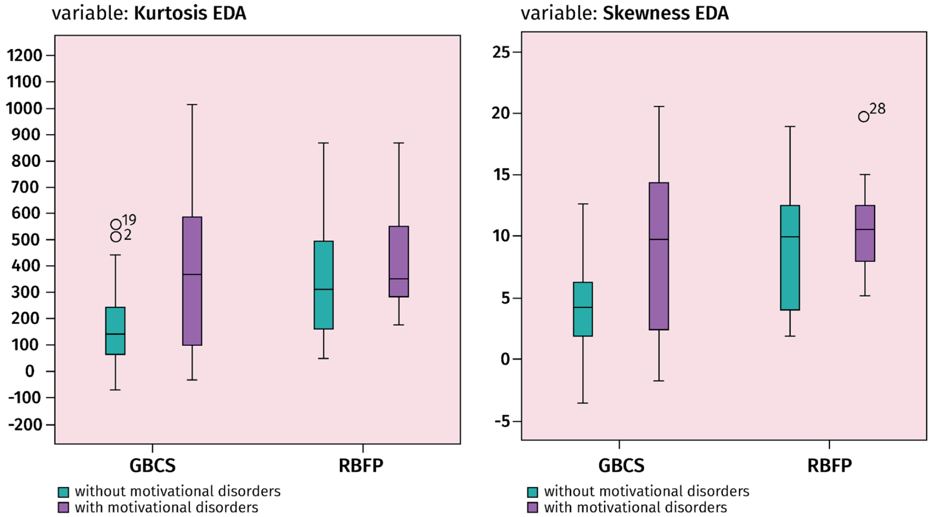


Figure F4. Effects of motivational disorders on EDA (significance< .05). GBCS: game-based cognitive stimulation; RBFP: robot-based free play

attitude toward the partner than participants with mild dementia during robot-based free play and (2) a main effect of dementia severity on SUM H EDA ($F(1,30)= 9.919$, $p< .005$, $\eta^2=.248$) with participants with moderate dementia having lower SUM H EDA with respect to participants with mild dementia across activities.

Appendix G. SEM Glossary

TERM	DEFINITION
Latent Variable	A construct that is not directly observable (e.g., intelligence, trust), but can be measured through observed variables. In path diagram notation, it is represented by an oval or circle.
Observed Variable	Also called *indicator*, it is a construct that can be directly measured. In path diagram notation, it is represented by a rectangle or square.
Endogenous Variable	A latent or observed variable that is caused by another latent or observed variable in the model (i.e., dependent variable). In path diagram notation, it has a directional path pointing at it.
Exogenous Variable	A latent or observed variable that is not caused by another latent or observed variable in the model (an independent variable). In path diagram notation, it does not have any path pointing at it.
Mediating Variable	A latent or observed variable through which a certain variable has an effect on another.
Directional Path	A regression path. In path diagram notation, it is represented by a single-headed linear arrow.
Covariance Path	A correlational path. In path diagram notation, it is represented by a double-headed curvilinear arrow.
Error Variance	The residual of a dependent variable. In path diagram notation, it is represented by a circular shape connected to the dependent variable with a directional path.
Model Specification	Development of the model based on relevant theory or past research with the specification of latent and observed variables and of the relationships between them.
Model Identification	The balance between known and unknown parameters (i.e., variables, paths) in the model. A model can be unidentified (knowns<unknowns), just identified (knowns=unknowns), and over-identified (knowns>unknowns).
Model Estimation	The estimation of unknown parameters in the model. It has to do with how the variance/covariance matrix given by the data relates to the variance/covariance matrix hypothesized by the model.
Model Testing	The assessment of whether the hypothesized set of pathways in the model are correct. It is achieved with the estimation of goodness of fit.
Goodness of Fit Index	*Absolute fit indexes* look at the ability of the model to reproduce the observed variance/covariance matrix. *Relative comparative fit indexes* involve comparing the hypothesized model with a baseline model. Good fit: X^2 with a $p > .05$; $RMSEA < .10$; $NFI > .90$; $CFI > .90$; $RFI > .90$
Model Modification	Modification of the model in case of scarce goodness of fit or misspecification.

Bibliography

Alam, M. A. U., Roy, N., Holmes, S., Gangopadhyay, A., & Galik, E. (2016). Automated functional and behavioral health assessment of older adults with dementia. In *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)* (pp. 140–149). IEEE.

Anderiesen, H. (2017). *Playful design for activation: co-designing serious games for people with moderate to severe dementia to reduce apathy*. Delft: Delft University of Technology (doctoral thesis).

Andreassi, J. L. (2010). *Psychophysiology: Human behavior and physiological response*. New York, New York: Psychology Press.

Attfield, S., Kazai, G., & Lalmas, M. (2011). Towards a science of user engagement (Position Paper). In *WSDM Workshop on User Modelling for Web Applications* (pp. 9-12). ACM.

Bakeman, R., and Gottman, J. (1987). Applying observational methods: A systematic view. In J. Osofsky (Ed.) *Handbook of Infant Development* (pp. 818–854). New York, New York: Wiley.

Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences*. Cambridge: Cambridge University Press.

Bakker, S., Hoven, E. van den, and Eggen, B. (2010). Design for the Periphery. In *Proceedings of the Eurohaptics 2010 Symposium on Haptic and Audio-Visual Stimuli: Enhancing Experiences and Interaction* (pp. 71-80).

Banerjee, S., Smith, S. C., Lamping, D. L., Harwood, R. H., Foley, B., Smith, P., Prince, M., Levin, E., Mann, M. & Knapp, M. (2006). Quality of life in dementia: more than just cognition. An analysis of associations with quality of life in dementia. *Journal of Neurology, Neurosurgery & Psychiatry*, *77*(2), 146-148.

Bartenieff, I. and Lewis, D. (2002). *Body Movement*. New York, New York: Routledge.

Bednarik, R., Eivazi, S., & Hradis, M. (2012). Gaze and conversational engagement in multiparty video conversation: an annotation scheme and classification of high and low levels of engagement. In *Proceedings of the 4th Workshop on Eye Gaze in Intelligent Human Machine Interaction* (pp. 10-16). ACM.

Beerens, H. C., Zwakhalen, S. M., Verbeek, H., Ruwaard, D., & Hamers, J. P. (2013). Factors associated with quality of life of people with dementia in long-term care facilities: a systematic review. *International Journal of Nursing Studies*, *50*(9), 1259-1270.

Bradley, K. (2008). *Rudolf Laban*. London: Routledge.

Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. In R. D. Lane, & L. Nadel (Eds.), *Cognitive Neuroscience of Emotion (*pp. 242-276). New York, New York: Oxford University Press.

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2013). A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology*, *49*(1), 1017-1034.

Brandtzæg, P. B., Følstad, A., & Heim, J. (2003). Enjoyment: lessons from Karasek. In Blythe, M. A., Overbeeke, K., Monk, A. F., Wright, P. C. (Eds.), *Funology: from Usability to Enjoyment* (pp. 56-65). Dodrecht: Springer.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research. Methodology in the social sciences*. New York, NY: The Guilford Press.

Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. In *CHI'04 Extended Abstracts on Human Factors in Computing Systems* (pp. 1297-1300). ACM.

Bryant, L. L., Corbett, K. K., & Kutner, J. S. (2001). In their own words: a model of healthy aging. *Social Science & Medicine*, *53*(7), 927-941.

Carvalho, M. B., Bellotti, F., Berta, R., De Gloria, A., Sedano, C. I., Hauge, J. B., Hu, J. & Rauterberg, M. (2015). An activity theory-based model for serious games analysis and conceptual design. *Computers & Education*, *87*, 166-181.

Castellano, G., Leite, I., Pereira, A., Martinho, C., Paiva, A., & McOwan, P. W. (2009a). It's all in the game: Towards an affect sensitive and context aware game companion. In *Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009* (pp. 1-8). IEEE.

Castellano, G., Pereira, A., Leite, I., Paiva, A., & McOwan, P. W. (2009b). Detecting user engagement with a robot companion using task and social interaction-based features. In *Proceedings of the 2009 International Conference on Multimodal interfaces - ICMI-MLMI '09* (pp. 119-126). ACM.

Cenfetelli, R. T., & Basselier, G. (2009). Interpretation of formative measurement in information systems research. *MIS Quarterly, 33*(4), 689–707.

Chapman, P., Selvarajah, S., & Webster, J. (1999). Engagement in multimedia training systems. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences, 1999. HICSS-32* (pp. 1-9). IEEE.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37-46.

Cohen-Mansfield, J. (2001). Managing agitation in elderly patients with dementia. *Geriatric Times*, *2*(3), 1-5.

Cohen-Mansfield, J., Dakheel-Ali, M., & Marx, M. (2009). Engagement in persons with dementia: The concept and its measurement. *The American Journal of Geriatric Psychiatry*, *17*(4), 299-307.

Cohen-Mansfield, J., Marx, M. S., Thein, K., & Dakheel-Ali, M. (2010a). The impact of past and present preferences on stimulus engagement in nursing home residents with dementia. *Aging and Mental Health*, *14*(1), 67-73.

Cohen-Mansfield, J., Thein, K., Dakheel-Ali, M., Regier, N. G., & Marx, M. S. (2010b). The value of social attributes of stimuli for promoting engagement in persons with dementia. *The Journal of Nervous and Mental Disease*, *198*(8), 586-592.

Cohen-Mansfield, J., Marx, M. S., Freedman, L. S., Murad, H., Regier, N. G., Thein, K., & Dakheel-Ali, M. (2011). The comprehensive process model of engagement. *The American Journal of Geriatric Psychiatry*, *19*(10), 859-870.

Cohen-Mansfield, J., & Werner, P. (1997). Management of verbally disruptive behaviors in nursing home residents. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *52*(6), 369-377.

Cruz, J., Marques, A., Barbosa, A., Figueiredo, D., & Sousa, L. (2011). Effects of a motor and multisensory-based approach on residents with moderate-to-severe dementia. *American Journal of Alzheimer's Disease & Other Dementias*, *26*(4), 282-289.

Csikszentmihalyi, M. (1975). Play and intrinsic rewards. *Journal of Humanistic Psychology*; *15*(3), 41–63.

Csikszentmihalyi, M. (1978). Attention and the holistic approach to behavior. In Pope, K. S. & Singer, J. L. (Eds.), *The stream of consciousness* (pp. 335–358). New York, New York: Plenum.

Csikszentmihalyi, M. (2000). The costs and benefits of consuming. *Journal of Consumer Research*, *27*(2), 267-272.

Csikszentmihalyi, M. (2008). *Finding flow*. New York, New York: Basic Books.

Csikszentmihalyi, M. (2014). *Flow and the Foundations of Positive Psychology*. Dordrecht: Springer.

Csikszentmihalyi, M., & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology*, *56*(5), 815-822.

Cummings, J., Mega, M., Gray, K., Rosenberg-Thompson, S., Carusi, D., & Gornbein, J. (1994). The Neuropsychiatric Inventory: Comprehensive assessment of psychopathology in dementia. *Neurology*, *44*(12), 2308-2308.

Dael, N., Mortillaro, M. and Scherer, K. (2012). The Body Action and Posture Coding System (BAP): Development and reliability. *Journal of Nonverbal Behavior*, *36*(2), 97-121.

Dautenhahn, K., & Werry, I. (2002). A quantitative technique for analysing robot-human interactions. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2002* (Vol. 2, pp. 1132-1138). IEEE.

David, R., Mulin, E., Friedman, L., Le Duff, F., Cygankiewicz, E., Deschaux, O., Garcia, R., Yesavage, J., Robert, P. and Zeitzer, J. (2012). Decreased daytime motor activity associated with apathy in Alzheimer's Disease: An actigraphic study. *The American Journal of Geriatric Psychiatry*, *20*(9), 806-814.

David, R., Rivet, A., Robert, P., Mailland, V., Friedman, L., Zeitzer, J. and Yesavage, J. (2010). Ambulatory actigraphy correlates with apathy in mild Alzheimer's disease. *Dementia*, *9*(4), 509-516.

Dawson, M. E., Schell, A. M., & Filion, D. L. (1990). The electrodermal system. In Cacioppo, J. T. & Tassinary, L. G. (Eds.), *Principles of psychophysiology: Physical, social, and inferential elements* (pp. 295-324). Cambridge: Cambridge University Press.

Dechamps, A., Alban, R., Jen, J., Decamps, A., Traissac, T., & Dehail, P. (2010). Individualized Cognition-Action intervention to prevent behavioral disturbances and functional decline in institutionalized older adults: a randomized pilot trial. *International Journal of Geriatric Psychiatry*, *25*(8), 850-860.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9-15). ACM.

Díaz-Boladeras, M, (2017) *Bonding with robotic pets. Children's cognitions, emotions and behaviors towards pet-robots. Applications in a robot assisted quality of life intervention in a pediatric hospital*. Barcelona: University of Barcelona (doctoral thesis).

