Guiding older mourners from online self-help to offline support using algorithmic mental health monitoring

Lena Brandl

GUIDING OLDER MOURNERS FROM ONLINE SELF-HELP TO OFFLINE SUPPORT USING ALGORITHMIC MENTAL HEALTH MONITORING

Lena Brandl

The research in this dissertation was financially supported by the AAL Programme under project number AAL-2019-6-168-CP (optimizing the menta**L** health and resilienc**E** of older **A**dults that ha**V**e lost th**E**ir spou**S**e via blended, online therapy - **LEAVES**) with funding by the European Union and the national funding agencies from the Netherlands, Portugal, and Switzerland: The Netherlands Organization for Health Research and Development (ZonMW), Fundação para a Ciência e Tecnologia (FCT), and Innosuisse–Swiss Innovation Agency.

The research reported in this dissertation was carried out at Roessingh Research and Development and at the Human Media Interaction group of the University of Twente. The publication of this thesis was financially supported by both organizations.

Cover design: Lena Brandl and Hieu Nguyen
Printed by: Ipskamp Printing
Lay-out: Tessa Beinema and Lena Brandl
ISBN (print): 978-90-365-6435-9
ISBN (digital): 978-90-365-6436-6
URL: https://doi.org/10.3990/1.9789036564366

The design of the thesis cover was inspired by Greek mythology: the flowing lines symbolize the river Styx. The yellow circle represents a mourner, standing beside river Styx, a silent spectator of colorful souls drifting toward the afterlife on the river.

GUIDING OLDER MOURNERS FROM ONLINE SELF-HELP TO OFFLINE SUPPORT USING ALGORITHMIC MENTAL HEALTH MONITORING

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof. dr. ir. A. Veldkamp,
on account of the decision of the Doctorate Board,
to be publicly defended
on Thursday the 30th of January 2025 at 12.45 hours

by

Lena Brandl
born on the 4th of October, 1993
in Essen, Germany

This dissertation has been approved by:

Promoter:
prof. dr. D.K.J. Heylen

Co-promoter:
dr. S.M. Jansen-Kosterink

Graduation committee

Chair / secretary:
prof. dr. ir. B.R.H.M Haverkort

Promoter:
prof. dr. D.K.J. Heylen
University of Twente, EEMCS, Human Media Interaction

Co-promoter:
dr. S.M. Jansen-Kosterink
University of Twente, EEMCS, Biomedical Signals and Systems &
Roessingh Research and Development

Committee Members:
prof. dr. ir. M. Tabak
University of Twente, EEMCS, Biomedical Signals and Systems

prof. dr. M.L. Noordzij
University of Twente, BMS, Psychology, Health & Technology

prof dr. A. A. Salah
Utrecht University, Computer Science

prof. dr. S. Ben Allouch
University of Amsterdam

dr. ir. D.J. Broekens
Leiden University, LIACS

Paranymphs:
dr. C. Rikhof
L. Schoot Uiterkamp, MSc.

Contents

General introduction

Prologue

While booting her computer, Orphea takes a sip of her freshly brewed herbal tea and looks out the window upon the driveway in front of the house. It is late April and small buds of purple and white announce the beginning of the imminent colorful metamorphosis of her front yard. “Look at that, dear, has it been a year already?", she muses while taking another sip. “Soon the rhododendrons will bloom. This will be the first time in...ever, I guess. The first time that we won’t see them come into bloom together." Orphea puts her cup on the desk and starts typing the address of LEAVES, the e-learning environment about re-defining one’s life after loss that Dr. Epione recommended. “Good morning, Orphea.", the virtual assistant on the web page welcomes her, “What do you want to work on today?".

1.1 eMental health to prevent the exacerbation of mental health problems

The loss of a spouse is an impactful and common occurrence in later life. Most bereaved people eventually come to terms with their loss and, after a period of adjustment, continue to lead a normal and fulfilling life. However, some people (about 9% of the bereaved population according to a recent prevalence study (Wilson et al., 2022)) have difficulties overcoming bereavement, resulting in an exacerbation of their grief symptoms over time. Starting from the 5th edition of the Diagnostic and Statistical Manual of Psychological Disorders (DSM-5) (American Psychiatric Association, 2013), this condition has been recognized as *persistent complex bereavement disorder*. Its equivalent *prolonged grief disorder* became a diagnostic code (6B42) in the 11th version of the World Health Organization's International Classification of Diseases (ICD-11) (World Health Organization, 2019). The fact that persistent grief has only recently been recognized as a mental health disorder exemplifies how challenging it can be to draw the line between individual coping and mental illness. Persistent grief is also a good example for how underlying and unresolved personal issues can gradually exacerbate into mental health problems if left unattended.

Despite being a life event that sooner or later affects most during their lifetime, we appear to live in a "grief-denying" society (Macdonald, 2019) in which socially awkward and maladapted grief responses are the norm. Bereaved people frequently experience that friends and colleagues lack compassion and the knowledge how to be supportive in the face of bereavement, often resulting in the mourner becoming socially and emotionally isolated after their loss (Breen et al., 2022). In a call to action to educate communities to become more grief literate, Breen et al. (2022) argue that the institutionalization and professionalization of bereavement care as something that is almost exclusively offered to those who have lost someone contributes to the inability of communities, colleagues and friends to be supportive and shifts responsibility away from support networks to the bereaved individual. As a consequence, on top of dealing with their loss, mourners often become responsible for managing social interactions with friends and colleagues to minimize the awkwardness that arises in their interactions due to their loss.

Taking some responsibility off the shoulders of the bereaved entails increasing grief literacy in the general population. According to Breen et al. (2022), grief literacy entails the capacity to gather and use knowledge regarding loss experiences, including the knowledge to facilitate understanding and reflection, the skills to take action, and values that inspire compassion and care. Regarding the knowledge component of grief literacy, earlier research suggests that all mourners benefit from information about grief and loss, regardless of the severity of their grief symptoms (Aoun et al., 2018; Breen et al., 2022). However, for those in need of professional mental health support, knowing that they are, in fact, dealing with a mental health issue becomes a pre-requisite for help-seeking. A person's mental health literacy strongly influences whether a person receives the help that they need (Clement et al., 2015; Elshaikh et al., 2023; Gulliver et al., 2010). The term mental health literacy complements the more specific term of grief literacy and it is commonly used to summarize a person's knowledge of positive mental health as well as mental illness, their attitudes towards mental illness and their capacity to seek mental health support if ever they find themselves in need of it (Gulliver et al., 2010).

In recent years, eMental health applications have received increasing attention for improving the accessibility of mental healthcare (Borghouts et al., 2021; Leung et al., 2022), and have been shown to be effective for treating various mental illnesses, including depression, anxiety and prolonged grief disorder (Brodbeck et al., 2019; Carlbring et al., 2011; Schröder et al., 2016). *eMental health application* is an umbrella term that encompasses a range of mental health smartphone apps, internet websites, wearable devices, virtual reality, or video games (Mohr et al., 2013) and self-guided online services as well as those blended with human support or traditional psychotherapy (Borghouts et al., 2021). Alternative terms that are often used interchangeably include *digital mental health interventions* (DMHIs) and *internet interventions*. Design, content, target groups and the extent to which the mental health benefits of the service have been tested empirically differ tremendously between online mental health services. Many eMental health applications combine information about a specific mental health issue, mental illness or general information about maintaining positive mental health with reflective exercises to apply the acquired knowledge to the user's situation.

The digital format of eMental health applications is promising for increasing the accessibility and convenience of mental healthcare, but online mental health services introduce challenges of their own, including low adherence (Borghouts et al., 2021; Jagayat et al., 2024; Leung et al., 2022). While adherence, or *drop out*, is not only a problem in eMental health, but a challenge for mental healthcare in general, attrition rates for some internet interventions have been alarming with up to 74% of study participants being lost to attrition (Richards & Richardson, 2012). Drop out undermines the potential usefulness of any kind of healthcare service. There is some evidence that the risk of discontinuing an online mental health service increases with mental illness severity (Christensen et al., 2009). Indeed, self-guided eMental health applications have repeatedly been shown to be most effective for treating mild to moderate mental health problems (Zuelke et al., 2021), indicating that professional intervention is warranted when people suffer from more severe and complex mental health problems.

A specific form of mental healthcare that tailors care based on symptom severity is *stepped care*. In stepped care, every user receives some form of symptom treatment, starting with the least intensive intervention in terms of care resources and costs, such as self-help materials and psychoeducation. More intensive levels of care can entail a form of (para-)professional intervention, such as a telephone call with a non-clinical eCoach, or an appointment with a psychiatrist (Jagayat et al., 2024). In eMental health, more intensive levels of stepped care are examples of blending online self-guided mental health services with professional support. Support on-demand (Dahlin et al., 2020; Oromendia et al., 2016) is another way of stepping up care based on symptom severity. It relies on recommendations given by the eMental health application and the user's initiative to seek support outside the online service.

A pre-requisite for delivering high quality stepped care in eMental health are well-developed risk detection systems that are tailored to the signs and symptoms of the user population and the specific mental health issue that the intervention is targeted at. Risk detection systems inform decision-making about advancing users of online mental health services to more intensive forms of care.

1.1.1 Monitoring mental health in online self-help services

Risk detection systems in eMental health usually combine a form of information gathering and a decision-making component, either performed by a human (e.g., Jagayat et al. (2024)) or automatically by the mental health application (Meurling et al., 2023; Tielman et al., 2019). Adjustments to digitally delivered care can be made at different moments, including screening procedures that determine the best treatment option before the user engages with an online service, and dynamic adjustments to the individual's care program based on their momentary needs. Dynamic risk detection systems monitor user behaviors and mental states continuously or in regular intervals using a variety of data sources, including self-report questionnaires, data about the user's interaction with the online service, and physiological measures such as monitoring sleep patterns and tracking the user's heart rate. The development of mental health monitoring systems is complex and, based on the research and experience accumulated in this dissertation, involves (1) the selection of reliable indicators for people's mental health and changes thereof, (2) a measurement protocol including decisions about the frequency of measurements and a choice of assessment tools, (3) the processing of collected monitoring data for a pre-defined purpose (e.g., presenting a trend summary to the user, informing decision-making regarding stepped care), (4) a decision-making component, human or computer-based, that is capable of taking action based on users' processed data, and (5) a strategy for communicating the results of the monitoring to users and/or care professionals.

This dissertation exemplifies and discusses each of the above five steps in the context of the development of a new online grief service that supports older mourners after the loss of their spouse.

1.2 The LEAVES service

The research conducted in the context of this dissertation contributed to the development of LEAVES, an online grief service to support older mourners to adjust their life after spousal bereavement, as part of the LEAVES project (optimizing the mentaL health and resiliencE of older Adults that have lost thEir spouSe via blended, online therapy) within the Active Assisted Living (AAL) programme (project number AAL-2019-6-168-CP) (van Velsen et al., 2020). As such, a brief overview of the LEAVES online grief service will help readers of this dissertation contextualize the described research in the development process of a real eMental health application.

The LEAVES online grief service consists of three main components: (1) an onboarding procedure that introduces the service to its users, (2) an evidence-based grief intervention called LIVIA (Brodbeck et al., 2019), and (3) a monitoring module that recommends help-seeking in case the mental health of its users deteriorates. An initial risk assessment during the onboarding and a bi-weekly self-report mental health check form the basis for decision-making about recommending users to proactively seek support outside the grief service (Brandl et al., 2023). The LIVIA program is a text-based online grief intervention targeted at older adults who have lost their spouse due to bereavement or divorce. It is based on two influential theoretical models of grief, the task model by Worden (Worden, 2018) and the Dual Process Model (DPM) of coping with bereavement by Schut and Stroebe (Schut, 1999; Stroebe & Schut, 2010). Its original target group were older adults who suffer from prolonged grief symptoms, or seek help for the emotional adaptation after the loss. LIVIA proved its effectiveness for reducing grief and depressive symptoms, loneliness, and embitterment for mourning older adults that were recruited from a general population sample (Brodbeck et al., 2019). LEAVES is a redesign of the LIVIA program that extends the original content with a virtual companion called “Sun" and the monitoring functionality.

LEAVES consists of ten content modules through which users work in their own pace. Each content module treats a grief-related topic (e.g., myths and truths about grieving, reflecting on the days leading up to the death of the deceased spouse) and provides scientifically supported information about grieving, rewritten in non-academic language. Each module offers exercises that encourage the mourner to apply what is learned to their own grief process and to reflect on their coping strategies. Table 1.1 summarizes the topic of each content module briefly. The first two modules are mandatory to work through first. Afterwards, the user can choose the order of the remaining modules freely. The scientific grief content is complemented with activity suggestions that encourage the user to reflect on and actively shape their daily routine in ways that promote their mental health and physical well-being. The LEAVES companion “Sun" guides the user through the service and introduces and wraps up the work the user does within the content modules. Figure 1.1 provides an impression of the visual design of the LEAVES grief service.

At the time of writing this dissertation, the LEAVES project that lead to the development of the LEAVES grief service was concluded and received a positive review from the AAL programme. The evaluation of the LEAVES service (Brodbeck et al., 2022) has been completed at two out of three evaluation sites (the Netherlands, Portugal, and Switzerland). The obtained results have been shared within the project consortium and are in the process of being disseminated via scientific publications. So far, older mourners have been positive about the LEAVES service. On the level of the individual, the service improved grief and depressive symptoms and reduced mourners' feelings of loneliness (Hurmuz et al., 2023). In Switzerland, the evaluation is ongoing until the necessary sample size has been reached to empirically test two delivery formats of the service: (1) a self-tailored program where mourners choose the order in which they work through the ten content modules, and (2) a pre-determined program structure where mourners work through the program starting from the first to the last module.

In the context of the LEAVES project, the research described in this dissertation led to the development and evaluation of the LEAVES monitoring module (Brandl et al., 2023) with the aim of guiding users to offline support whenever their mental health warrants more intensive care than the self-guided online service can provide.

Table 1.1: Overview of the 10 grief topic modules in the LEAVES online service.

LEAVES module	Topic summary
1. Grief	Information about grief reactions, common misconceptions about grief, complicated grief and its treatment.
2. Where am I today	Information and reflection about emotions in the context of the interpersonal loss, changes in life and obstacles for adjustment after the loss.
3. Fostering positive thoughts and emotions	Information and reflection about regulating one's thoughts and emotions. Cognitive-behavioural strategies to promote positive thoughts and emotions, including directing one's attention and cognitive reappraisal.
4. Finding comfort	Suggestions for and reflection about self-soothing strategies and exercises to promote positive feelings amidst the negative feelings caused by the loss (e.g., going out for a walk and taking in the experience with all senses).
5. Self-care	Suggestions for taking care of one's physical (e.g., regular nutritious meals, keeping clothing clean), emotional and intellectual well-being (e.g., spending time in nature) and one's living and work spaces (e.g., car or bike maintenance) to create an environment for healing after loss.
6. Accepting memories and pain	Guided writing tasks to integrate painful memories about the loss into the autobiographical memory by practicing telling the story of the loss.
7. Unfinished business	Reflection about unfinished business and regrets, guided writing exercises to identify regrets and to make them explicit.
8. Shaping the new life situation	Identifying changes in daily life since the loss and sources of support and strength by comparing daily activities before and after the loss. Reflecting on one's personal growth since the loss.
9. Social relationships	Clarifying current relationships using a sociogram, reflecting on how satisfying each relationship is and how the loss has impacted the relationship. Defining goals and strategies, e.g. for intensifying specific relationships, or building new relationships.
10. Farewell	Guided formulation of a farewell letter to the lost partner to facilitate redefining the the shared bond.

1.3 This dissertation

The structure of this dissertation follows the five steps to developing an eMental health monitoring system that we outlined above. In the next chapter, chapter 2, we describe how we arrived at an initial set of mental health indicators as decision basis for determining whether people should seek more intensive support to prevent the exacerbation of grief symptoms. The development of any mental health monitoring system is a multi-disciplinary endeavor (Pagliari, 2007) which is why we draw on the knowledge of an expert-panel with varying (non-)academic backgrounds including grief professionals, psychotherapists, and experts in eHealth.

In chapter 3, we will introduce fuzzy cognitive maps (FCMs), a method that produces models that can be used as automatic decision-making components for eMental health monitoring systems. FCMs combine aspects of cognitive maps for decision-making (Axelrod, 2015) and fuzzy logic (Zadeh et al., 1996), a form of logic that allows different levels of truth rather than endorsing a dichotomous perspective of truth as something that is either *true* or *false*. Next, we apply the FCM methodology to the development of an expert-based decision model for the LEAVES monitoring module. Chapter 4 summarizes the research that led to the construction of the measurement protocol, self-report measurement tools, the FCM decision model, and the embedding of the monitoring module in the LEAVES grief service, including how help-seeking recommendations are communicated to the user.

Chapters 5 and 6 go beyond the initial technical development of the LEAVES monitoring system and are concerned with evaluating monitoring systems in eMental health, including the LEAVES monitoring module, in meaningful ways. In both chapters, stakeholder perceptions and actual use of the monitoring are evaluated and provide valuable insights into people's rationale for (not) using the monitoring system and how it can support mourners in transitioning from self-help to more intensive forms of care. In chapter 6, we investigate how well the LEAVES monitoring identifies whether it is advisable to seek more intensive care. We unravel and discuss challenges for evaluating monitoring systems in meaningful ways that relate to the state-of-the art of model evaluation in clinical practice and the priority with which monitoring systems are developed in eMental health. Ultimately, this dissertation a) exemplifies the development of an eMental health monitoring system with the aim of guiding people from low-intensity online self-help to more intensive forms of care and b) critically appraises opportunities and challenges at each step of the development process, particularly regarding methodologies that are available and commonly used to develop eMental health monitoring systems. As such, this dissertation contributes to the rapidly emerging field of monitoring mental health in the context of online self-help services in conjunction with decision-making regarding stepping up care intensity.

Figure 1.1: Screenshots of the LEAVES online grief service.

1

Determining mental health monitoring parameters to guide users of online self-help services to offline support

Based on:
Brandl, L., Cabrita, M., Brodbeck, J., Heylen, D., & van Velsen, L. (2022). Consulting the Oracle: A Delphi study for determining parameters for a mental health user profile and a personalization strategy for an online service to aid grieving older adults. *Internet Interventions*;28, doi:10.1016/j.invent.2022.100534.

Prologue

“After the ‘grief work’ has been completed, there is a complete recovery of the psychological condition", Orphea reads from the screen. She is familiar with the idea that mourning requires some kind of ‘work’ on the part of the mourner. To her surprise, the e-learning rejects the idea of ‘grief work’ and other common ‘myths’ about grieving. It explains that the term ‘healing’ suggests that there is some definitive end to the grief process (there is not). Instead, the grief program advocates that it is completely normal and healthy that even after years, loss can be painful. Orphea sits back and crosses her arms as she mulls the idea over. “Then what exactly is ‘good’ grieving? When am I doing it ‘right’?", she asks exasperatedly as she continues reading about how the grief process develops over time, according to science.