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*(3), 285-290.

Douglas, Y., & Hargadon, A. (2000). The pleasure principle: immersion, engagement, flow. In *Proceedings of the eleventh ACM on Hypertext and Hypermedia* (pp. 153-160). ACM.

Drachen, A., Nacke, L. E., Yannakakis, G., & Pedersen, A. L. (2010). Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games* (pp. 49-54). ACM.

Eifert, G. H., & Wilson, P. H. (1991). The triple response approach to assessment: A conceptual and methodological reappraisal. *Behaviour Research and Therapy*, *29*(3), 283-292.

Ekman, P., & Rosenberg, E. L. (Eds.). (1997). *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. New York, New York: Oxford University Press.

Feng, Y., van Reijmersdal, R., Yu, S., Rauterberg, M., Hu, J., & Barakova, E. (2017). Dynamorph: Montessori inspired design for seniors with dementia living in long-term care facilities. In *International Conference on Intelligent Technologies for Interactive Entertainment* (pp. 49-58). Springer.

Feng, Y., Yu, S., van de Mortel, D., Barakova, E. I., Rauterberg, G. W. M., & Hu, J. (2018). Closer to nature: Multi-sensory engagement in interactive nature experience for seniors with dementia. In *Proceedings of the Sixth International Symposium of Chinese CHI* (pp. 49-56). ACM.

Fleiss, J. (1981). *Statistical Methods for Rates and Proportions*. New York, New York: Wiley.

Fosch-Villaronga, E. (2016). Legal Frame of Non-social Personal Care Robots. In *International Workshop on Medical and Service Robots* (pp. 229-242). Springer.

La France, M., & Ickes, W. (1981). Posture mirroring and interactional involvement: Sex and sex typing effects. *Journal of Nonverbal Behavior*, *5*(3), 139-154.

Gao, A. Y., Barendregt, W., & Castellano, G. (2017). Personalised human-robot co-adaptation in instructional Settings using reinforcement learning. In *Persuasive Embodied Agents for Behavior Change (PEACH2017) Workshop at the International Conference on Intelligent Virtual Agents (IVA2017)*.

Garbarino, M., Lai, M., Bender, D., Picard, R. W., & Tognetti, S. (2014). Empatica E3 – A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. In *Wireless Mobile Communication and Healthcare (Mobihealth), 2014 EAI 4th International Conference on* (pp. 39-42). IEEE.

Garson, G.D. (2008). *Logistic regression Stat Notes: Topics in multivariate analysis.* Retrieved from http://faculty.chass.ncsu.edu/garson/PA765/logistic.htm

Ge, B., Park, H. W., & Howard, A. M. (2016). Identifying engagement from joint kinematics data for robot therapy prompt interventions for children with autism spectrum disorder. In *International Conference on Social Robotics* (pp. 531-540). Springer.

Getzels, J. W., & Csikszentmihalyi, M. (1976). *The creative vision: A longitudinal study of problem finding in art*. New York, New York: Wiley.

Ghazali, A. S., Ham, J., Barakova, E., & Markopoulos, P. (2017a). The Influence of Social Cues and Controlling Language on Agent's Expertise, Sociability, and Trustworthiness. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 125-126). ACM.

Ghazali, A. S., Ham, J., Barakova, E. I., & Markopoulos, P. (2017b). Pardon the rude robot: Social cues diminish reactance to high controlling language. In *26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN* (pp. 411-417). IEEE.

Ghazali, A. S., Ham, J., Barakova, E., & Markopoulos, P. (2018a). The influence of social cues in persuasive social robots on psychological reactance and compliance. *Computers in Human Behavior*, *87*, 58-65.

Ghazali, A. S., Ham, J., Barakova, E. I., and Markopoulos, P. (2018b). Effects of robot facial characteristics and gender in persuasive Human-Robot Interaction. *Frontiers in Robotics and AI*, *5*, 73.

Giblin, L. B., De Leon, L., Smith, L. M., Sztynda, T., & Lal, S. (2013). Heart rate variability, blood pressure and cognitive function: Assessing age effects. *Journal of Green Engineering*, *3*, 347-361.

Giblin, L. B. (2016). *Effects of heart rate variability and blood pressure on cognition in healthy and clinical cohorts*. Sidney: University of Technology Sidney (Doctoral dissertation).

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (1998). *Multivariate data analysis*. Upper Saddle River, New Jersey: Prentice hall.

Hackney, P. (2002). *Making connections: Total body integration through Bartenieff fundamentals*. New York, New York: Routledge.

Hassenzahl, M. (2003). The thing and I: Understanding the relationship between user and product. In M. A. Blythe, A. F. Monk, K. Overbeeke, & P. C. Wright (Eds.), *Funology: from usability to enjoyment* (pp. 31–42). Dodrecht: Kluwer.

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. London: Sage.

Henriques, R., Paiva, A., & Antunes, C. (2013). Accessing emotion patterns from affective interactions using electrodermal activity. In *In 2013 Humaine Association Conference on Affective Computing and Interaction (ACII)* (pp. 43-48). IEEE.

IJsselsteijn, W., Poels, K., de Kort, Y.A.W. (2008). The Game Experience Questionnaire: Development of a Self-report Measure to Assess Player Experiences of Digital Games. Eindhoven: Eindhoven University of Technology.

Inoue, K., Lala, D., Nakamura, S., Takanashi, K., & Kawahara, T. (2016). Annotation and analysis of listener's engagement based on multi-modal behaviors. In *Proceedings of the Workshop on Multimodal Analyses enabling Artificial Agents in Human-Machine Interaction* (pp. 25-32). ACM.

Izard, C. E. (1971). *The face of emotion*. East Norwalk, Connecticut: Appleton-Century-Crofts.

Izard, C. E. (1972). *Patterns of emotions*. New York, New York: Academic Press.

Jacques, R. (1995). Engagement as a design concept for multimedia. *Canadian Journal of Educational Communication*, *24*(1), 49-59.

Jaques, N., McDuff, D., Kim, Y. L., & Picard, R. (2016). Understanding and predicting bonding in conversations using thin slices of facial expressions and body language. In *International Conference on Intelligent Virtual Agents* (pp. 64-74). Springer.

Järvelä, S., Kivikangas, J. M., Kätsyri, J., & Ravaja, N. (2014). Physiological linkage of dyadic gaming experience. *Simulation & Gaming*, *45*(1), 24-40.