Abstract

While much effort has been devoted to the development of mental eHealth interventions, the tailoring of these applications to user characteristics and needs is a comparatively novel field of research. The premise of personalizing mental eHealth interventions is that personalization increases user motivation and (thereby) mitigates intervention dropout and enhances clinical effectiveness. In this study, we selected user profile parameters for personalizing a mental eHealth intervention for older adults who lost their spouse. We conducted a three-round Delphi study involving an international and interdisciplinary expert panel (N=16) with two objectives. The first aim was to elicit adaptation strategies that can be used to dynamically readjust the intervention to the user's needs. The second aim was to identify a set of meaningful indicators for monitoring the user from within the grief intervention to escalate from self-help to blended care, whenever advisable. This Delphi study used as starting point an evaluated, text-based grief intervention composed of ten modules, including psychoeducation about grief and cognitive-behavioral exercises to support the user in adjusting their lives after bereavement. Every user follows this grief intervention in a linear fashion from beginning to end. The resulting conceptual adaptation model encompasses *dynamic* adjustments, as well as one-time adjustments performed at the *initialization* of the service. On the level of the application structure, the adaptations affect *when which* topic module is presented to the user. The adaptations further provide strategies for adjusting the text-based content of individual intervention modules dependent on user characteristics and for selecting appropriate reactions to user input. Eighteen monitoring parameters were elicited and grouped into four categories: *clinical*, *behavioral/emotional*, *interactive*, and *external*. Parameters that were perceived as most urgent to attend to for escalation were *Suicidality*, *Self-destructive behavior*, *Client-initiated escalation*, *Unresponsiveness* and *(Complicated) Grief symptoms*.

2

2.1 Introduction

The loss of a spouse is a frequent occurrence in later life. While most bereaved adults successfully process the loss and continue to lead a normal life, some (about 9% of the bereaved population according to a recent prevalence study (Wilson et al., 2020)) have difficulties overcoming bereavement and develop complicated grief. Complicated grief in adults is a condition where severe grief symptoms occur longer than six months after bereavement and frequently results in a multitude of mental and physical problems, such as depression, loneliness, cardiovascular problems and, in extreme cases, suicidal tendencies (Molina et al., 2019). Internet-based (mental eHealth) interventions have been shown to be effective in treating mental illnesses, including complicated grief (Brodbeck et al., 2017; Carlbring et al., 2018; Eisma et al., 2015; Wagner et al., 2006). Benefits of eHealth interventions compared to face-to-face therapy are low threshold accessibility, flexible usage at a self-determined pace, and lower costs (Schröder et al., 2016). Internet-based interventions often combine a web-based self-help program and minimal, but regular therapist contact. The inclusion of regular professional guidance has been shown to improve adherence and clinical outcomes compared to standalone self-help programs (Baumeister et al., 2014), but less is known about how this support needs to be delivered. Support on-demand, where contact is initiated by the client and focused on the specific needs they have at that moment, has been suggested to optimize the incorporation of therapists in blended internet interventions (Dahlin et al., 2020; Oromendia et al., 2016). Here, we consider a complementary strategy for support on-demand: automated monitoring of the user's symptoms and situation while they use the intervention with the purpose of timely escalation if they end up needing more intensive professional support. This escalation could suggest to schedule a telephone or face-to-face meeting with a professional, if advisable. For this, a user profile is needed consisting of relevant indicators of the user's ability to continue working on the intervention by themselves, without professional intervention. These indicators should be optimized for the specific type of mental eHealth intervention. Finally, a decision algorithm that combines these indicators into actionable advice needs to be developed.

A second consideration about mental eHealth interventions is that a client's journey through mental illness is inherently personal. Indeed, therapists personalize face-to-face therapy readily by skipping or modifying therapeutic protocols to adapt to the client's needs and preferences and to increase the client's adherence to therapy (van Dooren et al., 2020). Lack of adherence, i.e., intervention dropout, has been recognized as a core challenge for eHealth interventions and personalization is a primary strategy for mitigating dropout (Burley et al., 2020; Eysenbach, 2005). Co-design of personalization strategies together with (clinical) professionals and end-users is essential, inherently multidisciplinary and challenging (Pagliari, 2007; van Dooren et al., 2020). Different expertises (e.g., therapists, designers, software developers) and disciplines (e.g., cognitive-behavioral, psychoanalytical therapists) come together, sometimes in multicultural settings accompanied with language barriers. In this paper, we demonstrate the administration of the Delphi method (Okoli & Pawlowski, 2004; Schmidt, 1997) to early-stage personalization research for mental eHealth interventions, involving an multidisciplinary expert panel working in four European countries.

This chapter has two objectives. First, we aim to find a personalization strategy for an internet-based grief intervention for older adults who lost their spouse. Second, we strive to determine a set of indicators that can be used to monitor the user for the purpose of delivering professional support on-demand. In particular, we will answer the following research questions:

2

> **RQ1**: What is a suitable personalization strategy for tailoring an online grief intervention for older adults who lost their spouse to the characteristics and needs of the individual user?

> **RQ2**: Which user parameters should be included in a user profile of an online grief intervention for older adults with the purpose of delivering professional support on-demand if they end up needing more intensive support?

We use the Delphi method (Okoli & Pawlowski, 2004; Schmidt, 1997) to consult an expert panel of clinical professionals regarding how they personalize their therapy and combine their knowledge with the knowledge of eHealth experts to yield actionable ideas for adaptations and indicators for monitoring. The Delphi method is commonly employed to reach agreement when literature is inconclusive or incomplete (Schmidt, 1997; Straat et al., 2020) and it is suitable for administration in multidisciplinary, multilanguage expert panels as the researcher can facilitate the group communication process.

2.2 Background

2.2.1 Personalized (mental) eHealth

In recent years, the personalization of mental eHealth has received much interest. A variety of personalized mental eHealth interventions have been developed and evaluated, including web-based interventions for treating anxiety (Carlbring et al., 2011) and depression (Johansson et al., 2012) and mobile interventions, such as Woebot, targeted at young adults with symptoms of depression and anxiety (Fitzpatrick et al., 2017). Nevertheless, there is no unified definition of personalized eHealth and the term *personalization* is often used interchangeably (e.g., Lustria et al. (2013) and Ryan et al. (2019)). Noar (Noar et al., 2011) explains the *tailoring process* as follows: characteristics of the person are gathered, either by another person or self-administered, and represent the *input* for the tailoring process. The input is processed either by a human or a computer that uses an algorithm to select content from an expert-developed database, such as texts, images, recommendations, and intervention messages. This is called the *tailoring process*. Finally, the tailored material (*output*) is optimized for the delivery mode at hand and presented to the individual. According to Noar, the premise behind personalization is that it increases the relevance of the intervention for the individual, who is subsequently more likely to cognitively process and to adhere to health advice.

Personalization strategies that have been employed to increase the relevance of mental eHealth interventions vary greatly in complexity, ranging from inserting the individual's name in the intervention, to adapting content based on user characteristics (Morrison, 2015), including clinical characteristics. Berger, Boettcher and Casper (Berger et al., 2014) use cut-off scores from disorder-specific self-report questionnaires to tailor the selection of treatment content in an intervention targeted at several anxiety disorders. Carlbring et al. (2018) compose individually-tailored anxiety intervention programs targeted at co-morbidities on the basis of structured clinical interviews for DSM disorders (SCID) prior to the start of the intervention. Other personalization strategies include user preferences and demographics (e.g., relationship status, status of employment) to optimize the relevance of scenarios with which the user practices healthy thinking patterns (Burley et al., 2020).

In conclusion, personalized mental eHealth is a vast research discipline, with no clear guidelines regarding the development of personalization strategies. Nevertheless, it appears that effective personalized interventions combine a) an evidence-based therapeutic foundation, b) careful selection of user characteristics to base the tailoring process on, and c) an expert-informed choice of therapeutic content that is suitable for personalization (van Dooren et al., 2020).

2.2.2 Case study of an online grief intervention

Our efforts to unravel parameters for automated monitoring and a personalization strategy are evidence-based. They draw on a text-based online intervention for older adults who have lost their spouse due to bereavement or divorce, called LIVIA (Brodbeck et al., 2019). The therapeutic content of the intervention was designed by professionals based on theoretical models of grief and components of cognitive-behavioral therapy for treating complicated grief (Brodbeck et al., 2017). The intervention is divided in ten modules that the user works through in their own pace. Modules consist of psychoeducation about the grief trajectory and cognitive-behavioral writing exercises to support the user in adjusting their lives after bereavement. For instance, the first module -*Psychoeducation* - consists of information about grief reactions, emotional reactions after bereavement and the (clinical) treatment of grief. The second module -*Assessment of the current situation* - reflects on the user's emotional reactions after the loss, changes in life since the loss and obstacles for positive adaptation. The intervention further encompasses topics related to self-care, the identification of changes in the daily routine since the loss and unresolved issues in the relationship to the lost spouse. It concludes with writing a farewell letter. The results of this research will inform a re-design of the intervention as part of the AAL LEAVES project (van Velsen et al., 2020). The re-design will enable the resulting intervention, called *LEAVES* in the remainder of this paper, to a) monitor a users' situation, with the goal to deliver support on-demand in the form of telephone calls or face-to-face meetings, whenever advisable and b) to tailor the therapeutic content based on user characteristics and needs to enhance user adherence and thereby clinical effectiveness. Finally, as part of the re-design and to make the intervention more interactive, the text-based content will be re-written into a dialogue format and delivered by an embodied conversational agent (ECA).

2.3 Method

2.3.1 Study design

A Delphi study is a systematic polling of the opinions of an expert panel, knowledgeable on a specialist topic through an iterative survey, usually in an attempt to reach group consensus on a given topic (Schmidt, 1997). A three-phase Delphi study involving 16 experts was conducted between June and December 2020. Figure 2.1 shows the process of the Delphi study. The design followed the three-phase design of Okoli and Pawlowski (Okoli & Pawlowski, 2004) with two notable deviations. First, we distributed three instead of four questionnaires in total. Okoli and Pawlowski administered two questionnaires in the first, the *brainstorming phase* and used the second phase of their study to narrow down the number of items they extracted from the first phase. We devoted the first questionnaire to brainstorming and the second to validating the items we elicited in the first questionnaire, thereby leaving out the step of narrowing down the number of items, for two reasons. First, the multidisciplinary nature of the conceptual adaptation model required our experts to transfer their discipline knowledge. We considered a validation of the model essential for establishing common ground for the final phase of the Delphi study, the *ranking*. Second, a common reason for narrowing down the amount of items is that it tremendously decreases the cognitive load of ranking tasks (Alwin & Krosnick, 1985). In our case, there was no need to narrow down the number of items for ranking because we designed the ranking in such a way that its cognitive load was minimized.

In the third round - the *ranking* - we did not strive for consensus, which constitutes a second deviation from Okoli and Pawlowski. Reaching consensus is often regarded as a necessary criterion for finalizing data collection in a Delphi study and this has been criticized before for its inappropriateness and artificiality of results (Dajani et al., 1979; von der Gracht, 2012). We asked our expert panel to rank the extracted adaptation strategies and user parameters for the monitoring in terms of maximal effect for the intervention (adaptation strategies) or clinical relevance (user parameters).

2.3.2 Participants

Experts were recruited via professional networks and by searching on the web. Search criteria were a) the expertise of the experts (grief, eHealth or both) and b) the country where they work. Since the results of this research will inform the design of the LEAVES grief intervention that is targeted at Dutch, Portuguese and Swiss older mourners, we recruited experts originating from these three countries. The recruited experts worked as researchers in the field of grief or eHealth or they practiced grief therapy, or both. Prior to the start of the study, all participants provided written informed consent to partake in the study and to reveal their identity to the rest of the expert panel after data collection had been finalized. Following (Miaskiewicz & Kozar, 2011; Okoli & Pawlowski, 2004; West, 2011), we collected data electronically, via e-mail, so as to ensure timely data collection and analysis. This was of importance to allow for conducting the different rounds in the study within a timeframe that was acceptable to the participants. Indeed, panel fatigue has been recognized as a challenge in Delphi studies (Hasson et al., 2000; Schmidt, 1997), as has the generation of large amounts of data that the research team has to process between questionnaire rounds (Green et al., 1999; West, 2011). Implementing a questionnaire with open-ended questions and recommending that each panel member provides a specific number of items, as suggested by Schmidt (1997) and West (2011), allowed our panel members to generate ideas while allowing the research team to process their input within a reasonable timeframe. Table 2.1 provides an overview of the recruited expert panel. Between phase one and three of the study, we experienced five dropouts. As a result, the second questionnaire was administered among twelve experts and the final questionnaire among eleven experts. Reasons for discontinuing their participation in the study were a) lack of time ($n = 1$) b) personal circumstances ($n = 2$) and c) lack of confidence in their responses ($n = 2$).

Table 2.1: Descriptives of the recruited expert panel.

Participant number	**Gender** (Female = 10, Male = 6)	**Country** (GER = 2, NL = 7, CH = 5, PT = 4)	**Experience**	**Expertise**
1	Female	Germany	Research	Grief
2	Female	Germany	Research/Clinical	Grief
3	Female	The Netherlands	Research	Grief
4	Male	The Netherlands	Research	eHealth
5	Male	The Netherlands	Research	eHealth
6	Female	The Netherlands	Research/Clinical	Grief
7	Male	The Netherlands	Research	Grief
8	Female	The Netherlands	Research/Clinical	Grief
9	Female	The Netherlands	Research	eHealth
10	Male	Switzerland	Research	Grief
11	Female	Switzerland	Research/Clinical	Grief/eHealth
12	Male	Switzerland	Research	Grief/eHealth
13	Female	Portugal	Clinical	Grief
14	Male	Portugal	Clinical	Grief
15	Female	Portugal	Research/Clinical	Grief
16	Female	Portugal	Clinical	Grief

2.4 Delphi Round 1

2.4.1 Method

The first questionnaire served the purpose of brainstorming user parameters that are important for making well-founded decisions regarding monitoring users from within the grief intervention and about adaptations that can be performed to make the LEAVES intervention more personal. Experts were asked to list five characteristics of (older adult) mourners that they consider important to attend to for monitoring based on their clinical and/or research expertise. A second question asked the experts to consider how they adapt grief therapy to their own clients or how they would adapt therapy in practice based on their discipline knowledge to meet the individual needs of clients. Since the second question was quite abstract, an illustrative example was given for a plausible adaptation of the LEAVES service. The example was based on the influential Dual Process Model of Bereavement (DPM) (Schut, 1999; Stroebe & Schut, 2010), which considers oscillating between focusing on the loss and on restoration essential for the recovery of the bereaved person. We suggested that one way to adapt the service to the needs of the user was to monitor their orientation and to nudge the user to either orientation if they appear to focus too much on the other, that is, to facilitate the *oscillation* process that is central to the DPM. In the same spirit, participants were asked to list five suggestions for adaptations that could be used to tailor the intervention. The first questionnaire included a description of the LIVIA grief intervention, including an overview of the topics that the intervention treats in ten modules.

A thematic analysis (Braun & Clarke, 2006) was conducted on the qualitative *brainstorming* responses of the first questionnaire. A coding scheme was constructed in a bottom-up fashion and employed to the data. Two researchers independently coded the data using this coding scheme and a third researcher was involved to resolve conflicts.

For the adaptations, the initial coding scheme included codes for strategies that can be employed to tailor an online intervention. Additional codes covered suggestions regarding user parameters for personalization and theoretical and therapeutic frameworks to consider. For the monitoring parameters, each suggested parameter was coded. During the construction of the coding scheme, codes for overarching monitoring categories emerged and were included in the coding. Based on the coding, the adaptation model was constructed and user parameters were extracted and grouped into the overarching parameter categories.

2.4.2 Results

The two main outcomes of the first round of the Delphi study are the conceptual adaptation model depicted in Figure 2.2 and a set of user parameters for monitoring purposes.

Adaptations

The adaptation model consists of four types of adaptation strategies that can be employed to tailor LEAVES to the needs and characteristics of the user during the *initialization* of the service and *dynamically* while the user engages with the platform. The four types of adaptations are *Topic Selection*, *Program Structure*, *Content Version*, and *Coaching Style*. In the following, each adaptation type is presented in more detail, alongside some exemplary suggestions of the panelists that contributed to their conceptualization.

Topic Selection
On the level of intervention modules, *Topic Selection* concerns *what* content is presented to the client. In contrast to the linear structure of the grief intervention LIVIA, a LEAVES user is presented with an individualized program based on an initial assessment. Subsequently, their individualized program is adjusted based on regular assessments as they progress through the intervention. Within *Topic Selection*, two adaptation strategies were proposed. First, the *removal of intervention modules* from the default configuration. Second, the *dynamic adjustment* of the selection of intervention modules based on regular

re-assessment of the user's situation and progress. The latter strategy does include adding earlier removed intervention content to the user's personalized intervention program.

> "Many online interventions are 'hybrid' meaning, that some modules are mandatory but others are optional dependent on the needs of the participant. For the LIVIA intervention, I expect that psycho-education and assessment of the current situation are mandatory modules that every participant should follow, but this may not be the case for all modules (e.g., not every participant may have problems with self-care or personal relationships)." (Participant 9)

Program Structure
On the level of intervention modules, *Program Structure* affects the structure of the client's personalized selection of therapeutic content. It determines *when* content is presented to the client, given a selection of intervention modules determined by the *Topic Selection* adaptation. Structural adaptations occur in the beginning of the LEAVES program as well as dynamically when the client is already using the program. Specific strategies that were proposed were *adjusting the order* of modules, *adjusting the length* of the intervention by manipulating the time spent on a topic and finally, *adjusting the length* by manipulating the number of modules or exercises on a topic.

> "A possibility would be to suggest a priority list, and make participants begin the program by the factor that has been identified as the most salient or serious for each participant." (Participant 11)

Content Version
On the level of individual intervention modules, the *Content Version* adaptation impacts *how* the content is presented to the user. In particular, two strategies for adjusting the content to the characteristics emerged from the suggestions. First, when the client appears to get stuck with the content of a module or an exercise, or they indicate that they want to try a different approach, the same content is presented from a different angle. An alternative approach to regular psycho-education regarding a coping strategy could, for example, be a reflective exercise or a more playful "thought experiment" where the client tries out the benefits and consequences of the strategy first-hand in a type of adventure game. Another approach would be to identify concrete tasks and obstacles since the loss and to consider which resources in their daily life can be activated to address them:

> "With regard to 'unfinished business' and 'creating a new life', the client may be guided to explicitly identify tasks that the deceased partner had performed and that now remain undone (e.g., managing

> the family finances, providing emotional support). In the next step, the client should specify how and by whom these tasks can be done now or the needs can be met now." (Participant 2)

The second *Content Version* strategy is about re-writing the content of an intervention module depending on characteristics of the client. For example, the module about personal relationships could be better tailored by preparing a version for introvert versus more extrovert users.

Coaching Style
On the level of the conversations between the client and the embodied conversational agent (ECA) of LEAVES, *Coaching Style* concerns how the LEAVES program reacts to the input that the client provides. The adaptation encompasses strategies such as *acknowledging* hardship and reacting with appropriate empathy, including *personalized feedback* messages regarding regular monitoring assessments. Another strategy is keeping track of the story the client discloses to the system and to *point out changes or incoherencies* in the client's perceptions. A final strategy focuses on *interweaving* the input the user provides regarding different topics. For example, the ECA could build a conversation around what the user discloses regarding their hobbies and suggestions about how they can widen their social network. Regarding reflecting on changes in the discourse of the client, participant 12 suggested:

> "Memories are also changing over the course of time and grief. It might be important to 'store' such memories and ask from time to time whether these memories still have the same quality. I could imagine a program that 'reflects' on personal memories in exchange with adaptations from the client." (Participant 12)

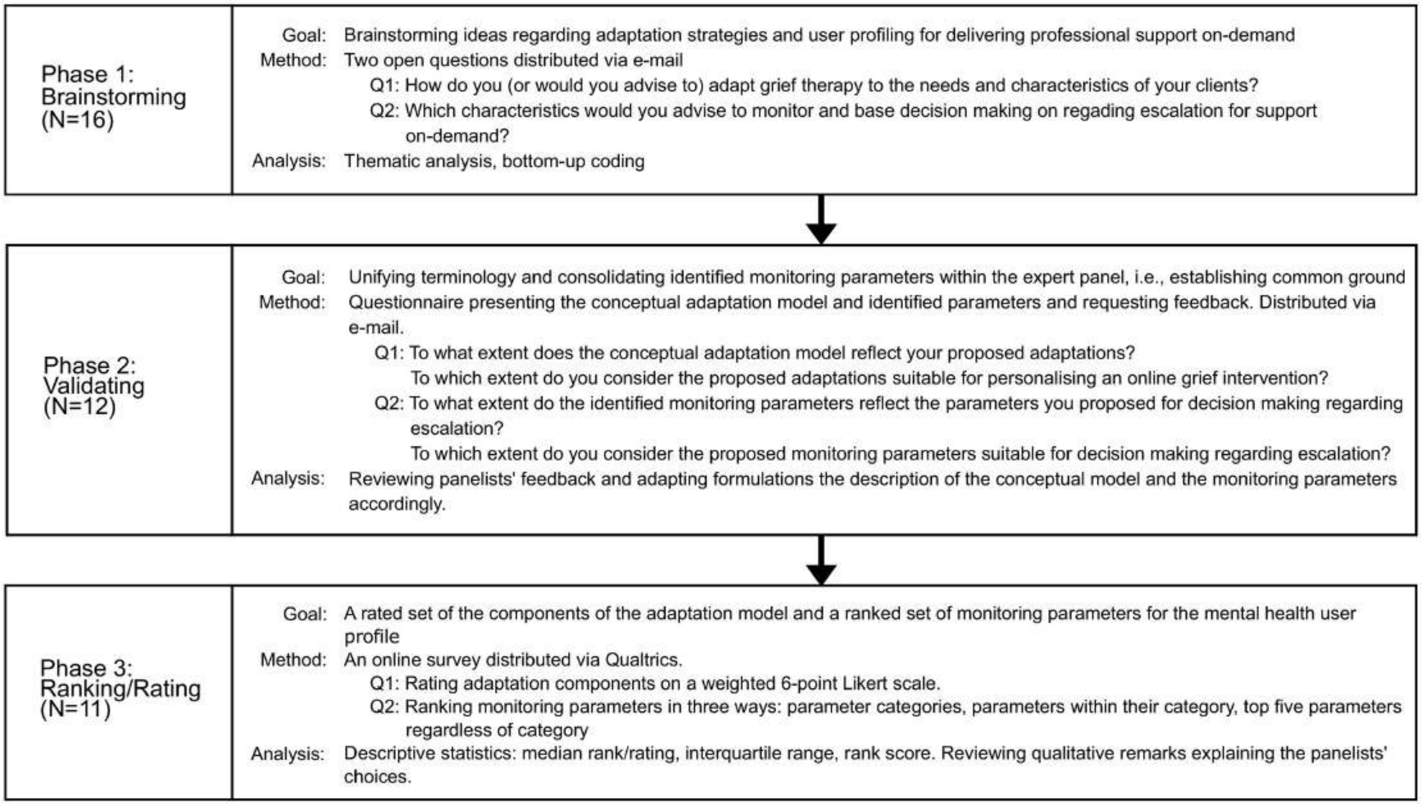


Figure 2.1: Overview of the three-phase Delphi design.

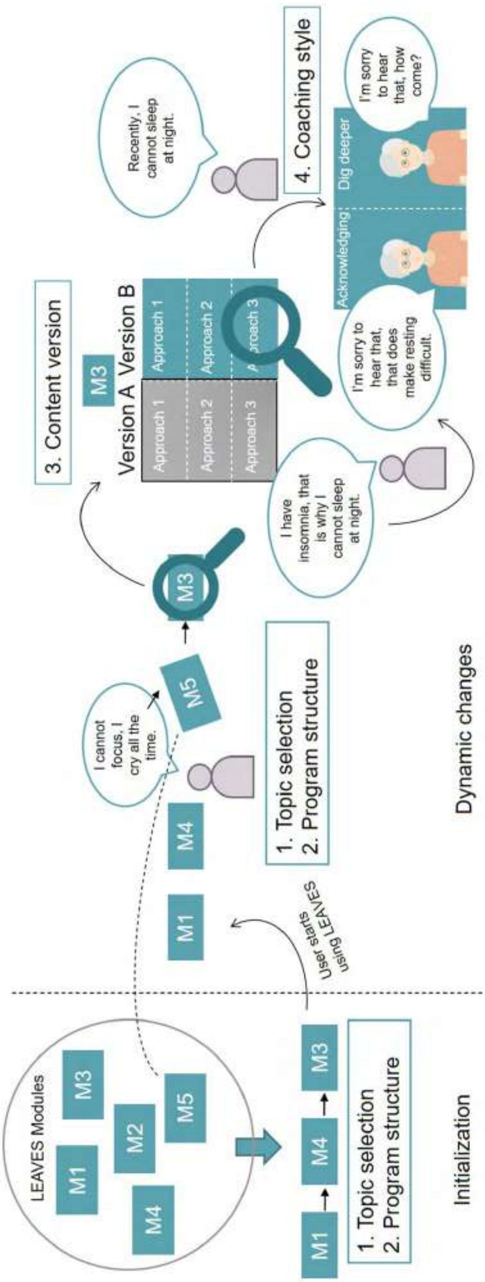


Figure 2.2: Visualization of the proposed adaptation model for personalising the LEAVES intervention.

Monitoring parameters

On the basis of the brainstorming phase, we extracted 18 monitoring parameters for the construction of a user profile that informs decision making about professional support on-demand in LEAVES. The parameters were subdivided into four categories: *clinical*, *behavioral/emotional*, *interaction*, and *external*. Clinical parameters included symptoms that frequently occur in mourners, such as depressive and complicated grief symptoms. The behavioral/emotional category summarizes relevant user behaviors, such as the extent to which they are functionally autonomous, and relevant user characteristics such as the extent to which they are able to look ahead in the future positively. Both represent frequent risk factors for deterioration in (older) adult mourners. The interaction category includes parameters that describe the interaction of the user with the LEAVES service. Finally, the external parameter category encompasses two parameters that involve either events or people in the physical world outside the LEAVES intervention. Table 2.4 shows the 18 parameters that were extracted from the brainstorming phase. Variations in how the experts formulated their suggested monitoring parameters were treated in questionnaire two when the experts were able to give feedback on the completeness and appropriateness of the extracted parameters. For instance, the following three expert suggestions contributed to the definition of the *social isolation* monitoring parameter:

> “If a client shows an increase in social avoidance or a reduction or complete lack of social contacts, blended treatment should be favoured." (Participant 10)

> “Ability to focus and derive comfort from other relationships (e.g., grandchildren)." (Participant 15)

> “It could also be good to assess isolation or withdrawal behaviors, that is if the participant does not follow the program anymore because they are withdrawing from any activity or contact." (Participant 11)

Concluding remarks

In the first round of the Delphi study, we elicited ideas for adapting the LEAVES intervention to the needs and characteristics of the user and parameters that can form the user profile for providing support on-demand. The first round yielded a conceptual adaptation model for LEAVES and an initial set of 18 monitoring parameters.

2.5 Delphi Round 2

2.5.1 Method

The aim of the second questionnaire was to validate the conceptual adaptation model and to confirm that we correctly understood the suggested user parameters for monitoring. For this, we presented the adaptation model to our expert panel and asked them to comment on the extent to which it represented their suggested adaptations. We also encouraged them to pose any questions or make any remarks they may have about the proposed model. Specifically, we provided a list of the adaptation types and strategies with their respective definitions. For each adaptation strategy, we included an example. We also provided a holistic visualization of the adaptation model that summarized the adaptations and how they interact with each other to tailor the LEAVES intervention.

Regarding the monitoring parameters, we presented the 18 parameters that we extracted from the first questionnaire alongside their definitions. The parameters were subdivided into their four overarching categories: clinical, behavioral/emotional, interactive and external. As for the adaptations, we asked our experts to review the parameters and to state how the list represented their input. We also encouraged them to add any parameters that we may have overseen. In terms of analysis, the research team reviewed the questions and remarks the panelists submitted as response to the second questionnaire.

2.5.2 Results

Overall, the conceptual adaptation model was received positively and the presented list of monitoring parameters turned out to be quite exhaustive. There were some confusions regarding components of the adaptation model. For instance, one participant believed that the *Content Version* adaptation type impacted the modality of the module and not the therapeutic content itself. They thought that *Content Version* was about whether audio or video were included in the user interface design. Minor changes to the description of the conceptual adaptation model and the monitoring parameters were performed as a result of the input we received in the second round. In addition, personalized clarifications and answers to questions posed by panlists were provided to achieve a solid common understanding of the adaptation model and the monitoring parameters.

Concluding remarks

In the second round of the Delphi, we established common ground regarding terminology and a common understanding of the components of the adaptation model and the monitoring parameters. This round represented a necessary intermediate step towards a well-informed and reliable weighting of the components of the adaptation model and the monitoring parameters.

2.6 Delphi Round 3

2.6.1 Method

The third questionnaire aimed to yield a weighting of the adaptations and parameters for monitoring. Given the abstract nature of the adaptation model defined in terms of personalization concepts, we chose to ask our mostly clinical expert panel for a *rating* of the adaptation strategies according to their perceived contribution to clinical outcome and to *rank* the overarching adaptation types. All adaptation strategies were rated on a 6-point Likert scale ranging from 1 (Not beneficial, even risky) to 6 (Extremely beneficial).

For the monitoring parameters, we asked the expert panel to rank the 18 parameters in three ways: By ranking the four parameter categories, by ranking all parameters within each parameter category, and finally, by selecting a top five most important parameters for decision making regarding escalation, regardless of their category. For all rankings, the order in which the monitoring parameters were presented was randomized for each panelist. We asked the panelist to briefly explain their rationale for the *rating* of the adaptations and the *ranking* of the parameters for monitoring.

The results of the third Delphi round were analyzed as follows. For both, the adaptations and the monitoring parameters, the median rating/rank was determined and the interquartile range (IQR) was calculated as a measure of dispersion. The IQR consists of the middle 50% of the observations and therefore, an IQR of less than 1 means that more than 50% of all rankings or ratings fall within 1 point on the scale. The IQR is frequently used in Delphi studies and it is generally accepted as an objective and rigorous method for assessing consensus (von der Gracht, 2012). For the top five selection, we calculated a ranking score. The score was a combination of the number of times a parameter was assigned a specific rank in the top five and a weight that was assigned to the rank, normalized by the size of the sample:

$$\frac{\sum_{i=1}^{5} x_i w_i}{N};\ w \in \{20, 16, 12, 8, 4\}$$

Where x_i is the count of how many times a parameter was ranked as rank i, w_i is the assigned weight to rank i and N is the sample size of the expert panel in the third round of the Delphi, $N = 11$. Adjusting the rank weights boiled down to balancing two values, the importance one assigns to being included in the top five selection versus the importance one assigns to a specific rank. We experimented with a number of weights and settled with $w \in \{20, 16, 12, 8, 4\}$ because a) we considered being chosen as one of the top five parameters out of 18 too important to let individual ranks dominate the ranking and b) the order of the parameters ranked by score stayed stable. For the ranking of the parameter categories and the ranking of the behavioral/emotional parameters, two participants refrained from participating in the ranking, one participant for each ranking. In both cases, the participant felt that they lacked the required expertise for this ranking task.

2.6.2 Results

Adaptations

Table 2.2 summarizes the ranking of the four adaptation types and the ratings for the eleven adaptation strategies. The following ranking order ranked highest to lowest can be extracted: Topic Selection, Program Structure, Content Version, and Coaching Style. Using the rule of thumb advocated by von der Gracht (2012) when using the inter-quartile range as a measure of consensus, the rankings for *Topic Selection* and *Program Structure* achieved reasonable consensus ($IQR \leq 1$), and there was less agreement on the other two adaptations ($IQR \geq 1$), *Content Version* and *Coaching Style*. For the adaptation strategies, scores were generally positive as all suggested strategies had a median rating of at least 3, *Somewhat beneficial*, five strategies had a median rating of 4, *Beneficial* and four a median rating of 5, *Very beneficial*. Regarding panelist agreement, three strategies were rated with good agreement ($IQR \leq 1$) (*Adjusting the order*, *Different versions*, and *Acknowledging*) and six adaptations with reasonable agreement ($IQR \leq 1.5$). The greatest divergence in expert opinion was obtained for *Adjusting the length (time)* of the intervention by manipulating the amount of time the user spends on a topic. Participant 12 is in favor of manipulating the time a user is advised to spend on a topic; Participant 8 is doubtful:

> “It is an excellent idea to play with time and shape the intervention accordingly. Especially, I like the idea to give the client more exercises when needed." (Participant 12)

> “The suggested interpretation of ‘clicking through the program’ here might be too small: it might be the client’s normal learning need to behave like this to answer a need to get an overview of what is available (understand the modules before determining a sequence). In such cases, a suggestion to slow down is not helpful to this client." (Participant 8)

Figure 2.3 shows how often the five highest rated adaptation strategies received which rating, regardless of their overarching adaptation category. The selection of the five most promising strategies was based on their median rating and IQR. Ratings were closely tied, but based on the counts, strategies *Dynamic adjustments* to the selected topics for a user's personalized program, dynamically manipulating the *Order* of intervention modules, and offering *Different versions* of the intervention content were rated highest.

Table 2.2: Median rank and inter-quartile range (IQR) for the four adaptation types. Median rating and IQR for the eleven adaptation strategies.

Adaptation		Median		
Type	**Strategy**	**Rank**	**Rating**	**IQR**
Topic Selection		2.0		1.0
	Dynamically changing topics		5.0	1.5
	Removing from default		4.0	1.5
Program Structure		3.0		1.0
	Adjusting the order		5.0	1.0
	Overall length (n modules)		5.0	1.5
	Overall length (time)		3.0	2.0
Content Version		3.0		1.5
	Different versions		5.0	1.0
	Repertoire of approaches		4.0	1.5
Coaching Style		3.0		2.5
	Acknowledging		4.0	1.0
	Personalized feedback messages		4.0	1.5
	Reacting to incoherencies		4.0	1.5
	Interweaving input		3.0	1.0

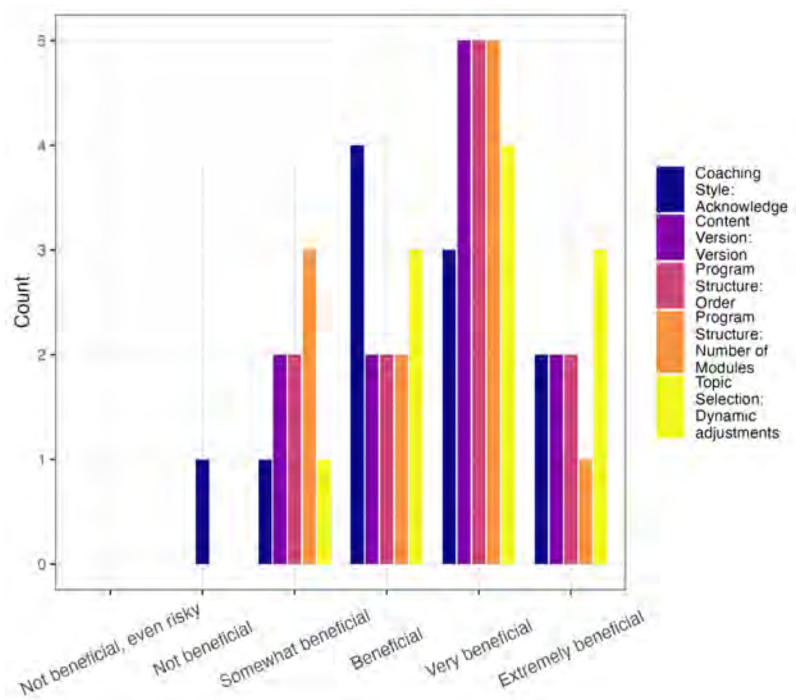


Figure 2.3: Bar chart of adaptation strategy ratings with comparable median and IQR ratings.

2

Monitoring Parameters

Table 2.3 summarizes the ranking for the four user parameter categories (*clinical*, *behavioral/ emotional*, *interaction*, and *external*). The intra-category ranking of individual user parameters and their calculated rank scores are depicted in Table 2.4. The ranking of the four monitoring parameter categories yielded a clear propensity towards the clinical parameters. Overall, consensus regarding parameter category ranks was good ($IQR \leq 1$). Regarding intra-category rankings, for the clinical parameter category, *Suicidality* was ranked unanimously as the most important parameter for escalation, followed by *(Complicated) Grief*, *Depressive* and *PTSD* symptoms. With the exception of the latter parameter, there was good agreement regarding the ranking of each clinical parameter ($IQR \leq 1$). The behavioral/emotional parameter category exhibited a less conclusive ranking. While having obtained the highest median rank, *Self-destructive behavior* exhibited the largest dispersion of assigned ranks. Based on the median rank and IQR, *Hopelessness* was ranked highest of all behavioral/emotional parameters for escalation, albeit with considerable disagreement ($IQR \geq 2$). Figure 2.4 shows how often each behavioral/emotional parameter was assigned which rank in the intra-category ranking. It appears that ranking *Self-destructive behavior* and *Functional Autonomy* had a polarizing effect. A subset of panelists ranked these two parameters high, while another subset ranked them low. *Hopelessness* was the only parameter with a trend towards higher ranks.