Jones, C., Sung, B., & Moyle, W. (2015). Assessing engagement in people with dementia: a new approach to assessment using video analysis. *Archives of Psychiatric Nursing*, *29*(6), 377-382.

Jøranson, N., Pedersen, I., Rokstad, A. M. M., Aamodt, G., Olsen, C., & Ihlebæk, C. (2016). Group activity with Paro in nursing homes: systematic investigation of behaviors in participants. *International Psychogeriatrics*, *28*(8), 1345-1354.

Judge, K., Camp, C. and Orsulic-Jeras, S. (2000). Use of Montessori-based activities for clients with dementia in adult day care: Effects on engagement. *American Journal of Alzheimer's Disease*, *15*(1), 42-46.

Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, *64*(5), 402-406.

Karasek, R. & Theorell. T. (1990). Healthy work: Stress, productivity, and the reconstruction of working life. *The American journal of Public Health*, *80*, 1013-1014.

Kawaguchi, Y., Wada, K., Okamoto, M., Tsujii, T., Shibata, T., & Sakatani, K. (2012). Investigation of brain activity after interaction with seal robot measured by fNIRS. In *In 21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 571-576). IEEE.

Klimova, B., & Kuca, K. (2016). Speech and language impairments in dementia. *Journal of Applied Biomedicine*, *14*(2), 97-103.

Kubey, R., Larson, R., & Csikszentmihalyi, M. (1996). Experience sampling method applications to communication research questions. *Journal of Communication*, *46*(2), 99-120.

Kuhlmei, A., Walther, B., Becker, T., Müller, U. and Nikolaus, T. (2013). Actigraphic daytime activity is reduced in patients with cognitive impairment and apathy. *European Psychiatry*, *28*(2), 94-97.

Laban, R. (1966). *Choreutics*. London: Macdonald & Evans.

Lawton, M. P., Van Haitsma, K., & Klapper, J. (1996). Observed affect in nursing home residents with Alzheimer's disease. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *51*(1), 3-14.

Leite, I., Henriques, R., Martinho, C., & Paiva, A. (2013). Sensors in the wild: Exploring electrodermal activity in child-robot interaction. In *Proceedings of the 8th ACM/IEEE international conference on Human-Robot Interaction (HRI)* (pp. 41-48). ACM/IEEE.

Leite, I., Castellano, G., Pereira, A., Martinho, C., & Paiva, A. (2014). Empathic robots for long-term interaction. *International Journal of Social Robotics*, *6*(3), 329-341.

Levenson, R. W., & Gottman, J. M. (1983). Marital interaction: physiological linkage and affective exchange. *Journal of Personality and Social Psychology*, *45*(3), 587-597.

Liang, R. H., Yu, B., Xue, M., Hu, J., & Feijs, L. M. (2018, April). BioFidget: Biofeedback for Respiration Training Using an Augmented Fidget Spinner. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-12). ACM.

Logsdon, R. G., Gibbons, L. E., McCurry, S. M., & Teri, L. (1999). Quality of life in Alzheimer's disease: patient and caregiver reports. *Journal of Mental Health and Aging*, *5*(1), 21-32.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, *3*(2), 1-43.

Lysack, C. L., & Seipke, H. L. (2002). Communicating the occupational self: A qualitative study of oldest-old American women. *Scandinavian Journal of Occupational Therapy*, *9*(3), 130-139.

Maletic, V. (1987). *Body - Space - Expression. The Development of Rudolf Laban's Movement and Dance Concepts*. Berlin, Berlin: Mouton de Gruyter.

Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. (2006a). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & Information Technology*, *25*(2), 141-158.

Mandryk, R. L., Atkins, M. S., & Inkpen, K. M. (2006b). A continuous and objective evaluation of emotional experience with interactive play environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1027-1036). ACM.

Martin, P., & Bateson, P. P. G. (2007). *Measuring Behaviour: An Introductory Guide*. Cambridge : Cambridge University Press.

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion, 5*(2), 175-190.

McIntyre, G., & Howie, L. (2002). Adapting to widowhood through meaningful occupations: A case study. *Scandinavian Journal of Occupational Therapy*, *9*(2), 54-62.

Mishler, W., & Rose, R. (2005). What are the political consequences of trust? *Comparative Political Studies*, *38*(9), 1050-1078.

Mislevy, J. L., Rupp, A. A. (2010). Concurrent Validity. In Salkind, N. J. (Ed.), *Encyclopedia of Research Design* (pp. 210-211). Thousand Oaks, California: Sage.

Monkaresi, H., Bosch, N., Calvo, R. A., & D'Mello, S. K. (2017). Automated detection of engagement using video-based estimation of facial expressions and heart rate. *IEEE Transactions on Affective Computing*, *8*(1), 15-28.

Mota, S., & Picard, R. W. (2003). Automated posture analysis for detecting learner's interest level. In *2003 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'03)* (Vol. 5, pp. 49-49). IEEE.

Moyle, W., Venturto, L., Griffiths, S., Grimbeek, P., McAllister, M., Oxlade, D., & Murfield, J. (2011). Factors influencing quality of life for people with dementia: A qualitative perspective. *Aging & Mental Health*, *15*(8), 970-977.

Mulin, E., Leone, E., Dujardin, K., Delliaux, M., Leentjens, A., Nobili, F., Dessi, B., Tible, O., Agüera-Ortiz, L., Osorio, R. S., Yessavage, J., Dachevsky, D., Verhey, F. RJ., Cruz Jentoft, A. J., Blanc, O., Llorca, P. M., & Robert, P. H. (2011). Diagnostic criteria for apathy in clinical practice. *International Journal of Geriatric Psychiatry*, *26*(2), 158-165.

Mwangi, E., Barakova, E. I., Díaz-Boladeras, M., Mallofré, A. C., & Rauterberg, M. (2018). Directing attention through gaze hints improves task solving in Human-Humanoid Interaction. *International Journal of Social Robotics*, *10*(3), 1-13.

Nacke, L. E., Grimshaw, M. N., & Lindley, C. A. (2010). More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computers*, *22*(5), 336-343.

Nakamura, J., & Csikszentmihalyi, M. (2009). Flow theory and research. In S. Lopez & C. Snyder (Eds.), *Oxford Handbook of Positive Psychology* (pp. 195-206). New York, New York: Oxford University Press.

Nakano, Y. I., & Ishii, R. (2010). Estimating user's engagement from eye-gaze behaviors in human-agent conversations. In *Proceedings of the 15th International Conference on Intelligent User Interfaces* (pp. 139-148). ACM.