The ranking of the five interaction parameters was more conclusive, albeit with considerable disagreement regarding the rank of three out of five parameters ($IQR = 2$). There was good agreement regarding *Client-initiated escalation* and *Defence mechanisms* ($IQR = 1$). Regarding the final parameter category, external parameters, there was no clear preference for either external parameter. In sum, based on the three-fold ranking, the five highest ranked monitoring parameters are *Suicidality*, *Self-destructive behavior*, *Client-initiated escalation*, *Unresponsiveness*, and *(Complicated) Grief symptoms*, closely followed by the behavioral/emotional parameter *Hopelessness*. However, there is considerable disagreement regarding the importance of *Self-destructive behavior*, *Unresponsiveness* and *Hopelessness* when these parameters are ranked against other parameters within the same category.

Table 2.3: Median rank and inter-quartile range (IQR) for the four parameter categories. ***Median and IQR values based on the depicted number of experts.

Parameter category (N=10*)	**Median**	**IQR**
Clinical	1.0	0.5
Behavioral/Emotional	2.0	1.0
Interaction	3.0	0.5
External	4.0	1.0

Table 2.4: Median rank and inter-quartile range (IQR) for the intra-category ranking and calculated ranking scores for each parameter. ***Median and IQR values based on the depicted number of experts.

Parameter	**Median**	**IQR**	**Score**
Clinical parameters			
Suicidality	1.0	0.0	15.64
(Complicated) Grief symptoms	2.0	1.0	5.45
Depressive sypmtoms	3.0	1.0	2.91
PTSD symptoms	3.0	1.75	3.27
Behavioral/Emotional parameters (N=10*)			
Hopelessness	2.0	2.25	4.73
Self-destructive behavior	1.5	4.5	6.55
Social isolation	3.5	1.0	0.73
Affective state	3.5	1.75	0.36
Functional autonomy	5.0	2.0	1.45
Physiological	6.5	1.75	0
Relation to the deceased	5.5	3.75	0
Interaction parameters			
Client-initiated escalation	2.0	1.0	6.18
Unresponsiveness	2.0	2.0	5.82
Coherence of discourse	3.0	2.0	1.09
Defence mechanisms	4.0	1.0	0
Too many questions	4.0	2.0	0
External parameters			
Events inducing vulnerability	1.0	1.0	2.91
Peer assessment	2.0	1.0	2.91

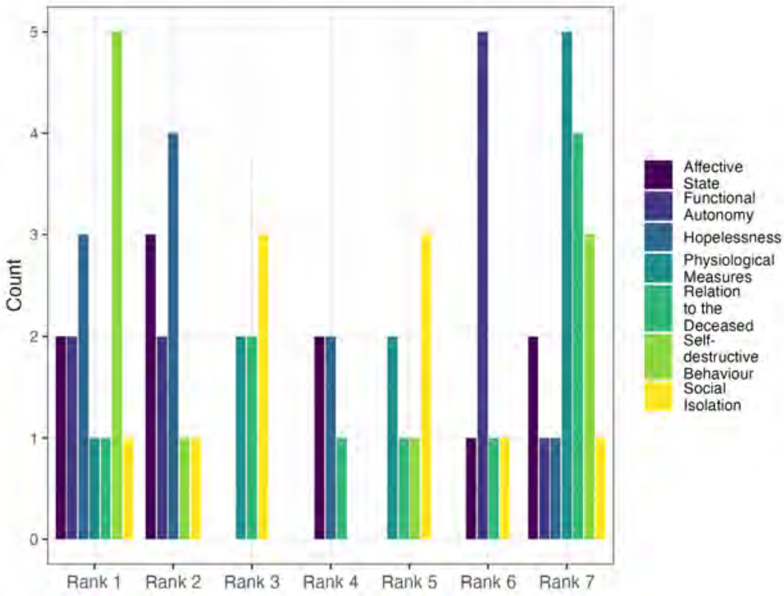


Figure 2.4: Bar chart of behavioral/emotional user parameter ranks. Counts are based on the responses of ten experts.

Concluding remarks

The third round of the Delphi, yielded a *rating* of the components of the proposed adaptation model according to their expected capacity to increase clinical effectiveness and a *ranking* of the monitoring parameters according to their importance for decision making regarding professional support on-demand. There was considerable disagreement among panelists which emphasizes the importance of qualitative accounts of the panelists' rationale in addition to ratings and rankings in the Delphi process.

2.7 Discussion

This chapter describes the process and the results of a three-round Delphi study involving 16 grief and eHealth experts to determine strategies for adapting an online grief intervention for older adults who lost their spouse and to identify parameters for a user profile for decision making regarding support on-demand. The Delphi study yielded a conceptual adaptation model whose components were rated by the expert panel according to their potential for increasing the clinical effectiveness of a text-based intervention. A preference emerged for dynamic *Topic Selection* and *Program Structure* adjustments as well as the possibility to offer tailored versions of the therapeutic text-based content. In contrast, adaptations that impact the *Coaching Style* of an embodied virtual agent (ECA) that guides the user through the intervention were received with scepticism regarding technical feasibility and health risks if not employed with utmost caution. A set of 18 monitoring parameters was elicited out of which *Suicidality*, *Self-destructive behavior*, *Client-initiated escalation*, *Unresponsiveness*, *(Complicated) Grief symptoms* and *Hopelessness* were ranked as most important to attend to for decision making regarding the intensity of professional support outside the online self-help service.

Two types of adaptations emerged from the suggestions of our expert panel: adaptations that impact the configuration of the service at initialization based on an initial assessment and dynamic adaptations that continue to re-adjust the intervention to the changing needs and preferences of the user. In a meta-analysis of tailored interventions for health behavior change, Krebs et al. (2010) found *dynamic tailoring* (assessing intervention parameters prior to each feedback provided to the user) to outperform *static tailoring* (basing all intervention feedback on one baseline assessment) regarding long-term intervention effect. While the effects for both, *static* and *dynamic* tailoring decrease over time, the authors found *dynamic tailoring* to still be statistically effective at twelve months after the intervention and attributed this to the enhanced relevance of feedback that reflects a person's change. This strengthens our panel's advice to focus on *dynamic* adjustments to enhance clinical effectiveness of the intervention. Regarding the user profile for delivering professional support on-demand, the panel's strong preference for clinical parameters can be attributed to the predominantly (clinical) grief expertise in the panel and the unanimous high ranking of *Suicidality* by it being life-threatening to the user and requiring immediate professional intervention, if present.

We were able to brainstorm ideas for adaptation and elicit monitoring parameters in a sample of 16 experts working in academia and clinical practice across four countries and to give a weighting to the elicited ideas to guide our future research efforts. The Delphi approach has been used at later stages of the development of personalized systems, for example, in rehabilitation to match recommendations regarding physical activity to the user's capabilities (Straat et al., 2020). However, the Delphi approach is rarely used in eHealth research according to a recent review of human-centred methods for eHealth development (Kip et al., 2022). The authors highlight two challenges for conducting Delphis in eHealth. First, the recruitment of experts can be challenging because experts may be scarce in a new or specific field and regarding topics that involve novel applications of eHealth technology. Second, reaching consensus can be time-consuming and complex. This paper shows how practical design choices can address these challenges when using the Delphi approach in early-stage personalization research for (mental) eHealth. Regarding the first challenge, we chose to recruit an interdisciplinary expert panel including experts in grief and eHealth instead of focusing on finding participants that possess expertise in both fields. Since we did not require a specific level of consensus as a stopping criterion and employed a fixed number of Delphi rounds instead, we limited the impact of consensus-seeking on the time effort required of the participants and the research team. The trade off between time and monetary investments and the extent to which consensus can be achieved in Delphis has been acknowledged (Hasson et al., 2000; von der Gracht, 2012).

Regarding the first strategy, recruiting an interdisciplinary expert panel, our subsequent choice to treat the experts as a single panel had implications for the level of detail of our results. We exposed grief experts to concepts from personalization and user profiling and subsequently asked them to rate these concepts. We also asked eHealth experts to consider specific characteristics of mourners. In both cases, the experts in this research had to transfer their own domain knowledge. To maintain the accessibility of the adaptation model to all panelists, it was formulated in generic personalization terms and requires a considerable specification for the development of the LEAVES grief intervention. This constitutes a limitation of the chosen strategy. However, it does increase the transferablility of the resulting adaptation model to other mental eHealth interventions for which a personalization strategy needs to be determined.

The latter strategy, refraining from using consensus in the rating and ranking tasks as stopping criterion and pre-determining the number of Delphi iterations instead, constitutes a deviation from Schmidt (1997) and Okoli and Pawlowski (2004) whose Delphi approach otherwise guided our study design. The common focus on reaching consensus in ranking-type Delphis has been criticized before for producing artificial consensus results and being an inadequate stopping criterion by itself (Dajani et al., 1979; von der Gracht, 2012). The current research endorses Gracht's statement about the primary goal of the Delphi approach: "the efficient structuring of a group communication process" (von der Gracht (2012), page 1527). Pre-determining the number of Delphi iterations implied that consensus would be difficult to achieve in the rating and ranking tasks. There was indeed considerable disagreement regarding the rating of some adaptations and user parameters for monitoring.

Consequently, a limitation of the obtained rating and ranking results is that they should not be treated as an objective order of importance based on expert consensus, but rather as a starting point for subsequent efforts to specify a personlization strategy and for constructing a user profile to deliver support on-demand. Any follow-up efforts should scrutinize the suitability of highly rated adaptations and highly ranked user parameters for their specific purpose. Decision criteria that are relevant to their specific eHealth application should be established. One criterion for determining the suitability of any monitoring parameter for any specific eHealth application should be its sensitivity to changes in the user's situation (Gokalp & Clarke, 2013). The parameter's sensitivity must be compatible with the application's monitoring measurement interval to timely detect a deterioration of the user's situation.

Researchers who consider taking a Delphi approach for early-stage personalization research should consider the implications of the practical choices presented in this paper to address common challenges of Delphi studies in eHealth. Specifically, if establishing consensus is desired in a Delphi approach with pre-determined iterations, the research team may consider replacing the brainstorming phase with pre-determined statements (e.g., Yap et al. (2017)). When the number of iterations is not pre-determined, a hierarchical stopping criterion combining measures of group response stability over the course of several rounds and a consensus measure such as the interquartile range should be considered, as advocated by Schmidt (1997) and Okoli and Pawlowski (2004).

In conclusion, this chapter set out to determine a personalization strategy for an online grief intervention targeted at older adults who lost their spouse and to identify parameters for a user profile for delivering blended professional support on-demand. A conceptual adaptation model was constructed. Based on the ratings of eleven grief and eHealth experts, *dynamic* adjustments, informed by regular assessment of user characteristics, to the selection of intervention topics and the order in which topics are presented emerged to be a promising personalization strategy. Indicators that capture perceived danger for the client and their ability to continue the intervention by themselves should be included in the user profile, including measures of *Suicidality*, *Self-destructive behavior* and *Client-initiated escalation*, *Unresponsiveness* and *(Complicated) Grief symptoms*. Based on our experiences from this research, the Delphi approach can be a useful early-stage research tool for exploring a personalization strategy for mental eHealth interventions and for unraveling user parameters for decision-making about providing professional support on-demand.

Fuzzy cognitive maps for algorithmic decision making in eMental health

Based on:
Brandl, L., Jansen-Kosterink, S., Hofs, D., & Heylen, D. Fuzzy cognitive maps for algorithmic decision making in eMental health. (*Submitted for publication*).

Prologue

Orphea feels restless. She planned to sit down and go through another module in the e-learning, box some of Justus' clothing and bring them up to the attic before visiting the bank to close her husband's account. Instead, she did not feel like breakfast, vacuum-cleaned the house, solved at least four crosswords, drank three instead of her usual only cup of morning coffee (which did not help with the restlessness), did some laundry and just as Orphea was considering cleaning the windows as well, the clock starts to chime 12PM in the living room. As the chiming fades, Orphea slowly sits down on the couch. It is the third time that she had planned to go to the bank. "This is not working", Orphea sighs. She picks up her phone and dials her daughter's number. Her daughter picks up after the second ring. "Honey, do you remember that computer program that Dr. Epione recommended?" Her daughter remembers. "I found it a little silly at the time, but it asked me to write down the names of people that I can reach out to. No, don't worry, I'm fine. Honey, I know that your busy, but can you meet me later at the bank? I have to close dad's account and I don't think I can."

Abstract

Background: eMental health is on the rise to meet growing global mental healthcare needs and so is the need for applications with automatic decision-making capabilities to support clinical practitioners and to personalize eMental health interventions to their users. However, due to unresolved challenges and clinicians' reluctance to adopt artificial intelligence in mental healthcare, the technology's current state in mental health is still in the phase of proof-of-concepts rather than providing decision making algorithms that are ready to create clinical impact. Fuzzy cognitive maps (FCMs), a soft computing technique that leverages expert knowledge by combining cognitive maps with fuzzy logic, can be an alternative approach to equipping eMental health with automatic decision-making capabilities.

Method: This chapter provides an introduction to FCMs for developing decision algorithms for eMental health. It demonstrates step by step how FCMs can be developed, based on an example from mental health, while discussing challenges when applying the methodology in eMental health. By situating the demonstrated FCM approach in the current research landscape of the methodology, we identify a research agenda for FCMs in the mental health domain.

Results: FCMs are a viable option at the disposal of developers and researchers in the field of eMental health that come with their own set of advantages and challenges: (1) they allow the construction of initial mental health decision models without depending on the availability of large amounts of historical client data, (2) they are explainable, and (3) they involve clinicians in the construction process, inherently promoting professionals' acceptance of the resulting algorithm. Regarding challenges, (4) there is a pronounced trade-off between the simplicity of FCMs and the accurate representation of mental health constructs since their approach offers limited ways for modelling beyond direct causality and for including temporal aspects of mental health problems, (5) FCM decision models are not usually on par with data-driven black-box AI models in terms of classification accuracy.

Conclusion: FCMs are an interesting and promising alternative modelling approach for developing decision-making algorithms in eMental health. Their inherent expert-basis, explainability and independence of large clinical data sets to arrive at initial usable decision models makes them especially attractive for the mental health domain. However, a recent surge in technical and methodological contributions to the FCM discipline can make it challenging for novel applied researchers to enter the field because FCM research is dispersed, progressing fast, and sometimes, contradictory.

3.1 Introduction

The realization that in-person professional mental healthcare is no longer and will never again in the future be able to meet the global need for mental health support has become known as the mental health treatment gap (Teachman et al., 2022). Digital mental (eMental) health services (DMHS) in general and those leveraging recent advances in artificial intelligence (AI) in particular (Graham et al., 2019; Higgins et al., 2023), hold much promise for increasing the accessibility and scalability of mental healthcare (Baños et al., 2022; Leichsenring et al., 2019; Teachman et al., 2022). However, AI-based DMHS introduce new challenges of their own, including the availability of high quality clinical data sets to train models, model transparency, and the smooth integration of AI tools into clinicians' workflows, which in turn affect clinicians' trust in AI technology and their willingness to use it in practice (Higgins et al., 2023). The most prominent purposes of AI tools in mental health are quality assessments of mental health treatments (Tornero-Costa et al., 2023) and the detection of symptoms and risks of mental illness, and subsequent mental disorder diagnosis (Thieme et al., 2023; Tornero-Costa et al., 2023). All in all, the current state of AI in mental health is still considered to be in the phase of proof-of-concepts and showcasing technical feasibility rather than providing systems that are ready to create clinical impact (Thieme et al., 2023).

One type of model that addresses some of the common challenges that prevent AI tools from moving from experimental proof-of-concepts to systems that are used in mental healthcare practice are fuzzy cognitive maps (FCMs). FCMs combine aspects of cognitive maps for decision making (Axelrod, 2015) and fuzzy logic (Zadeh et al., 1996) and are in healthcare most prominent in the AI branch of clinical decision support (Bennett & Doub, 2016; Luxton, 2016). The incorporation of fuzzy logic into clinical decision support has been motivated by its capacity to reason with ambiguous terms and imprecision (Kumar et al., 2022; Masri & Jani, 2012). FCMs are a form of expert-based system that uses fuzzy logic for computation based on vague, experience-based linguistic terms provided by domain experts. While FCMs commonly do not outperform other AI-based decision-making algorithms in terms of prediction accuracy, their advantage over other decision-making architectures is their independence of large (mental) health client data sets to arrive at an initial, usable version of the algorithm and their comparable simplicity and interpretability for domain experts (Farahani et al., 2021) such as healthcare professionals. Consequently, FCMs have been used in many different decision-making systems for mental health (Billis et al., 2014; Kumar et al., 2022; Masri & Jani, 2012), sometimes in conjunction with other AI techniques to leverage the strengths of multiple methods (e.g., Chattopadhyay (2017) and Kumar et al. (2022)).

This chapter should be read as an introduction to FCMs and represents a starting point for developing decision algorithms for DMHS using FCMs. It demonstrates the development of FCMs based on a mental health example while highlighting some considerations for using the method and identifying future directions for research based on unresolved challenges and recent developments in the field of FCMs. We aspire to stir ongoing discussions and research approaches to making technologies and applications in mental health more intelligent, as needed to reduce today's and the future global burden of mental illness, on individuals and the healthcare system as a whole, and to move the mental health discipline forward (Thieme et al., 2023). To this end, this chapter can be read as a) a hands-on tutorial for those who want to use FCMs for modelling mental health problems and b) as a nuanced discussion of the method's strengths, challenges, and some recent developments in the field for those looking for a decision modelling approach in the absence of available training data, or those who wish to leverage the knowledge of domain experts during the construction process of such a decision model.