Negini, F., Mandryk, R. L., & Stanley, K. G. (2014). Using affective state to adapt characters, NPCs, and the environment in a first-person shooter game. In *Games Media Entertainment (GEM)* (pp. 1-8). IEEE.

Norman, K. L. (2013). Geq (game engagement/experience questionnaire): a review of two papers. *Interacting with Computers*, *25*(4), 278-283.

O'Brien, H. L. (2010). The influence of hedonic and utilitarian motivations on user engagement: The case of online shopping experiences. *Interacting with Computers*, *22*(5), 344-352.

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, *59*(6), 938-955.

O'Brien, H. L., Cairns, P., & Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies*, *112*, 28-39.

Olsen, C., Pedersen, I., Bergland, A., Enders-Slegers, M. J., & Ihlebæk, C. (2016). Engagement in elderly persons with dementia attending animal-assisted group activity. *Dementia*, *0*(0), 1-17.

Overbeeke, K., Djajadiningrat, T., Hummels, C., Wensveen, S., & Frens, J. (2003). Let's make things engaging. In M. A. Blythe, A. F. Monk, K. Overbeeke, & P. C. Wright (Eds.), *Funology: from usability to enjoyment* (pp. 7-17). Dordrecht: Kluwer.

Peters, C., Asteriadis, S., Karpouzis, K., & de Sevin, E. (2008). Towards a real-time gaze-based shared attention for a virtual agent. In *Workshop on Affective Interaction in Natural Environments (AFFINE), ACM International Conference on Multimodal Interfaces (ICMI'08)* (pp. 574-580). ACM.

Peters, C., Castellano, G., & de Freitas, S. (2009). An exploration of user engagement in HCI, in *Proceedings of the International Workshop on Affective Aware Virtual Agents and Social Robots* (pp. 9-11). ACM.

Peters, C., Pelachaud, C., Bevacqua, E., Mancini, M., & Poggi, I. (2005). A model of attention and interest using gaze behavior. In *International Workshop on Intelligent Virtual Agents* (pp. 229-240). Berlin, Berlin: Springer.

Petter, S., Straub, D., & Rai, A. (2007). Specifying Formative Constructs in Information Systems Research. *MIS Quarterly*, *31*(4), 623-656.

Phillips, A. C. (2011). Blunted cardiovascular reactivity relates to depression, obesity, and self-reported health. *Biological Psychology*, *86*(2), 106-113.

Phinney, A., Chaudhury, H., & O'connor, D. L. (2007). Doing as much as I can do: The meaning of activity for people with dementia. *Aging and Mental Health, 11*(4), 384-393.

Picard, R. W., Fedor, S., & Ayzenberg, Y. (2016). Multiple arousal theory and daily-life electrodermal activity asymmetry. *Emotion Review*, *8*(1), 62-75.

Pino, M., Boulay, M., Jouen, F., & Rigaud, A. S. (2015). "Are we ready for robots that care for us?" Attitudes and opinions of older adults toward socially assistive robots. *Frontiers in Aging Neuroscience*, *7*, 141.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715-734.

Prince, M., Wimo, A., Guerchet, M., Gemma-Claire, A., Wu, Y.-T., & Prina, M. (2015). *World Alzheimer Report 2015: The Global Impact of Dementia - An analysis of prevalence, incidence, cost and trends*. London: Alzheimer's Disease International.

Prince, M. J., Wu, F., Guo, Y., Robledo, L. M. G., O'Donnell, M., Sullivan, R., & Yusuf, S. (2015). The burden of disease in older people and implications for health policy and practice. *The Lancet*, *385*(9967), 549-562.

Ravaja, N. (2004). Contributions of psychophysiology to media research: Review and recommendations. *Media Psychology*, *6*(2), 193-23.

Ravaja, N., Turpeinen, M., Saari, T., Puttonen, S., & Keltikangas-Järvinen, L. (2008). The psychophysiology of James Bond: Phasic emotional responses to violent video game events. *Emotion*, *8*(1), 114-120.

Read, J. C., MacFarlane, S. J., & Casey, C. (2002). Endurability, engagement and expectations: Measuring children's fun. In *Interaction design and children*. Shaker Publishing Eindhoven.

Reisberg, B., Ferris, S., De Leon, M. and Crook, T. (1982). The Global Deterioration Scale for assessment of primary degenerative dementia. *American Journal of Psychiatry*, *139*(9), 1136-1139.

Rich, C., Ponsler, B., Holroyd, A., & Sidner, C. L. (2010). Recognizing engagement in human-robot interaction. In *Proceedings of the 5th ACM/IEEE international conference on Human-Robot Interaction (HRI)* (pp. 375-382). ACM/IEEE.

Ritchey, L. H., Ritchey, L. H., & Dietz, B. E. (2001). Clarifying the measurement of activity. *Activities, Adaptation & Aging*, *26*(1), 1-21.

Robert, P. H., Verhey, F. R., Byrne, E. J., Hurt, C., De Deyn, P. P., Nobili, F., Riello, R, Rodriguez, G., Frisoni, G., Tsolaki, M., Kyriazopoulou, N., Bullock, R., Burns, A. & Vellas, B. (2005). Grouping for behavioral and psychological symptoms in dementia: clinical and biological aspects. Consensus paper of the European Alzheimer disease consortium. *European Psychiatry*, *20*(7), 490-496.

Robert, P., Mulin, E., Malléa, P., & David, R. (2010). Review: Apathy Diagnosis, Assessment, and Treatment in Alzheimer's Disease. *CNS Neuroscience & Therapeutics*, *16*(5), 263-271.

Rudovic, O., Lee, J., Dai, M., Schuller, B., & Picard, R. (2018). Personalized machine learning for robot perception of affect and engagement in autism therapy. *arXiv preprint arXiv:1802.01186.*

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161-1178.

Russell, J. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145-172.

Šabanović, S., Bennett, C. C., Chang, W. L., & Huber, L. (2013). PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. In *2013 IEEE International Conference on Rehabilitation Robotics (ICORR)* (pp. 1-6). IEEE.

Salam, H., Celiktutan, O., Hupont, I., Gunes, H., & Chetouani, M. (2017). Fully automatic analysis of engagement and its relationship to personality in human-robot interactions. *IEEE Access*, 5, 705-721.

Sanghvi, J., Castellano, G., Leite, I., Pereira, A., McOwan, P. W., & Paiva, A. (2011). Automatic analysis of affective postures and body motion to detect engagement with a game companion. In *Proceedings of the 6th International Conference on Human-Robot Interaction (HRI)* (pp. 305-312). ACM/IEEE.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. New York, New York: CRC press.