3.2 Background: Fuzzy Cognitive Maps

3.2.1 The basics: What are fuzzy cognitive maps?

Fuzzy cognitive mapping (FCM) is a soft computing technique that combines aspects of cognitive maps for decision making (Axelrod, 2015) and fuzzy logic (Zadeh et al., 1996). Kosko (1986) was the first to propose FCMs in response to limitations encountered when using cognitive maps for causal reasoning. He developed an algebra for fuzzy causal reasoning. Fuzzy reasoning is different from other forms of logic in that it allows varying degrees of truth rather than conceptualizations of something being either true or false (Zadeh et al., 1996). FCMs represent knowledge as directed graphs where elements of this knowledge are represented as graph nodes and their causal relationships as weighted edges (Papageorgiou, 2011, 2013). Each graph node has a numeric state which represents the magnitude of its presence in the specific knowledge system. A high numerical value means that the concept is strongly present, whereas a small numerical value means that it is weakly present. The strength of relationships between concepts in the knowledge domain represents the change in one concept caused by a change (an increase or a decrease) in another concept. The strength of relationships between concepts can be denoted with linguistic terms, such as *strong*, *sometimes*, or *a bit*. The capability of FCMs to deal with vague, linguistic descriptions of numerical values is one of the technique's major advantages over other modeling techniques in a discipline that is dominated by latent, inherently subjective and qualitative concepts such as mental health.

Figure 3.1 shows a simple FCM in its graph representation. In this FCM, the causal influence of concept C_1 on concept C_2 is indicated with weight w_{12}. Dur-

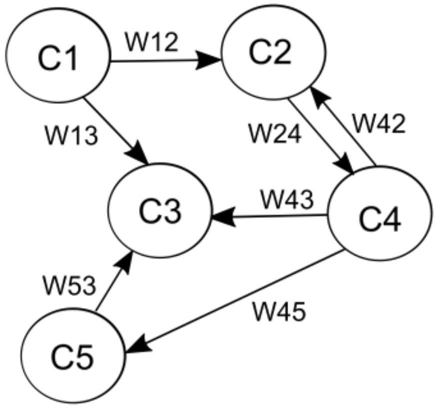


Figure 3.1: A simple Fuzzy Cognitive Map.

ing the construction of a FCM, linguistic weights are transformed into numerical weights using fuzzy inference. There are three types of numerical weights. Weights can equal zero, meaning that that there is no causality between two concepts. Weights larger than zero indicate causal increase (i.e. if w_{12} is larger than zero C_2 increases as C_1 increases and C_2 decreases as C_1 decreases). Weights smaller than zero indicate causal decrease (i.e. if w_{12} is smaller than zero C_2 increases as C_1 decreases and C_2 decreases as C_1 increases). Weights are usually expressed as numeric values in the interval [-1, 1] (Nápoles et al., 2024). -1 corresponds to the strongest causal decrease and 1 to the strongest causal increase. The other values express different levels of influence. FCM weights are commonly summarized in a weight matrix W of size $M \times M$ for M concepts in the knowledge domain. Each weight w_{ij} in W represents the causal influence of concept i on concept j with $w_{ij} \in [-1, 1]$. Once the concepts of an FCM and the weight matrix W are established, simulations can be initiated in which FCM concepts interact with each other according to their weights. For this, an update equation, or reasoning rule, is applied to the concept values at an initial state t to derive their respective values at a next time step $t + 1$. Commonly, concept values are summarized in a state vector $A^{(t)}$ of size $M \times 1$ with each value $A_i^{(t)}$ representing the value of concept i at time step t.

3.2.2 Classic fuzzy cognitive maps

Kosko (Felix et al., 2019; Kosko, 1986) developed an equation to update state vector A^t. This inference equation has been extended throughout the years to accommodate the idiosyncrasies of different FCM architectures and has been adopted widely in FCM applications (e.g. Amirkhani et al. (2017), Rahimi (2018) and Szwed (2021)). Kosko's modified inference equation is depicted in equation

3.1.

$$A_i^{(t+1)} = f\left(\sum_{\substack{j=1\\ i\neq j}}^{M} A_j^{(t)} w_{ji} + A_i^{(t)}\right) \tag{3.1}$$

where $A_i^{(t)}$ is the value of concept i at the current time step t, $A_i^{(t+1)}$ is the value of concept i at the next time step $t+1$, and f is a transfer function. The transfer function f ensures that the activation value of each FCM concept stays in a desired interval I, where $I = [0,1]$ or $I = [-1,1]$ depending on the knowledge system (Bueno & Salmeron, 2009). Some authors prefer to work with fixed sets instead of intervals to represent activation vs. inactivity of a concept, for example by allowing concepts to take values from the set $\{0,1\}$. Common transfer functions include the sigmoid function, hyperbolic tangent function and bivalent function (Nápoles et al., 2020). At this point, it should be noted that traditionally, inference time steps in FCMs are abstract and distinct from time in the real world. They do not correspond to measurable amounts of time such as minutes, hours, or years in the real world. This has implications for the interpretation of FCM predictions and has been discussed extensively in the recent FCM literature as we will see later in this tutorial. FCM simulations continue either until their concept values converge, meaning that the FCM produces the same state vector after a t-th iteration, or for a fixed T number of iterations. To determine whether the FCM converges, a residual ϵ is defined. The FCM converges if the difference between the updated concept values is smaller than ϵ in two consecutive iterations. The resulting concept values represent the final prediction of the FCM. It is possible that the FCM does not converge on a fixed state vector after t iterations. Alternatively, the FCM may produce a limit cycle, meaning that it periodically produces the same state vector after t iterations. A final possible outcome is that the FCM's configuration yields chaotic state vectors, meaning that after each inference iteration, it produces a novel state vector. For decision making and classification problems, convergence is desirable because predictions become consistent and thereby, easier for the domain expert to understand. For other use cases such as time series forecasting, convergence is less desirable (Felix et al., 2019).

3.2.3 Recent developments and implications for applying FCMs

Recently, a new perspective about the importance of interpreting the dynamic behavior of FCMs as part of the initial model development process has emerged (Nápoles et al., 2024). Nápoles et al. (2024) explain that the same FCM can produce very different simulation results, depending on the choice of the transfer function and that counterintuitively, a non-converging model can be a more truthful representation of the modelled dynamic system. More importantly, simulating the model's dynamic behavior can show whether it converges to a unique fixed state vector, no matter the input. In that case, the model is unsuitable for decision-making since it will eventually give the same output regardless of the provided input and no decision-making takes place. In such cases, applying another transfer function, or even another reasoning rule is recommended. Unique fixed-points have recently been addressed extensively by the FCM research community. They have been tied to the use of the sigmoid and hyperbolic tangent activation functions (Nápoles et al., 2018, 2020; Vergini & Groumpos, 2017). Another common issue with sigmoid and hyperbolic tangent activation is the saturation of activated concept values (Nápoles et al., 2024). As mentioned earlier, the purpose of activation functions is to keep concept values in a desired interval, such as $[0, 1]$, especially after incoming influences have been taken into account while iteratively updating the state vector $A^{(t)}$ using a reasoning rule. However, in practice, most activation functions, and in particular the widely adopted sigmoid and hyperbolic tangent functions, quickly saturate towards the extremes of their respective activation intervals, even for relatively small absolute input values. This behavior is problematic because it limits the extent to which different input values produce meaningful outputs, limiting the range of meaningful model inputs. Since activation functions have been tied to many detrimental issues with FCMs that limit the method's practical utility, especially for decision-making models, new methodological approaches have been developed that no longer make use of traditional activation functions, or Kosko's inference equation (Nápoles et al., 2024; Vergini & Groumpos, 2017).

In sum, FCMs iteratively calculate a future state of the knowledge system that they represent given an initial set of numeric concept values, a weight matrix that describes how their concepts interact with each other, and a chosen inference equation. In recent years, much effort has been devoted to addressing common issues with the traditional way FCMs have been used, pertaining to the interpretation of causal weights, side effects of using activation functions and interpreting dynamic behaviors of FCMs, among other topics (Nápoles et al., 2024; Vergini & Groumpos, 2017). The FCM research community is gradually moving towards new standards for employing FCMs for decision-making, as recently summarized by Giabbanelli and Nápoles (2024). Those wishing to apply the method in practice are strongly advised to consider recent developments in the field next to familiarizing themselves with the basics of the FCM methodology (section 3.2.2). Some issues that can arise when applying the method in practice may have been addressed by recent advances in the field.

3.2.4 What kind of problems do fuzzy cognitive maps solve?

FCMs are commonly used to represent knowledge about a specific dynamic system, including knowledge systems in the area of decision support, time series forecasting and pattern classification (Felix et al., 2019), and to model the development of the system over time using either experts' experiences, historical data, or both. In recent years, FCMs have been applied in clinical decision-making (Amirkhani et al., 2017), including the mental health domain (Billis et al., 2014; Brandl et al., 2023; Farahani et al., 2021; Papageorgiou et al., 2013). For example, Farahani et al. (2021) have developed a FCM to investigate the causal relationships between factors that influence psychological well-being in university students, including social support, resilience, and personal values such as hedonism and benevolence. The authors regard FCMs as a useful tool to predict psychological well-being based on a limited number of measurable factors. In addition, psychological counsellors at the university could easily interpret the constructed FCM. Recently, Billis et al. (2014) combined several FCMs in a modular architecture to detect pathological depressive tendencies in older adults living at home, with the goal of early and minimally obtrusive estimation of older adults' risk to develop clinical depression. The modular decision algorithm takes information about the senior's sleeping patters, based on sensor data extracted from wearable technology, as well as cognitive-emotional data as input and provides a depression diagnosis classification, severity level and an indication of the risk to develop pathological depression. The authors tested their decision algorithm preliminary using synthetic data and could detect geriatric depression based on early pathological signs with 78.66% accuracy. While FCMs do not commonly outperform other models in terms of classification accuracy, a major advantage of FCMs in mental health decision settings is their comparable simplicity, interpretability by domain experts, and explainable causal simulation of future states of the knowledge system that they represent (Apostolopoulos & Groumpos, 2023; Farahani et al., 2021).

3

3.2.5 Expert-based and data-driven FCM development

To construct a FCM, the most important concepts that represent the knowledge domain of interest have to be identified and the relationships between these concepts have to be determined so that the weight matrix W can be constructed. There are many ways to accomplish this task and they can be subdivided into two main approaches, *expert-based* and *data-driven* approaches. In expert-based FCMs, people who are knowledgeable about the components that make up the knowledge system provide the concepts of the FCM and their respective relationships to each other. Knowledge extraction methods commonly used in expert-based FCM development include focus groups and interview studies (Farahani et al., 2021; Olazabal & Reckien, 2015). Data-driven approaches usually assume that the set of relevant FCM concepts is provided a priori by an expert and focus on extracting the weight matrix W from other sources of information, most notably from historical clinical data. Data-driven approaches use a variety of learning techniques (Felix et al., 2019), including those based on Hebbian learning rule and error-driven approaches (e.g., regression, support vector machines, neural networks, and random forests), which are popular in other AI approaches for mental health. While error-driven approaches typically produce more accurate classification results, they exchange accuracy for causality and interpretability, usually resulting in black-box algorithms whose reasoning is difficult to grasp for mental health domain experts, fostering clinicians' skepticism and hesitation to use AI-based decision tools (Higgins et al., 2023). In addition, any data-driven approach to constructing FCMs still depends on expert involvement at some point (Felix et al., 2019). This in combination with the absence of available historical data in many mental health settings (Thieme et al., 2023), makes the expert-based approach most applicable for constructing an initial version of FCMs for mental health decision problems. Therefore, in this tutorial, we will focus on the construction of FCMs for eMental health by taking an expert-based approach.

3.3 Tutorial: Developing an eMental health FCM using the classic approach

In this section, the development of an expert-based FCM for an eMental health case study is demonstrated step by step, starting with considerations about how to model the mental health problem at hand and finishing with the transformation of linguistic weights such as *very high* to numerical weights using fuzzy inference. The section concludes with a demonstration of how a FCM arrives at a decision given an initial numeric state of its concepts.

3.3.1 Case study: Severity of loneliness

We will model the severity of experienced loneliness as an example to showcase the construction of an expert-based mental health FCM. The FCM was constructed during a single online focus group with two mental health experts. The focus group was held via Microsoft Teams and facilitated using Mural[1], an online whiteboard tool. We used FCMpy (Mkhitaryan et al., 2022), an open-source Python module for developing FCMs, to construct the weight matrix and to run scenarios with the constructed FCM. The research that led to the construction of this example was approved by the ethical board of the University of Twente, Enschede, The Netherlands (application nr. 230278).

[1]https://mural.co/

3.3.2 Preface: Model architectures for decision-making in mental health

Prior to constructing the FCM, we need to define the decision problem at hand. Traditionally, FCMs have been used to model the development of dynamic systems over time, i.e. to predict a future state of the knowledge system that they represent. Therefore, the decision how to act on the predicted future state of the knowledge system traditionally lies outside the FCM and is typically made by another actor, an algorithm or a human expert. In practice, however, many mental health decision problems boil down to some form of classification. For example, classifying someone's depressive symptoms as pathological and assessing their severity (Billis et al., 2014), or screening for a mental disorder (Masri & Jani, 2012). According to Felix et al. (2019), there are two basic types of FCM architectures that model a classification problem: class-per-outcome architectures and single-outcome architectures. In the former, each decision outcome is included in the FCM as a concept and these concepts are designated as output nodes, meaning that we can predict a decision class from their final inferred value. Output nodes are influenced via their connected weights by the remaining concepts in the FCM, the input concepts. In class-per-outcome architectures, the predicted decision class corresponds with the label of the output node with the highest value (e.g., Georgopoulos et al. (2003)).

Alternatively, in single-output architectures, there is one output node in the FCM. Consequently, we need to interpret the final state of the output node and translate it into an actionable decision based on some additional logic outside the FCM (Papakostas & Koulouriotis, 2010). Felix et al. (2019) distinguish between thresholding approaches and clustering approaches to derive a decision from a single-output architecture. For example, in a binary decision setting that uses a thresholding approach, output values exceeding 0.6 could be assigned one decision class while all predicted values below are assigned the alternative decision class. Determining a suitable threshold requires a good understanding of the decision problem that the FCM represents. In this tutorial, we will demonstrate the construction of a single-output FCM architecture. However, the steps involved in developing a single-output architecture FCM also apply to class-per-outcome FCMs.

3.3.3 Step 1: Determining FCM concepts

The first step after chosing the FCM architecture for the mental health decision problem is to gather the most important input concepts that influence the outcome in the FCM. At this step, experts are commonly involved in the construction of the FCM. Several knowledge extraction methods can be used to arrive at an initial set of input concepts together with experts, including in-depth interviews, focus groups, and surveys. Regardless of the chosen methodology, experts are asked to name and describe the most important symptoms, factors, or concepts that, according to their experience, matter when modelling the mental health problem at hand. The main output of this data collection are individual FCMs, commonly in their graph form, each representing the concepts and linguistic weights provided by one expert (steps 1 to 3). These individual FCMs are then aggregated to construct the weight matrix W (step 4).

For the loneliness case study, we asked two mental health experts to name psychological constructs and factors that influence the extent to which people feel lonely. In this simplified mental health example, we look at loneliness in general, without a specific target group (e.g., university students, the elderly) in mind. Since both experts were present in the focus group, they brainstormed concepts together and discussed their relevance before arriving at the final set of concepts, consisting of nine input concepts. All concepts included in the resulting loneliness FCM are summarized in Table 3.1.

Table 3.1: Summary of concepts in the loneliness FCM.

Concept	Abbreviation	Definition
Severity of Loneliness	SoL	The extent to which someone experiences loneliness. Output concept in the loneliness FCM.
Secure Attachment Style	SAS	The extent to which someone has an secure attachment style.
Independence	I	The extent to which one feels independent from other people and social norms.
Depressive and Anxious Mood	DAM	The extent to which one feels depressed and anxious.
Satisfaction with one's Appearance	SwA	The extent to which one feels good about their looks.
Self-efficacy	SE	The extent to which one feels in control of managing meaningful relations.
Social Skills	SS	One's ability to build and maintain social relations.
Quantity of Social Contacts	QNS	The number of social relations one has.
Quality of Social Contacts	QLS	The quality of the social relations one has.
Identification with a Community	IwC	The extent to which one feels that they belong, e.g., to a community, either in the physical world or in spirit.

3.3.4 Step 2: Determining relations between FCM concepts

In the next step, for each pair of concepts, the experts indicate whether there is a causal relation between the two concepts. If the concepts are causally related, the experts clarify whether the relation is one-sided or whether the concepts influence each other mutually. During this step, drawing a schematic graph representation of the FCM, either digitally using a tool such as Mural, or physically on a sheet of paper, is strongly recommended to facilitate the thinking process and to keep an overview of the connections between FCM concepts. Once causal relations have been identified, for each relation between two FCM concepts, the experts indicate whether the relation is excitatory (e.g., an increase in Secure Attachment Style (SAS) increases one's Independence (I)) or inhibitory (e.g., an increase in SAS decreases one's Severity of Loneliness (SoL)). In the graph representation of an FCM, excitatory relations are indicated with a $+$ sign and inhibitory relations with a $-$ sign. Excitatory relations are also called causal increase and inhibitory relations are called causal decrease. Typically, experts use IF-THEN rules to describe the relation between concepts, such as

> IF the value of concept C_i is A THEN the value of concept C_j is B.

One such an IF-THEN rule for our exemplary FCM is

> IF the value of concept SAS is increased THEN the value of concept SoL is decreased.

Filling in the relations in our exemplary FCM results in the graph shown in Figure 3.2.

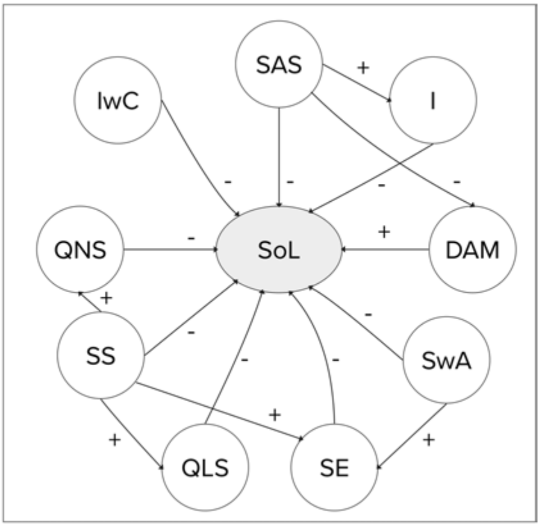


Figure 3.2: Identified causal relations in the exemplary loneliness FCM.

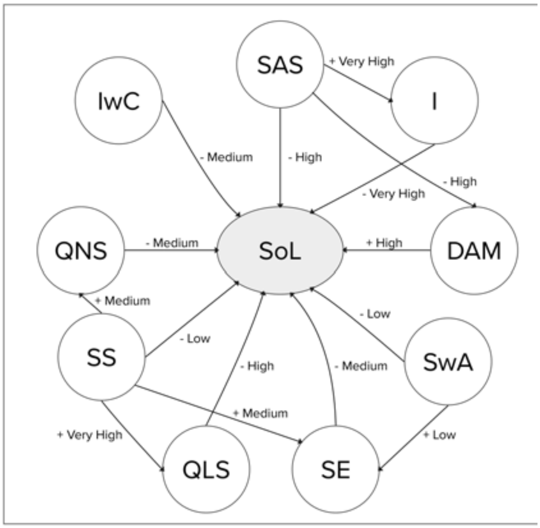
3.3.5 Step 3: Determining the strength of relations

Once the causal influences between concepts have been identified, experts are asked to indicate how strong each influence is. In practice, this step is often executed in conjunction with the previous step. Mental healthcare practitioners often deal with an imprecise and vague decision basis. In contrast to many other modelling approaches, FCMs embrace imprecision by allowing the strength of causal influences in the FCM to be described in fuzzy terms. For this, a linguistic variable $Influence$ is declared that is represented by a set of linguistic terms T that experts may use to describe the extent to which concepts impact each other causally:

$$T(Influence) = \{\text{Very very low, Very low, Low,}$$
$$\text{Medium, High, Very High, Very very high}\}$$

The linguistic terms can be adapted to the terminology of the knowledge system that the FCM represents. However, different gradations of high and low are most commonly used to describe the degree of causal influences between concepts. The resulting graphical representation of this step is shown in Figure 3.3.

Figure 3.3: Identified strength of causal relations in the exemplary loneliness FCM using linguistic terms to express influence. This is the resulting graphical representation of step 3 in the FCM construction process.



3.3.6 Step 4: Weight aggregation

The final step involves aggregating the linguistic weights obtained from all experts and transforming them into numerical weights, essentially yielding the weight matrix W. Since the causal influence between FCM concepts commonly takes values in the interval $[-1, 1]$, the linguistic terms used to describe influence take values in between these interval limits. Where linguistic terms fall in the interval is expressed through membership functions μ_{vvl}, μ_{vl}, μ_{l}, μ_{m}, μ_{h}, μ_{vh}, μ_{vvh}. Figure 3.4 depicts the triangular membership functions commonly applied for the seven-item set of linguistic terms ranging from *Very very low* to *Very very high*.

Once the membership functions of the linguistic terms are defined, the linguistic weights are commonly aggregated in three steps: 1) applying a fuzzy implication rule, 2) combining the membership functions of individual experts, and 3) defuzzifying the aggregated membership functions to derive numerical weights for each causal relation in the FCM (Mkhitaryan et al., 2022). We refer the interested reader to Papageorgiou (2013) and Felix et al. (2019) for the more in-depth theoretical and mathematical foundations for each of these steps. For this tutorial, we used commonly applied methods for each step: we applied Mamdani's fuzzy implication rule to determine the extent to which each membership function was activated for each causal relation (Giabbanelli, 2014; Nandi, 2012), the algebraic SUM aggregation operation (Giabbanelli, 2014;

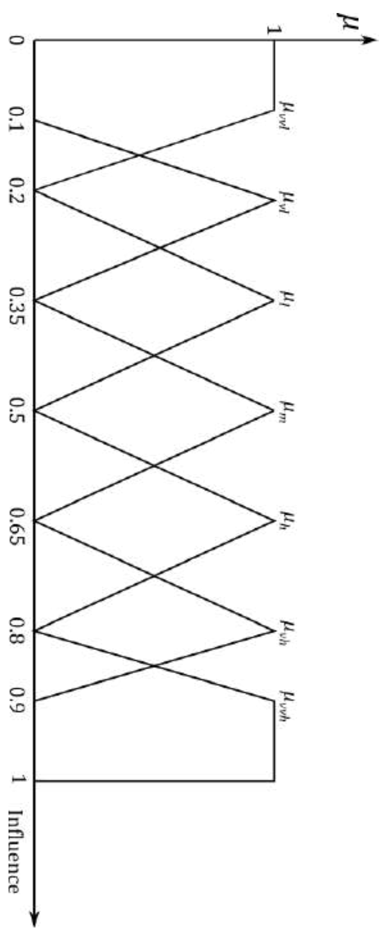


Figure 3.4: The seven membership functions corresponding to linguistic weights commonly used to describe the grade of causal influence in FCMs.

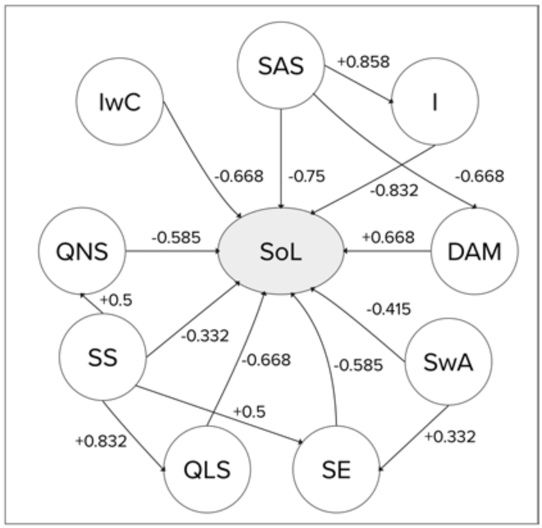


Figure 3.5: Identified strength of causal relations in the exemplary loneliness FCM using numeric weights to express influence. This is the resulting graphical representation of step 4 in the FCM construction process.

Papageorgiou, 2011) for combining the linguistic weights provided by individual experts, and center of gravity (COG) defuzzification (Papageorgiou, 2013; Stylios & Groumpos, 2004) to derive numeric weights. The resulting weight matrix W is shown below and the respective graphical representation of this step is shown in Figure 3.5

$$
W = \begin{array}{c c}
 & \begin{array}{cccccccccc} \text{SAS} & \text{I} & \text{DAM} & \text{SwA} & \text{SE} & \text{SS} & \text{QNS} & \text{QLS} & \text{IwC} & \text{SoL} \end{array} \\
\begin{array}{c} \text{SAS} \\ \text{I} \\ \text{DAM} \\ \text{SwA} \\ \text{SE} \\ \text{SS} \\ \text{QNS} \\ \text{QLS} \\ \text{IwC} \\ \text{SoL} \end{array} &
\begin{bmatrix}
 & 0.858 & -0.668 & & & & & & & -0.75 \\
 & & & & & & & & & -0.832 \\
 & & & & & & & & & 0.668 \\
 & & & & 0.332 & & & & & -0.415 \\
 & & & & & & & & & -0.585 \\
 & & & & 0.5 & & 0.5 & 0.832 & & -0.332 \\
 & & & & & & & & & -0.585 \\
 & & & & & & & & & -0.668 \\
 & & & & & & & & & -0.415 \\
 & & & & & & & & &
\end{bmatrix}
\end{array}
$$

3.3.7 Simulating mental health scenarios using FCMs

After constructing the FCM weight matrix W with the help of experts, the FCM can be used to simulate scenarios based on varying input values for its concepts. This is done by iteratively applying Kosko's inference equation, or a variation of the updating equation that better fits the idiosyncrasies of the represented knowledge system. Following the classic FCM approach, we use Kosko's modified equation. To update the initial state of the FCM, the extent to which the input concepts are present has to be determined and transformed into a numeric input. For the loneliness FCM, we decided to let numeric input values as well as the values of the output concept SoL take values in the interval $[0, 1]$ with 0 indicating that the concept is minimally present and 1 indicating that it is profoundly present. While FCMs are flexible when it comes to the interval in which their inputs vary, to preserve the interpretability of the FCM, one should consider input values that make sense intuitively. For the loneliness FCM, for instance, a value of 0.9 for the concept DAM can be interpreted intuitively as that the person feels very depressed and anxious. In contrast, a value of 0.1 can be interpreted as that the person does neither feel depressed nor anxious. A negative value of -0.9, however, would be difficult to make sense of intuitively and therefore, does not serve the interpretability of the FCM.

For illustrative purposes, we calculate a first iteration for a fictitious scenario of a person who feels moderately lonely ($SoL = 0.5$), but who has a stable support network, good social skills and feels confident. We apply the $sigmoid$ transfer function to keep activated values in the $[0, 1]$ interval.

$$f(x) = \frac{1}{1 - e^{-x}}$$

to the matrix multiplication between the initial state vector $A^{(t)}$ and the weight matrix W, which is what applying an updating equation essentially boils down to (Zhong et al., 2008).

$$A^{(t+1)} = f\left(A^{(t)} \times W + A^{(t)}\right)$$

$$= f\left(\begin{bmatrix} 0.8 \\ 0.7 \\ 0.3 \\ 0.5 \\ 0.6 \\ 0.8 \\ 0.9 \\ 0.9 \\ 0.8 \\ 0.5 \end{bmatrix} \times \begin{bmatrix} 0 & 0.858 & -0.668 & 0 & 0 & 0 & 0 & 0 & 0 & -0.750 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.832 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.668 \\ 0 & 0 & 0 & 0 & 0.332 & 0 & 0 & 0 & 0 & -0.415 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.585 \\ 0 & 0 & 0 & 0 & 0.500 & 0 & 0.500 & 0.832 & & -0.332 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.585 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.668 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.415 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0.8 \\ 0.7 \\ 0.3 \\ 0.5 \\ 0.6 \\ 0.8 \\ 0.9 \\ 0.9 \\ 0.8 \\ 0.5 \end{bmatrix}\right)$$

The first iteration step yields the updated state vector $A^{(t+1)}$.

$$A^{(t+1)} = \begin{bmatrix} 0.800 \\ 0.797 \\ 0.442 \\ 0.500 \\ 0.762 \\ 0.800 \\ 0.786 \\ 0.827 \\ 0.800 \\ 0.069 \end{bmatrix}$$

This calculation is repeated using the updated state vector until the FCM concept values converge with $\epsilon = 0.001$. For the fictitious initial state $A^(t)$, our example FCM predicts a decrease in loneliness from 0.5 to 0.048 and converges after 8 iterations.

At this point it should be noted that based on recent insights about the importance of making a thorough investigation of the dynamic behavior of the FCM part of the initial model development process, it is recommended to continue simulation experiments, as illustrated by Nápoles et al. (2024), to ensure that the obtained simulation results do not represent a unique fixed point before applying the FCM in practice. This step is omitted in this tutorial since we do not apply the loneliness FCM in practice. In addition, any actionable conclusion based on simulation results would depend on the specific use context of the loneliness FCM. Taking an eMental health service as an example aimed at detecting a need for professional intervention, this preliminary result could be interpreted as requiring no further intervention from a professional.

3.3.8 Software Tools for Developing FCMs

In recent years, a number of software tools have been developed and made available by researchers working with FCMs. We recommend that any researcher who wishes to use FCMs familiarizes themselves with the available tools as they can facilitate the FCM development process tremendously. Felix et al. (2019) and Mkhitaryan et al. (2022) provide a detailed overview of the different tools available. Here, we will highlight a few. The existing software solutions for FCMs can be distinguished based on some desired characteristics, such as supporting scenario analyses, enabling data-driven learning, or having a graphical interface for less programming-savvy users. Table 3.2 shows an overview of the software solutions highlighted in this tutorial.

Table 3.2: Overview of selected software solutions for developing fuzzy cognitive maps.

Software	Reference	Scenarios	Graphical UI	Data-driven Learning	Open Source
Mental Modeler	Gray et al. (2013)	✓	✓	×	×
JFCM	De Franciscis (2014)	✓	×	✓	×
FCMExpert	Nápoles et al. (2018)	✓	✓	✓	×
FCMpy	Mkhitaryan et al. (2022)	✓	×	✓	✓

Mental Modeler (Gray et al., 2013) is a web-based FCM modeling tool designed to support the decision-making of a group of experts. Using Mental Modeler, they can collaboratively design and test a FCM with simple scenarios. Mental Modeler is targeted at domain experts with little programming experience. However, its source code is not open-source and it only has limited options for scenario analysis, making it unsuitable for integration in other software applications. On the other end of the spectrum, Java Fuzzy Cognitive Maps (JFCM) (De Franciscis, 2014), is an open-source library for the Java programming language that offers standard FCM components that can be used for a variety of FCM problems and can be extended programmatically to accommodate more complex projects. At the same time, it lacks a graphical user interface, making it unsuitable for non-programmers.

3

FCMExpert (Nápoles et al., 2018) has been introduced as a downloadable[2] software tool for designing, training and executing FCMs, including an easy-to-use graphical interface and a variety of learning and experimental options. Recently, Mkhitaryan et al. (2022) have contributed FCMpy, an open-source Python module that facilitates aggregating the input received from experts, allows the simulation of scenarios given an existing FCM structure, and constructing a weight matrix based on historical data. Since it is open-source researchers can extend the code base with their own code to accommodate the idiosyncrasies of their research and conveniently integrate the module into their own machine learning pipelines.

3.4 Considerations when developing FCMs for eMental Health

In addition to common considerations for the FCM development process, such as making a sound decision about the classification architecture for the FCM, choosing an interval in which concept values can take values, and making an informed choice about the reasoning rule that is used to simulate the FCM's dynamic behavior, applying FCMs for the mental health domain comes with specific considerations and challenges that we will highlight and discuss in this section. We will first discuss some considerations about extracting the FCM architecture and linguistic weights from domain experts. Afterwards, we

[2]Accessed 2023-05-25: https://sites.google.com/view/fcm-expert

will highlight some of the advantages and limitations of FCMs for constructing decision algorithms for eMental health and put them into the context of some recent developments in the field.

3.4.1 Expert involvement

When choosing a study setup to extract experts' knowledge, the researcher should consider the implications of collecting data simultaneously from several experts (e.g., during a focus group), compared to collecting data in separate sessions (e.g., one-on-one interviews). When extracting knowledge from several experts simultaneously, group dynamics and professional hierarchies can affect data collection. For example, without active moderation by the researcher, experts may refrain from providing input until the most senior expert in the group has explained their point of view, or they may adjust their responses to accommodate the viewpoints of a more senior expert. This can ultimately lead to results produced by group think. In contrast, an advantage of interviewing several experts simultaneously is that they can discuss their experiences with each other and thereby, arrive at a more coherent representation of their joint knowledge, essentially moving part of the aggregation of their individual inputs into the data collection session.

There are different gradations of conducting asynchronous data collection for FCM development and usually, they compromise flexibility during data collection and ease of managing the aggregation process. The researcher can decide to keep components of the FCM constant between asynchronous sessions which generally facilitates aggregation. FCM components that the researcher may wish to keep constant are a) FCM concepts, and b) which causal relations experts should evaluate. FCM concepts that are pre-determined and thus, not part of the knowledge extraction process, may originate from earlier research, including literature reviews on the mental health problem under investigation. Keeping the specific causal relations that experts evaluate constant counteracts situations in which some relations are evaluated by most experts while others may only be evaluated by a single expert which can raise questions regarding the reliability of the specific weight. Olazabal and Reckien (2015) propose to exclude experts' weights that differ too much from the mean to reduce uncertainty in the aggregation process, or alternatively, to only incorporate weights that are suggested by a minimum number of experts, which they refer to as "2p+" or "3p+" networks. Alternatively, Stylios and Groumpos (2004) have proposed adjusting an individual expert's weights if they divert too much from the input of the remaining experts by applying a credibility factor during weight aggregation.

Since aggregating the weights of multiple experts is essentially a quantitative process, the question arises how many experts are needed to construct a valid and reliable FCM. FCM studies that extract the concepts of the FCM alongside the weights can use the extent to which additional sessions add new concepts as criterion to determine the usefulness of including more experts (Özesmi & Özesmi, 2004). However, there is no standard way to deduce the sample size of a cognitive mapping study prior to data collection and consequently, sample sizes differ dramatically between studies (e.g., Papageorgiou et al. (2013) and Reckien (2014)). However, rather than sample size, the careful selection of experts and the quality of their input tends to determine the usefulness of the resulting FCM (Reckien, 2014).

3.4.2 History of user measurements and FCMs

In mental health decision-making, much like in medical decision-making (Bourgani et al., 2013), the progression of symptoms forms a crucial decision criterion for determining the severity of the mental health problem alongside the intensity with which the client experiences the symptoms. Decisions are thereby dependent on the client's current as well as past symptoms. For modelling the client's mental state using FCMs, this means that the FCM should include information about the past values of its concepts in the decision-making. However, including temporal information in FCMs is not straightforward (Olazabal & Reckien, 2015) and quickly exhausts the modelling possibilities of classic FCMs. For example, to model the influence of a person's past measurement of some mental health variable on a current outcome, such as the influence of one's past *depressive and anxious mood* on their current experience of *severity of loneliness*, one will likely conclude that the influence depends on the person's current mental state. Therefore, modelling past influences on current states goes beyond direct causal influences. The temporal limitations of FCMs have been widely acknowledged (Acampora & Loia, 2011; Carvalho, 2013; Nair et al., 2020; Zhong et al., 2008). Over the years, several efforts have been made to enhance the capabilities of FCMs to include a temporal dimension, as summarized by Acampora and Loia (2011) and later, by Nair et al. (2020). However, these efforts have been made with different definitions of *time* and different purposes in mind. As a consequence, the available literature about temporal FCMs can be confusing, and at times, misleading for the novel reader.

For example, Kosko's modified inference equation (Stylios & Groumpos, 2004) that we introduced as part of this tutorial can be misinterpreted regarding its relation to time since it considers concept values from a previous time step. However, this has nothing to do with the definition of time as a mental health researcher would intuitively understand it. This modified rule was developed to force the updating of concepts that are not influenced by other concepts in the FCM, to counteract their unintentional decay throughout the iterative inference process. The internal time steps of an FCM are abstract and distinct from time as we experience it. Therefore, we cannot simply assume that FCM inference iterations correspond to specific (hypothetical) measurement moments in the physical world. Carvalho and Tomè (2000) coined the term base time (*b-time*) to describe the time dimension in which FCM inference iterations occur. Their rule-based extension to FCMs (RB-FCM) addresses a fair amount of the temporal issues in FCMs, but exchanges the simplicity of traditional FCMs for semantic and temporal soundness, making RB-FCMs theoretically more complex and difficult to use in practice (Carvalho, 2013; Carvalho & Tomè, 1999).

Other efforts have been directed towards modelling long-term developments of the knowledge system that the FCM represents by introducing the idea of multiple developmental stages that the FCM undergoes (called *cognitive eras*) (Acampora & Loia, 2011). Between different cognitive stages, the structure of the FCM can change, including the removal and addition of concepts and changes in FCM weights. The conditions for shifting from one cognitive era to another are pre-defined. This approach, called timed automata FCMs (TAFCMS), is methaphorically comparable to the developmental stages that humans undergo as they develop from childhood to adolescence, to adulthood and finally, late adulthood, for which age intervals serve as triggers that initiate the transition from one stage to the next.