Schell, J. (2014). *The art of game design*. New York, New York: A K Peters/CRC Press.

Schelle, K. J., Gomez Naranjo, C., ten Bhömer, M., Tomico, O., & Wensveen, S. (2015). Tactile dialogues: Personalization of vibrotactile behavior to trigger interpersonal communication. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction* (pp. 637-642). ACM.

Scherder, E. J., Bogen, T., Eggermont, L. H., Hamers, J. P., & Swaab, D. F. (2010). The more physical inactivity, the more agitation in dementia. *International Psychogeriatrics*, *22*(8), 1203-1208.

Schulter, G., & Papousek, I. (1992). Bilateral electrodermal activity: reliability, laterality and individual differences. *International Journal of Psychophysiology*, *13*(3), 199-213.

Schumacker, R. and Lomax, R. (2016). *A beginner's guide to structural equation modeling*. New York, New York: Routledge.

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: an Introduction. *American Psychologist*, *55*(1), 5-14.

Sidner, C. L., Lee, C., Kidd, C. D., Lesh, N., & Rich, C. (2005). Explorations in engagement for humans and robots. *Artificial Intelligence*, *166*(1-2), 140-164.

Silveira, F., Eriksson, B., Sheth, A., & Sheppard, A. (2013). Predicting audience responses to movie content from electro-dermal activity signals. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 707-716). ACM.

Silverstein, M., & Parker, M. G. (2002). Leisure activities and quality of life among the oldest old in Sweden. *Research on Aging*, *24*(5), 528-547.

Slater, M., Lotto, B., Arnold, M. M., & Sánchez-Vives, M. V. (2009). How we experience immersive virtual environments: the concept of presence and its measurement. *Anuario de Psicología*, *40*, 193-210.

Stone, D., Jarrett, C., Woodroffe, M., & Minocha, S. (2005). *User interface design and evaluation*. San Francisco, California: Elsevier.

Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics*. Boston, Massachusetts: Pearson/ Allyn & Bacon.

Taelman, J., Vandeput, S., Spaepen, A., & Van Huffel, S. (2009). Influence of mental stress on heart rate and heart rate variability. In *4th European conference of the international federation for medical and biological engineering* (pp. 1366-1369). Springer.

Takayanagi, K., Kirita, T., & Shibata, T. (2014). Comparison of verbal and emotional responses of elderly people with mild/moderate dementia and those with severe dementia in responses to seal robot, PARO. *Frontiers in Aging Neuroscience*, *6*, 257.

Thompson, I. (1987). Language in dementia part I: A review. *International Journal of Geriatric Psychiatry*, *2*(3), 145-161.

Tiberio, L., Cesta, A., Cortellessa, G., Padua, L., & Pellegrino, A. R. (2012). Assessing Affective Response of Older Users to a Telepresence Robot Using a Combination of Psychophysiological Measures. In *21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN).* (pp. 833-838). IEEE.

Tickle-Degnen L (2006) Nonverbal behavior and its functions in the ecosystem of rapport. In: Manusov V, Patterson M (eds) *The SAGE Handbook of Nonverbal Communication*. Thousand Oaks, California: Sage.

Tickle-Degnen, L. & Rosenthal, R. (1990). The Nature of Rapport and Its Nonverbal Correlates. *Psychological Inquiry*, *1*(4), 285-293.

Tomkins, S.S. (1984). Affect theory. In P. Ekman (Ed.), *Emotion in the human face* (2nd ed.) (pp. 353-395). New York, New York: Cambridge University Press.

Treusch, Y., Page, J., van der Luijt, C., Beciri, M., Benitez, R., Stammler, M., & Marcar, V. L. (2015). Emotional Reaction in Nursing Home Residents with Dementia-associated Apathy: A pilot study. *Geriatric Mental Health Care*, *3*(1), 1-6.

UK Department of Health (2009). *Living well with dementia*. London: Dept. of Health.

Van Den Bosch, I., Salimpoor, V., & Zatorre, R. J. (2013). Familiarity mediates the relationship between emotional arousal and pleasure during music listening. *Frontiers in Human Neuroscience*, *7, 534*.

Vila-Miravent, J., Villar, F., Celdrán, M., & Fernández, E. (2012). El modelo de la atención centrada en la persona: análisis descriptivo de una muestra de personas mayores con demencia en centros residenciales. *Aloma: Revista de Psicologia, Ciències de l'Educació i de l'Esport Blanquerna*, *30*(1), 109-117.

Vinyoles Bargalló, E., Vila Domènech, J., Argimon Pallàs, J., Espinàs Boquet, J., Abos Pueyo, T. and Limón Ramírez, E. (2002). Concordancia entre el Mini-Examen Cognoscitivo y el Mini-Mental State Examination en el cribado del déficit cognitivo. *Atención Primaria*, *30*(1), 5-13.

Volkers, A. C., Tulen, J. H., van den Broek, W. W., Bruijn, J. A., Passchier, J., & Pepplinkhuizen, L. (2003). Motor activity and autonomic cardiac functioning in major depressive disorder. *Journal of Affective Disorders*, *76*(1), 23-30.

Vygotsky, L. (1987). *Mind in society: The development of higher psychological processes*. Cambridge, Massachusetts: Oxford University Press.

Wada, K., Ikeda, Y., Inoue, K., & Uehara, R. (2010). Development and preliminary evaluation of a caregiver's manual for robot therapy using the therapeutic seal robot Paro. In *19th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 533-538). IEEE.

Wada, K., Shibata, T., Musha, T., & Kimura, S. (2005). Effects of robot therapy for demented patients evaluated by EEG. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005 (IROS)* (pp. 1552-1557). IEEE.

Wada, K., & Shibata, T. (2008). Social and Physiological Influences of Robot Therapy in a Care House. *Interaction Studies*, *9*(2), 258-276.

World Health Organization (2015). *World Health Report on Ageing and Health*. Geneva: World Health Organization.

World Health Organization & Alzheimer's Disease International (2012). *Dementia: A Public Health Priority*. Geneva: World Health Organization and Alzheimer's Disease International.

Wile, C., & Cook, R. (2010). *Moving About. Capturing Movement Highlights Using Motif Notation.* New York, New York: Charlotte Wile Publisher.