Finally, a recent re-interpretation of how concepts influence each other using causal weights may offer new ways to incorporate a person's mental health history in FCMs. Vergini and Groumpos (2017) point out that rather than the absolute values of concepts, their observable change Δ between measurement moments triggers change in other concepts and therefore, Δ should be the basis for simulating the dynamic behavior of FCMs, not absolute measurement values:

$$\Delta C_i^{(t)} = C_i^{(t)} - C_i^{(t-1)}$$

where $\Delta C_i^{(t)}$ is the change in concept C_i at time step t, based on the observable and absolute values of C_i at time steps t and $(t-1)$. The authors introduce a new reasoning rule that reflects this change and illustrate its use. Future work should investigate whether this subtle change in interpretation that has implications for the choice of a reasoning rule and the interpretation of time in FCMs can satisfy the temporal needs of FCMs in mental health. For example, by extending the use of Δ beyond consecutive measurement moments, FCMs may be able to reflect changes in concepts that lie further in the past than one time step and may thereby potentially represent a history of measurements.

At the time of writing this tutorial, the efforts described above (Acampora & Loia, 2011; Carvalho & Tomè, 2000; Vergini & Groumpos, 2017; Zhong et al., 2008) do not offer a straightforward and pragmatic way to incorporate a client's history of mental health symptoms in a decision model using fuzzy cognitive mapping. Instead, the mental health researcher must make pragmatic design choices to accommodate temporal information in their FCM-based decision algorithm. For example, they may add handcrafted rules to the FCM algorithm that can take consecutive measurements into account. More research is needed to establish ways in which FCMs can accommodate the history of a client's mental health symptoms effectively.

3.4.3 Modelling beyond direct causal influences

Related to the temporal limitations of FCMs is their limited capacity to model relations between concepts beyond simple symmetric and monotonic causal influences (Apostolopoulos & Groumpos, 2023; Carvalho, 2013; Özesmi & Özesmi, 2004). Taking the loneliness FCM as an example, we identified a positive causal influence between a person's social skills and the quality of their social relationships. It is questionable, however, whether an improvement in one's social skills influences the quality of one's relationships in the exact same way as a decrease in one's social skills would impact the quality of one's relationships (i.e., it is questionable whether the influence is symmetric). The terminology and types of influences that FCMs offer limit how experts can express their knowledge. One particular limitation that we encountered in the context of mental health is the lack of options to model interactions that arise from two or more concepts, i.e., the co-occurrence of multiple causes. There is no option to model individual incoming causal influences on a concept, or the interaction between specific influences from multiple concepts. In FCMs, all incoming influences are aggregated and subjected to the same transfer function, or another function that is used to normalize incoming influences. Again, taking the loneliness FCM as an example, while *Quantity of social relationships* (QNS) and *Quality of social relationships* (QLS) indeed both influence one's experienced *Severity of loneliness* (SoL) directly, our experts explained that part of their influence arises from their combination, which is difficult to represent using FCMs. One can have many acquaintances, but still feel lonely. Likewise, having a few physically distant good acquaintances can make people feel lonely. This implies that direct influences only partially capture the influence these concepts have on SoL, and this issue is likely relevant for other mental health FCMs, too. Another modelling limitation that we encountered in the context of mental health is the lack of conditionality in FCMs, for example by encoding IF-THEN rules in the influences between concepts. In the loneliness FCM, our experts explained that it would be more accurate to model the influence of *Identification with a Community* (IwC) on SoL as increasing in importance in the absence of other concepts, such as *Independence* (I) and the number and quality of social relationships (QNS, QLS). Put differently, for small values of other concepts in the FCM, the weight between IwC and SoL should be larger. This, however, cannot be modelled with traditional FCMs. It is therefore challenging to model complex mental health problems accurately with only direct causal influences at one's disposal (Carvalho & Tomè, 1999). The current state of the FCM methodology quickly puts the mental health researcher in a situation where they

3

have to compromise the simplicity and ease of use of traditional FCMs and the more adequate representation of mental health problems. However, the impact of these modelling limitations on the accuracy and usefulness of FCMs in eMental health is largely unknown and needs further investigation. Despite their modelling limitations, however, FCMs tend to perform reasonably well in other domains (Carvalho, 2013; Nápoles et al., 2020) and it remains to be seen whether this holds for the mental health domain.

3.4.4 Embedding FCMs in Digital Mental Health Services

FCMs are not stand-alone decision support systems. Depending on their architecture and embedding in any specific eMental health service, FCMs rely on other actors, algorithmic or human, to interpret their final states that result from iterative simulation and to arrive at actionable advice. Taking the loneliness FCM developed as part of this tutorial as an example; it does not provide actionable advice in its current state of development. It predicts a decrease in loneliness based on its architecture and an initial state, but whatever real-world actions follow the model's prediction must be considered separately and depend on the context in which the FCM model is used. In addition, data input streams that provide the FCM with initial states for simulation must be defined for the digital mental health service and developed alongside measurement protocols including decisions about the frequency with which input data is collected and which data collection tools are to be used. Exemplary data collection tools are questionnaires, interaction data logs, or sensory data extracted from wearables. In this context, it is important to realize that existing validated measurement tools for mental health may not be suitable for detecting small changes over time and may therefore be less suitable for automatic monitoring of mental states using eMental health solutions, including those based on FCMs. In the absence of validated measurement tools that fit the application's measurement interval, it may be necessary to develop customized questionnaires, or to consider other, less obtrusive data collection tools. Depending on the chosen data collection tools, situations may occur where not every FCM concept is measured in the same measurement interval (e.g., the concept is not expected to change in the short term and therefore, the same concept value is used for multiple consecutive inference moments), or for some other reason, a new measurement is missing. In such a situation, a reasonable and pragmatic decision must be made based on the specific use context of the FCM, such as re-using a previously measured value for the concept in question. Developing a FCM-based decision algorithm therefore goes beyond the construction of the FCM itself. To exploit their usefulness to the fullest, its integration in the data streams of the digital health service, the service model and remaining functionalities need to be given as much concern as the construction of the model itself.

3.4.5 Fuzzy mental health concepts

Another interesting considerations is the underuse of fuzzy properties in current FCM approaches (Carvalho, 2013; Nápoles et al., 2024). The assumption that FCM concepts (e.g., *depressive and anxious mood* in the loneliness FCM) increase linearly underlies most existing FCM modelling approaches (Nápoles et al., 2020). To perform algebraic calculations with the FCM, defuzzified numeric state values are needed that represent the magnitude with which a concept is present in the knowledge system at any given state. These concept values are for example obtained via psychometric questionnaires. Current psychometric approaches to measuring psychological constructs tend to model constructs as existing on a spectrum and by incorporating them as data collection tools into FCMs, current FCM approaches miss out on the capacities of true fuzzy systems that do not only allow fuzzy weights, but also fuzzy input and output concepts (Carvalho & Tomè, 2000). The impact of introducing true fuzzy psychological concepts instead of psychometric numeric approximations on the accuracy of FCM predictions is another interesting consideration to be explored, as is the impact of such fuzzy concepts on the resulting ease of interpretation for domain experts.

3.4.6 Evaluating fuzzy cognitive maps in eMental health

Finally, evaluating FCM-based decision models for eMental health deserves some consideration. FCMs are evaluated in many different ways, including scenario analyses (Papageorgiou et al., 2013), classic machine learning evaluation methods based on prediction errors (e.g., Felix et al. (2019)), and more recently, through extensive exploration of the model's dynamic behavior using simulations (Nápoles et al., 2024). While developing a FCM to support decision-making in an eMental health setting, one will likely be confronted with questions such as "How accurate is the algorithm?", "How reliable is it?", and "How accurate is good enough?". A well-informed choice should be made about the evaluation metric, taking into account the limitations and opportunities of evaluating a FCM in standard evaluation approaches in mental health, including randomized controlled trials, and other experimental settings involving real client mental health data. However, regardless of the specific chosen evaluation metric, there is no golden performance standard to decide when an algorithm is "good enough" to move from the proof-of-concept stage into eMental health practice. Consequently, the mental health researcher will need to decide on a case-by-case basis if the FCM algorithm fulfills its requirements adequately in the specific use context at hand, often involving an interdisciplinary view on the model's predictions and the extent to which it enhances rather than obstructs existing clinical practices (Thieme et al., 2023).

3.5 General Discussion

In this chapter, we introduce fuzzy cognitive mapping, a soft computing technique that can be used to develop decision-making models for eMental health. We discuss a hands-on example for applying the methodology to a mental health problem and identify advantages and unresolved challenges when applying FCMs to eMental health. Regarding the advantages of FCMs, (1) they allow the construction of a first usable version of a mental health (decision) model without depending on the availability of large amounts of historical client data. This makes FCMs a viable modelling approach for situations in which no training data is available, which is a common challenge in mental health research. Indeed, as recently investigated in a review of the methodological and quality flaws in AI research for mental health (Tornero-Costa et al., 2023), current research efforts mostly use private data, aiming to develop new AI models rather than validating existing models using public data sets. The availability of public mental health data sets is likely cause and effect in maintaining a preference for private data in AI research for mental health. In addition, the distribution of AI research in mental health is unbalanced across mental health problems. Using the International Classification of Diseases 11th Revision (ICD-11), 77 out of 153 included studies were related to mood disorders, including depressive disorders and bipolar or related disorders. This suggests that finding training data for mental health categories that are less frequently researched may prove even more challenging. A second advantage (2) of FCMs, in their traditional form, is their explainability (Apostolopoulos & Groumpos, 2023). The black-box nature of conventional and especially, deep AI models has been linked to clinicians' reluctance to trust the predictions of AI models and consequently, their reluctance to adopt them in clinical practice (Apostolopoulos & Groumpos, 2023; Sendak et al., 2020; Tornero-Costa et al., 2023). Likewise, (3) the reliance of FCMs mental health experts such as psychological therapists and psychiatrists during their construction process inherently promotes professionals' understanding and acceptance of the resulting decision model (Higgins et al., 2023; Thieme et al., 2023).

Regarding the disadvantages of FCMs, as discussed in more detail earlier in this tutorial, (4) there is a pronounced trade-off between the model's simplicity and the accurate representation of mental health constructs since FCMs offer limited ways for modelling influences beyond direct causality and the temporal dimensions of the represented mental health construct. Finally, (5) though promising progress has been made over the years (Billis et al., 2014) and recently, in terms of addressing well-known methodological issues, including practical limitations tied to the use of certain activation functions (Nápoles et al., 2024) and the interpretation of causal influences (Vergini & Groumpos, 2017), classification algorithms based on fuzzy cognitive mapping are usually not on par with their black-box AI counterparts in terms of classification accuracy

(Farahani et al., 2021; Felix et al., 2019). Finally, a recent surge in technical and methodological contributions to the FCM discipline can make it challenging for novel applied FCM researchers in mental health to enter the field because the current FCM research landscape is dispersed, progressing fast, and sometimes, producing contradictory results. Giabbanelli and Nápoles (2024) portrayed the current FCM research landscape and the interested reader is recommended to consult their work as a starting point for delving into recent developments in the field.

3.6 Conclusion

The current state of AI in mental health is still considered to be in the phase of proof-of-concepts and showcasing technical feasibility rather than providing systems that are ready to create clinical impact. Existing AI tools are not commonly optimized for integration in clinical workflows in mental healthcare. Additionally, the black-box nature of data-driven algorithms and the exclusion of clinicians in the development of AI tools foster clinicians' reluctance to integrate AI-based eMental health into their practice. Fuzzy cognitive maps (FCMs) are an alternative modelling approach for developing decision-making algorithms in eMental health. Their inherent expert-basis, explainability and independence of large clinical data sets to arrive at initial usable decision models makes them especially attractive for the mental health domain. With eMental health being on the rise to meet growing mental health care needs, the need for applications with decision-making capabilities will also rise. This article demonstrates the development of FCMs for eMental health and highlights what the technique can offer for the discipline, but also discusses unresolved challenges and topics for future research, in particular, given recent developments in the field.

Developing a monitoring module to guide older mourners to offline support in an online grief service

Based on:
Brandl, L., van Velsen, L., Brodbeck, J., Jacinto, S., Hofs, D., & Heylen, D. (2023). Developing an eMental health monitoring module for older mourners using Fuzzy Cognitive Maps. *Digital Health*; 9. doi:10.1177/20552076231183549.

Prologue

“In the past two weeks, how often have you avoided getting in touch with friends and/or family?", Orphea mouths silently while reading the question from her computer screen. She leans back in the chair and crosses her arms,“ ‘Avoided’...did I avoid anybody?". A light breeze enters the room through the patio door and Orphea reaches for Justus’ sweater hanging over the back of the chair. As she reaches for it, she pauses. The sweater was a birthday gift from their daughter years ago and it was one of her husband’s favorites. “It is not that I avoid them. Some days, it is just easier not to see anybody.". Orphea clicks the checkbox next to the “A few days" response as she pulls the sweater over her head."

Abstract

Objective: Effective internet interventions often combine online self-help with regular professional guidance. In the absence of regularly scheduled contact with a professional, the internet intervention should refer users to professional human care if their condition deteriorates. The current chapter presents a monitoring module to recommend proactively seeking offline support in an eMental health service to aid older mourners.

Method: The module consists of two components, a user profile that collects relevant information about the user from the application. The user profile enables the second component, a fuzzy cognitive map (FCM) decision-making algorithm to detect risk situations and to recommend the user to seek offline support, whenever advisable. In this chapter, we show how we configured the FCM with the help of eight clinical psychologists and we investigate the utility of the resulting decision tool using four fictitious scenarios.

Results: The current FCM algorithm succeeds in detecting unambiguous risk situations, as well as unambiguously safe situations, but it has more difficulty classifying borderline cases correctly. Based on recommendations from the participants and an analysis of the algorithm's erroneous classifications, we propose how the current FCM algorithm can be further improved.

Conclusion: The configuration of FCMs does not necessarily demand large amounts of privacy-sensitive data and their decisions are scrutable. Thus, they hold great potential for automatic decision-making algorithms in mental eHealth. Nevertheless, we conclude that there is a need for clear guidelines and best practices for developing FCMs, specifically for eMental health.

4.1 Introduction

The loss of a partner is a common occurrence in later life. While most older mourners cope with the loss of their spouse, some (about 10% according to Lundorff et al. (2017)) struggle with bereavement and develop prolonged grief (Aoun et al., 2015; Spahni et al., 2015). Severe grief symptoms that persist longer than six months after bereavement are characteristic of prolonged grief and can result in other mental and physical problems, including depression and cardiovascular problems, and in extreme cases, suicidal tendencies (Molina et al., 2019). Internet-based interventions have been shown to be effective for treating (prolonged) grief (Brodbeck et al., 2019; Eisma et al., 2015; Wagner et al., 2006).

Some eMental health services combine a web-based self-help program and minimal, but regular therapist contact (Baumeister et al., 2014). One form of eMental health that blends self-help with professional contact is support on-demand. In support on-demand, where no regular contact with a healthcare professional is planned, the client initiates contact with a professional. These client-initiated contacts are focused on the client's specific needs at that moment, while otherwise following the eMental health service on their own (Dahlin et al., 2020; Oromendia et al., 2016). In settings where no regular contact with a professional is scheduled, eMental health services have the responsibility to refer users to professional human care if the condition of the user deteriorates (Tielman et al., 2019). Based on existing mental health safety protocols (Knight et al., 2015), an analysis of professionals' core competencies in mental health telephone triage (Sands et al., 2013) and in-depth discussions with experts, Tielman et al. (2017) have developed safety protocols for referring users of eMental health services to human care in case of a risk situation. This referral can take the form of an automatic system action. Altrnatively, the system can aim at convincing clients to take the initiative in contacting a professional, promoting a self-referral process. Tielman et al. (2017) distinguish two stages in this (self-)referral process: information gathering and decision-making.

The goal of the information gathering stage is to identify whether a risk situation exists. In most current eMental health services (Robinson et al., 2015; Spence et al., 2014), a healthcare professional is involved at this stage. For example, in the Reframe IT intervention to reduce suicide risk in secondary students, human caregivers regularly screened the students' scores on a distress check-up and immediately alarmed the school staff and other appropriate healthcare authorities if necessary (Robinson et al., 2015). However, human assessment, such as 24/7 messaging services and screening of distress, are unfeasible for self-help eMental health services that do not have people at their disposal to make such assessments. Alternatively, monitoring whether the user mentions the specific risk (e.g., monitoring suicide ideation) while interacting with the service, combined with routine screening via rating scales have been proposed

(Belnap et al., 2015) and have shown good results. In self-help eMental health services, the user data collected to identify whether a risk situation exists , such as the user mentioning the specific risk, can be stored and updated in regular measurement intervals while the user interacts with the service. By doing so, the eMental health service collects a profile of the user's mental health, specific to the risk that is relevant in the eMental health service. By updating the variables stored in this profile regularly, continuous monitoring of risk situations is enabled (e.g., Wolters et al. (2013)).

In the second stage, the decision-making, it is assessed whether any detected risk situation is severe enough to warrant professional intervention. Most routine mental health safety protocols rely on a human professional that combines client-data with protocol guidelines (Tielman et al., 2019). When involving a human professional to assess risks is not feasible, an eMental health service needs a decision-making algorithm that combines the gathered user information and arrives at actionable advice for the client. Applying Tielman et al.'s distinction of the information gathering and decision-making stages to eMental health services, a mental health user profile and an algorithm that detects and assesses the severity of risk situations emerge as two prerequisites of a monitoring module that aims to deliver the technical infrastructure for support on-demand.

As such, eMental health monitoring modules have much in common with Decision Support Systems (DSS). In the past two decades, DSSs have become ubiquitous in medical research and in medical care (Papageorgiou et al., 2013). The main role of a DSS is to support practitioners in decision-making. According to Papageorgiou et al. (2013), common inference engines used in medical DSSs are rules, Bayesian theory, Bayesian Belief networks, heuristics, semantic networks, neural networks, genetic algorithms and other case-specific algorithms, and combinations of these inference mechanisms. The authors identify fuzzy logic to be especially promising for medical diagnosis tasks and reasoning. By allowing differing degrees of truth instead of representing something as either true or false, fuzzy logic is a mathematical representation of vagueness and imprecise, uncertain information (Zadeh et al., 1996). Fuzzy logic can be implemented in medical decision-making using fuzzy cognitive maps (FCMs), a soft computing tool that synergizes fuzzy logic and neural networks.