Wrobel, J., Wu, Y. H., Kerhervé, H., Kamali, L., Rigaud, A. S., Jost, C., Le Pévédic, B. & Duhaut, D. (2013). Effect of agent embodiment on the elder user enjoyment of a game. In *ACHI 2013-The Sixth International Conference on Advances in Computer-Human Interactions* (pp.162-167). IARIA.

Wu, Y., Fassert, C., & Rigaud, A. (2012). Designing robots for the elderly: Appearance issue and beyond. *Archives of Gerontology and Geriatrics*, *54*(1), 121-126.

Yazdanyar, A., & Newman, A. B. (2009). The burden of cardiovascular disease in the elderly: morbidity, mortality, and costs. *Clinics in geriatric medicine*, *25*(4), 563-577.

Yu, B., Hu, J., Funk, M., Liang, R. H., Xue, M., & Feijs, L. (2018). RESonance: Lightweight, Room-Scale Audio-Visual Biofeedback for Immersive Relaxation Training. *IEEE Access*, *6*, 38336-38347.

Yu, R., Hui, E., Lee, J., Poon, D., Ng, A., Sit, K., Ip, K., Yeung, F., Wong, M., Shibata, T. & Woo, J. (2015). Use of a therapeutic, socially assistive pet robot (PARO) in improving mood and stimulating social interaction and communication for people with dementia: Study protocol for a randomized controlled trial. *JMIR Research Protocols*, *4*(2).

Zajonc, R. B. (1965). Social facilitation. *Science*, *149*(3681), 269-274.

Acknowledgements

As I look back to these four years, a lot of images, stories, and names come to my mind. I am grateful to the people that walked this path with me, to those that I encountered at the crossroads along the way and decided to accompany me for a while, to those that did not manage to see me cross the finish line but made their way into my heart.

First of all, I would like to thank my parents. I am beyond grateful to you. I would have never managed to complete this PhD without your relentless support. Thanks for tolerating me during our infinite phone calls on the way to the university, for listening to me during my outbursts of anxiety, and for believing in me before I was able to do it myself. This accomplishment is yours as much as mine.

To Rolf, my partner and love. Thank you for teaching me the power of kindness and the value of silence. For showing me that actions speak louder than words and do not need to be noticed to bring us pleasure. Thanks for being the safe haven to my stormy character.

I felt privileged to have such a team of supervisors. Prof. Andreu Català-Mallofré, thank you for your fatherly supervision, for your ability to understand my work ethic from the very beginning, and, most notably, for believing in crazy ideas. Prof. Marta Dìaz-Boladeras, thank you for your tireless guidance, for the brainstorming sessions that put me on the right track (I still remember when this model was just a sketch on a whiteboard), thanks for pushing me to reflect on any single detail of my research, even on those that appeared to be self-evident, and, also, thanks for supporting me as a human being when I felt discouraged and insecure. Dr. Emilia Barakova, thanks for your availability and reliability, whenever I had a problem, I knew I could lean on you, thanks for the countless unscheduled meeting that we had to discuss every minute element of my project. Prof. Matthias Rauterberg, thanks for the challenges you put on my way, for ensuring that I always had clear in my mind the final objective of my research and that I could defend my positions in a scientific and factual way. Also, I would like to thank my future supervisor, Prof. Ginevra Castellano, thanks for offering me a wonderful postdoc position at Uppsala University and for giving me the time to complete this PhD in the best possible way.

I would like to express my gratitude to the reviewing committee: Prof. Panos Markopoulos, Prof. Anne-Sophie Rigaud, and Prof. Loe Feijs. Thanks for taking the time to read my thesis and provide so many insightful and encouraging comments. Thanks also to Prof. Rosalind Picard, who greatly inspired this work with her research on autism, thanks for taking the time to read my papers and offer me a precious feedback.

I am beyond grateful to three very special persons that collaborated with me during this PhD: Dr. Daniel Rodríguez-Martín, Roos van Berkel, and Dr. Ruth Dolado i Guivernau. Thanks for the endless exchanges on this research project, for the inspiring conversations, and for the fruitful feedback on my work.

Also, I would like to thank the nursing homes Redós de Sant Josep i Sant Pere and La Mallola for the invaluable support and inspiration. In particular, I am deeply grateful to four special persons without whom this research project could have never been accomplished: Anna Barea, Laia Aranda, Elisabeth Seguer, and Neus Sanchez. I take this opportunity to thank all the participants to my studies, I cherish every moment I spent with you, thanks for your patience with my bad Spanish, for sharing your amazing stories with me, and for holding my hand in search of comfort. I sincerely appreciated the help of Prof. Josep Vila-Miravent and Elena Fernàndez-Gamarra of the foundation Alzheimer Catalunya, thanks for understanding the potential of this research project and assist me in finding nursing homes willing to participate in my studies.

I dedicate a special thought to my cousin Sara, her husband Reza, and the little Leila, in a world that tries to build walls to divide us, you are a wonderful message of hope. Thanks to my uncles - Pino and Sandra - and to two special persons that are family to me - Lucia and Laura. You were always so interested in my work and found the time to cheer me up.

I would like to thank Bibi and Bernie, thanks for teaching me what it means to love and support each other through hardship, with courage and discretion. Thanks for treating me like family and for always ensuring that I had the best German experience, full of food and ice cream. I am also thankful to Gaby and Mario, for the cheerful dinners and the endless jokes. A special thanks goes to Oma Maja and Oma Rosa, not everybody has the luck to have two additional grandmas in its life.

This PhD was an incredible experience, precious, and nourishing also for the friends I encountered on the way. It was an honor for me to come across so many cultures, religions, and gastronomic backgrounds. Thanks to Sarang, Sophia, Kadian, Aimi, Fatihah, Begum, Kyra, Fabienne, Jessica, Ruud, Girmaw, Mehrnoosh, Yuan, Maira, Lavender, Eunice, Oscar, Sebastian, and Le. You really made me a richer person.

My last but not least important thanks goes to Sabine. Thanks for the way you carry out your work, in a precise and timely manner, but, above all, thanks for the luminous person you are, always ready to worry for the others and help them as much as you can.

To Future PhD Students

The founder of Lego - Ole Kirk Kristiansen - used to say: “Only the best is good enough” (in Danish “Det Bedste er ikke for godt”). I guess this motto summarizes very well my attitude towards research. This PhD has been very tough, probably the toughest thing I have done so far. When I started, I seriously doubted I could manage to complete it.

I am not an engineer, I am not a designer, and I am not a computer scientist. Probably, the closest thing to what I am is a psychologist. However, I am not a psychologist either. During these years, people often reminded me that I did not belong, that I was some kind of strange by-product of different scientific backgrounds. In those moments, I kept in mind the words that my supervisor Andreu told me at the very beginning of my PhD - “Your background is your jewel, it is the most precious thing that you have, you need to treasure it”. Those words resonated in me in the darkest moments and pushed me to keep working harder and harder in spite of challenges.

Since I was a child, my parents always taught me not to give up and to be independent. I distinctly remember one episode from my childhood. In the kitchen, we had this cabinet with drawers. There were six or seven drawers in the cabinet, the top two drawers were very hard to open and I could only reach them by staying on my toes. One day, I needed something from the top drawer. I tried several times to open it without success when I turned to my mom asking her to do it for me. Her reply was, as always - “Try and try! Never give up! Sooner or later you will manage!” (words that actually rhyme in Italian: “Vai e vai! Non ti arrendere mai! Prima o poi ce la farai!”). I kept forcing the drawer until it opened up and she shouted: “… and you managed!” (“E ce l’hai fatta!”). You cannot imagine how much I hated this tune as a child, but it built my character. It taught me that you cannot reach anything without sacrifice and persistence. I identify myself in those words now.

Doing an international PhD was a dream for me. It did not come to me casually, I spent a lot of time preparing for it and looking for it. Plenty of unpaid internships, lots of extracurricular study and tutoring for other students. Once, in a job interview, I was asked: “What is your dream job for the future?” I replied: “Being paid to study!” I guess I made it.

To conclude, I have a message for all the PhD students that are starting now: always remember who you are and what you love, because with an idea, somebody that believes in you, and stamina you can achieve anything. In a recent Ted Talk, Angela Lee Duckworth - a researcher from the University of Pennsylvania - explained that the most successful students are not necessarily the most talented ones, but those that show grit: “Grit is passion and perseverance for very long-term goals. Grit is having stamina. Grit is sticking with your future, day in, day out, not just for the week, not just for the month, but for years, and working really hard to make that future a reality. Grit is living life like it’s a marathon, not a sprint”. It will be hard, it is meant to be hard, it has been hard for all of us, but there is good news: YOU CAN MANAGE.

Curriculum Vitae

Giulia Perugia was born on April 11th, 1985, in Rome, Italy. She received a Bachelor's degree (BA) in Literature and Linguistics from the University of Roma Tre, Rome, in 2011. She continued her education by pursuing a Master's Degree (MSc) in Cognitive Science at the University of Siena and graduated magna cum laude in 2013 with a thesis focused on the use of social robots to improve language production in people with dementia. During her Master's years, she collaborated in several research projects conducted by the Robotics and Learning Technologies Lab of the University of Siena and was appointed university tutor for the inmate students in the prison of San Gimignano.

Since October 2014, she has been a doctoral candidate in the Erasmus Mundus Joint Doctorate in Interactive and Cognitive Environments in the research area of Assistive Technologies. She spent the first two years of her PhD at the Technical Research Centre for Dependency Care and Autonomous Living of the Technical University of Catalonia (UPC), in Spain. During the second half of her doctoral studies, she moved to the Netherlands and became part of the Designed Intelligence research group at the Department of Industrial Design of Eindhoven University of Technology. In her PhD project, she focused on building and testing a measurement framework of engagement for people with dementia. During her PhD, in 2017, she was awarded the Distinguished Interdisciplinary Research Award by the Korean Robotic Society (KROS) and the Robotics Society of Japan (RSJ) for her pioneering in the field measurement of psychophysiology in people with dementia.

Publications

Journals

1. Perugia, G., Díaz-Boladeras, M., Català-Mallofré, A., Barakova, E., & Rauterberg, M. (2018). An interdisciplinary critical review on engagement: Definitions, components, and measures. *The Journal of Positive Psychology* (to be submitted).

2. Perugia, G., Díaz-Boladeras, M., Català-Mallofré, A., Barakova, E., & Rauterberg, M. (2018). ENGAGE-DEM: A model of engagement of people with dementia. *IEEE Transactions on Affective Computing* (to be submitted).

3. Perugia, G., van Berkel, R., Díaz-Boladeras, M., Català-Mallofré, A., Rauterberg, M., & Barakova, E. (2018). Understanding engagement in dementia through behavior. The Ethographic and Laban-Inspired Coding System of Engagement (ELICSE) and the Evidence-Based Model of Engagement-Related Behavior (EMODEB). *Frontiers in Psychology*, *9*, 690.

4. Perugia, G., Rodríguez-Martín, D., Díaz-Boladeras, M., Català-Mallofré, A., Barakova, E., Rauterberg, M. (2017) Quantity of movement as a measure of engagement for dementia. The influence of motivational disorders. *American Journal of Alzheimer's Disease and Other Dementias*, *33*(2), 112-121.

Conferences (peer-reviewed)

1. Perugia, G., Rodríguez-Martín, D., Díaz-Boladeras, M., Català-Mallofré, A., Barakova, E., & Rauterberg, M. (2017). Electrodermal activity: Explorations in the psychophysiology of engagement with social robots in dementia. In *26th IEEE International Symposium on Robot and Human Interactive Communication* (RO-MAN) (pp. 1248-1254). IEEE. **[KROS/RSJ Distinguished Interdisciplinary Research Award].**

2. Perugia, G., Díaz-Boladeras, M., Català-Mallofré, A., Rauterberg, M., & Barakova, E. (2017). Modelling engagement in dementia through behaviour. Contribution for socially interactive robotics. In *2017 IEEE International Conference on Rehabilitation Robotics* (ICORR) (pp. 1112-1117). IEEE.

3. Perugia, G., Díaz-Boladeras, M., Barakova, E., Català-Mallofré, A., & Rauterberg, M. (2017). Social HRI for People with Dementia: One Size Fits All? In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 257-258). ACM.

Other

1. Perugia, G., Díaz-Boladeras, M., Barakova, E., Català-Mallofré, A., & Rauterberg, M. (2016) ENGAGE, Modeling engagement in people with dementia. In *4th IEEE-EURASIP Summer School on Signal Processing 2016*.

2. Perugia, G., Díaz-Boladeras, M., Barakova, E., Català-Mallofré, A., & Rauterberg, M. (2015) Distinguishing apathy and depression through their behavioral symptoms: the role of smart sensing technologies in dementia. In *ICE Summer School 2015*.

3. Perugia, G., Díaz-Boladeras, M., Català-Mallofré, A., (2015). ENGAGE: an Evidence-based Model of Engagement for Dementia. In *Actas de las XVII Jornadas de ARCA 2015* (pp. 22-25).