This chapter presents a technical infrastructure for support on-demand in an eMental health service for mourning older adults and the construction process of one of the infrastructure's components. This component constitutes a decision algorithm to detect risk situations using FCMs. Section 4.2.1 gives an overview of a monitoring module developed for the eMental health service, consisting of two components. It describes the first component, a (mental health) user profile that serves the purpose of information gathering. The remainder of Section 4.2 outlines the design and execution of a research study to develop the second component, the risk detection algorithm based on the FCM methodology. The resulting fuzzy cognitive map and the corresponding algorithm are presented in Section 4.3 and implications from initial scenario experiments and further results from the research study are discussed in Section 4.4, alongside recommendations for future work.

4.2 Method

The monitoring module described in this chapter has been developed in the context of LEAVES, a self-help online intervention designed to soothe the mourning process of older adults who lost their spouse (van Velsen et al., 2020). The service, based on cognitive behavioral therapy, combines psychoeducation about grief, cognitive behavioral exercises for coping with grief and creating a new life without the spouse with activity suggestions to foster self-care and promote resilience. The monitoring module is intended to help the user reflect about their mood and recommends proactively seeking either professional offline support, or reaching out to their support network in times of need when suffering becomes highly disruptive. At the time of writing this chapter, the effectiveness and technology acceptance of the LEAVES intervention are investigated in separate research efforts, as described in detail in Brodbeck et al. (2019).

4.2.1 Background: Overview of the monitoring module

The construction of the first component of the monitoring module, the monitoring user profile, has two phases; 1) an initial risk assessment aimed at identifying whether the service is adequate to meet users' needs at the moment and establishing a baseline for their emotional state; and 2) the continuous assessment of users' emotional state and behaviors to identify risk situations while using the program. Both phases correspond to a questionnaire that measure relevant user parameters for deciding whether recommending the user to seek offline support is appropriate.

The first questionnaire, the initial risk assessment (IRA), is completed during the introduction of the service to the user. Based on the users' responses, the service may display recommendations for using the program that, under certain circumstances, suggest seeking further human care to receive adequate support while using the self-help service. These IRA recommendations were categorized into three urgency levels of human help seeking, which translate into more or less pressingly formulated recommendations. Rules formulated by a team of clinical experts determine the urgency level based on the user's risk assessment responses. For instance, the highest level of urgency is triggered when users' Suicidality responses exceed a specific threshold. This recommendation requests users to get into contact with a regional suicide hotline immediately and strongly advises them to confide in a professional or another person they trust. The medium urgency level is triggered when the loss has occurred in the past twelve months, the loss has been violent (accident, suicide, homicide) and the user has had an inpatient psychological or psychiatric treatment within the last year. The medium urgency level is also tiggered if the loss has occurred longer than twelve months ago, but only if users' Suicidality responses exceed a threshold. This Suicidality threshold is lower than for the first urgency level. The lowest level of urgency is triggered when the loss has occurred more than twelve months before starting to use the service and the user has either experienced a violent loss or the user has had an inpatient psychological or psychiatric treatment within the last year, or both. Additionally, if the loss has occurred less than one month ago, users are given the advice to wait a bit longer before starting the program as the need to deal with practical aspects of the loss and emotional turmoil are common in this period (M. Stroebe & Schut, 1999; M. Stroebe & Schut, 2010). Except in this case, regardless of the urgency level for the recommendations, the user can start using the program. If no IRA recommendation is triggered, the user is encouraged to start using the service instead.

During the second phase, the continuous risk assessment (CRA) is repeatedly administered and a fuzzy cognitive map (FCM) decision algorithm determines each time whether a recommendation to seek offline support is needed. If a risk situation is detected and a recommendation is triggered, the system displays support options, such as a regional telephone hotline and the explicit suggestion to seek human care. Otherwise, the user is encouraged to continue using the service without further recommendations. The CRA is administered for the first time at the end of the user's introduction to the service. Afterwards, it is administered every second week. The first CRA measurement serves as a baseline. Consequently, the FCM decision algorithm is run starting from the second measurement, once the user has used the service for at least two weeks. The IRA and CRA represent entries to the user profile for the information-gathering stage, according to Tielman et al. (2019)'s model. 4.1 shows an overview of the different components of the monitoring module. The recommendation to seek offline support is depicted as *Escalate* and the recommendation to keep using the service without taking any further action is depicted as *Do not escalate*.

Monitoring user profile

As introduced in 4.2.1, the first component of the monitoring module is the user profile, consisting of the initial and the continuous risk assessment questionnaires. The questionnaires are based on a ranked set of 18 user parameters for monitoring identified by a Delphi study (Brandl et al., 2022) involving 16 experts in grief and eHealth.

The set included clinical parameters such as Suicidality and (Complicated) Grief

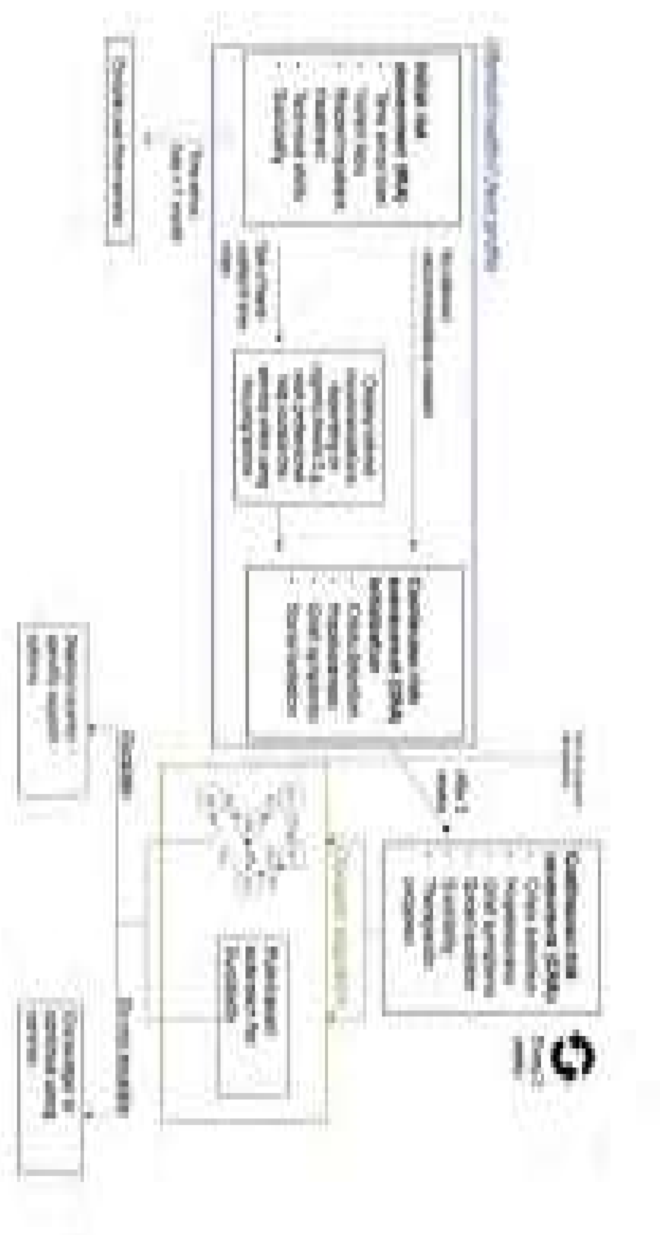


Figure 4.1: Overview of the monitoring module of an online service for older mourners to recommend seeking offline support whenever advisable.

symptoms, behavioral and emotional parameters such as Social isolation and Hopelessness, parameters describing the interaction between the user and the service, such as Unresponsiveness and finally, parameters external to the service, such as the estimation of the user's situation from the perspective of a close one. The ten highest-ranking parameters from the study were scrutinized for a) their suitability to detect changes in users' situation in a bi-weekly measurement interval and b) for reliable assessment that uses a minimal number of questions to reduce cognitive demands on the user. A selection of five parameters from the Delphi research was extended with five parameters for the initial risk assessment and one parameter for the continuous risk assessment. 4.1 lists the monitoring parameters assessed in the initial risk assessment (IRA) and in the continuous risk assessment (CRA).

For each CRA parameter, except for Suicidality, two questions were designed. The parameters Crisis Detection, Hopelessness, Grief Symptoms, and Social Isolation are completed on a 4-point Likert scale (Not at all, Several days, More than half of the days, and (Nearly) every day) measuring how frequent they disrupted users' daily activities in the last two weeks. The design of the response options was based on the PHQ-9 that can be used for diagnosing Major Depression Disorder according to the DSM-5 (Kroenke et al., 2001). Therapeutic progress is measured on a 4-point Likert scale ranging from Strongly disagree to Strongly agree. Suicidality follows a two-step approach. The first item In the past two weeks, I have considered committing suicide serves as a filter. The user can either endorse the statement with Yes, or respond No. Only if the user replied Yes, the remainder of the Suicidality items is presented to the user. The four conditional Suicidality items were adapted from the Scale for Suicide Ideation (SSI) (Beck et al., 1979) and assess the extent to which the user has explicit suicide plans. The complete initial risk assessment and continuous risk assessment questionnaires are included in the Appendix.

4

4.2.2 Background: Fuzzy cognitive maps

The second component of the LEAVES monitoring module is a risk detection algorithm based on fuzzy cognitive maps (FCMs). FCMs were first proposed by Kosko (1986) as an extension of cognitive maps to model dynamical systems. FCMs are directed graphs that use fuzzy logic to represent causal relations between graph nodes using weighted edges. 4.2 shows a simple FCM.

In this simple FCM, the causal influence of, for example, concept C1 on concept C2 is indicated with weight W12. There are three types of weights. Weights can either be equal to zero, meaning that that there is no causality between two concepts. Weights larger than zero indicate causal increase (i.e. C2 increases as C1 increases and C2 decreases as C1 decreases if W12 is larger than zero). Weights smaller than zero indicate causal decrease (i.e. C2 increases as C1 decreases and C2 decreases as C1 increases if W12 is smaller than zero).

Table 4.1: Summary of the monitoring parameters in the questionnaires for the initial risk assessment (IRA) and the continuous risk assessment (CRA).

Monitoring parameter	Definition
Time since loss (IRA)	The number of months since the loss. Measured in four intervals: Less than a month, 1-6 months, 7-12 months and more than 12 months ago.
Violent loss (IRA)	Violent losses considered in the monitoring are accidents, suicide, and homicide.
Recent inpatient treatment (IRA)	Whether or not the user has undergone, in the previous year, a psychiatric inpatient treatment.
Technical skills (IRA)	Lack of digital literacy can impact the users' motivation to use the service. Basic computer skills are sufficient to use the service.
Crisis detection (IRA & CRA)	The extent to which users have experienced a crisis in the past two weeks, meaning that they were impaired in their daily activities due to emotional distress or felt that they cannot cope alone.
Hopelessness (IRA & CRA)	The extent to which users feel negatively about their future and feel that they cannot do anything about it.
Grief symptoms (IRA & CRA)	Intense feelings of grief, such as being stunned or sad due to emotional distress specifically related to the loss, to the extent that users have been impaired in their daily functioning.
Suicidality (IRA & CRA)	The extent to which users have specific plans for a suicide attempt.
Social isolation (IRA & CRA)	The extent to which users feel burdensome towards their social contact and their actual withdrawal behavior.
Therapeutic progress (CRA)	The users' perception of the extent to which they are making progress in processing their loss.

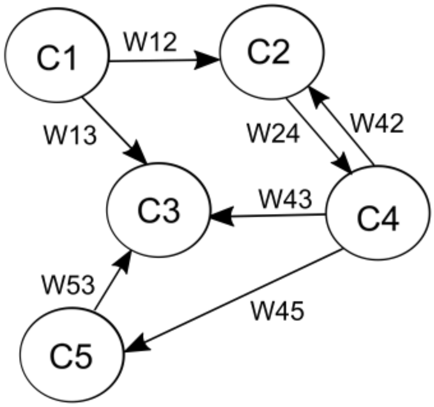


Figure 4.2: A simple fuzzy cognitive map.

Commonly, the relation between two concepts takes a value in the interval [-1, 1]. -1 corresponds to the strongest causal decrease and 1 to the strongest causal increase. The other values express different levels of influence. The main objective of building a FCM is to predict an outcome state of the FCM concepts by letting them interact with each other according to their weights until their values converge. The concept values at convergence represent the final prediction of the FCM concepts. FCMs have been developed in the context of mental health before. For example, Papageorgiou et al. (2013) developed a FCM-based decision support system (DSS) to diagnose depression in older adults and conclude that one of the strongest points of the FCM-based DSS is that it provides insight into feedback loops between symptoms by making them explicit. For example, one of the symptom concepts identified in the authors' FCM-based DSS for geriatric depression was Fatigue. After constructing the FCM, all connections towards Fatigue as well as all ways in which Fatigue influences other symptoms in the FCM were known, according to the experts who had helped construct the FCM. That way, the FCM was able to make explicit how an increase in Fatigue can bring about an increase or decrease in other symptoms in the model, such as Depressive mood or Indecisiveness. The configuration of causal relations in FCMs draws on expert knowledge and thereby reflects their reasoning, making them, ultimately, understandable by other human observers, such as healthcare professionals.

4.2.3 Procedure for developing the FCM decision algorithm

To configure the second component of the monitoring module for the LEAVES intervention, an automatic FCM decision algorithm, interviews with experts in grief were conducted. The interview protocol followed a standard procedure for configuring FCMs with the help of domain experts, described by, for example, Papageorgiou (2011b). The procedure has three main steps. During step 1, key concepts to be included in the FCM are identified. In step 2, the causal relations between the identified concepts are determined. And in step 3, the strength of the relations between concepts is estimated.

Since user parameters for monitoring risk situations in an online grief intervention were available based on existing research (Brandl et al., 2022), we used them as the FCM's symptom concepts, corresponding to step 1 in the procedure for constructing FCMs. The monitoring parameters in the initial risk assessment and continuous risk assessment make up the symptom concepts in the current FCM. In addition to the symptom concepts, there are two outcomes modelled in the FCM, Escalate and Do not escalate. They correspond to the two recommendations that the system can give based on the algorithm's calculations. Escalate corresponds to the system's recommendation to seek offline support and Do not escalate corresponds to recommendation to keep using the service as is. This FCM structure is referred to as a competitive fuzzy cognitive map (CFCM) because the two outcomes compete with each other, only one of the two is chosen in the end (Stylios et al., 2008). Figure 4.5 shows the structure of the FCM monitoring decision algorithm developed in this study.

In step 2, experts determine the causal relation between any two concepts in the FCM as either positive, negative or neutral. The latter case is equivalent to declaring that there is no influence between the two concepts in question. For step 3, commonly, a linguistic variable called Influence is declared to represent the strength of relations between concepts (i.e., weights) in fuzzy terms, such as low, medium and high. The set T of qualitative terms that experts in this study used to describe the influence among FCM concepts is: $T(Influence) = Veryverylow, Verylow, Low, Medium, High, Veryhigh, Veryveryhigh$. All qualitative terms occurred either in combination with a positive sign to represent a positive causal relation, or in combination with a negative sign to represent a negative causal relation (e.g., -Very high or +Medium). After these initial three steps to obtain qualitative weights based on expert knowledge, the weights provided by individual experts are aggregated. The process of weight aggregation is described in more detail in section 4.3.2.

Experts recruitment

For the interviews, eight experts were recruited via the researchers' professional networks and via snowball sampling. As inclusion criteria, experts were required to have experience with treating or coaching bereaved adults, preferably older adults. Consequently, psychotherapists as well as grief coaches were the primary focus of the recruitment. If a candidate expert expressed interest in the study, the researchers provided them with more detailed information about the research, including the informed consent and a digital copy of the two monitoring questionnaires (IRA and CRA). All participants provided written informed consent. As a token of gratitude, everyone who participated in the study received a gift card amounting to 25€.

Data collection

The interview protocol was supported by an online, interactive one-on-one session conducted and recorded via Microsoft Teams. Specifically, the interview was designed to ask experts to, based on their professional experience, determine the predictive value of each FCM concept, such as Social isolation, in the assessment of the users' risk situation. In other words, the goal was determining the weights of the FCM concepts in the decision algorithm that identifies risk and triggers recommendation messages to users when further support is needed. Considering the demanding nature of the interview's core task, before the interview, experts were asked to read a brief description of the monitoring module and to read through the monitoring questionnaire items. During the interview, experts were first asked to describe their professional experience with older mourners and to explain to what extent they observe differences between mourners that generally cope well with their loss and those who do not.

Then, experts were asked to attribute qualitative values (i.e., weights) from the pre-determined Influence set T ranging from Very, very low to Very, very high to each relation between symptom and outcome concepts and among symptom concepts (see Figure 4.3). Higher weights were indicative of higher risk to develop severe mental health symptoms (i.e., the parameter had a higher predictive value) and lower values were indicative of lower risk (lower predictive value). This part of the one-on-one interviews was conducted with the support of the online collaboration tool Mural (Mural.co., 2022). Mural is comparable to a whiteboard where different content such as text and pictures can be pinned and moved around freely. Via two exercises, experts attributed weights to the relations between the FCM concepts. The first exercise focused on the weights from the eight symptom concepts to the two algorithm outcomes Escalate and Do not escalate. The second exercise solely focused on relations between symptom concepts. Experts were invited to use coloured sticky notes to indicate the relation type between two concepts (e.g., Hopelessness and Social isolation), which could be positive, negative, or neutral. Our experts attributed weights to all relations between concepts in the first exercise (symptoms to FCM outcomes). Due to time restrictions, in the second exercise, they were asked to assign weights to only the most important relations (symptom to other symptom concepts), according to their professional experience. A snapshot of the Mural board, depicting the visualization of the first exercise, is shown in Figure 4.3. After experts finished both weighing exercises, the final interview question was to describe how difficult the exercises had been for them. A pilot session was conducted with one psychotherapist. As a result of the pilot, Suicidality was excluded from the FCM weighting exercises. The pilot participant explained that considering the serious consequences of active suicidal ideation, Suicidality already has relatively clear thresholds for advising a user to seek immediate offline support. Indeed, cut-off scores have readily been used for suicide risk assessment scales to identify people at (immediate) risk for attempting suicide (McCall et al., 2021; Sokero et al., 2003). As a consequence, in the final monitoring module, Suicidality is evaluated using decision rules prior to running the FCM. To evaluate Suicidality during the continuous risk assessment, the same escalation threshold is applied as for determining a recommendation of the highest urgency in the initial risk assessment.

Data analyses

Interviews were transcribed and analysed via an inductive coding scheme, developed and applied by one researcher and then verified by a second researcher. Any discrepancies were discussed until agreement was reached. Two examples from the coding scheme are:

- Q2.1CDToOutcome: Any explanations and remarks participants make about the relation between the symptom concept Crisis Detection to either of the two outcome concepts (Do not) Escalate.
- FCMModel: Any explanations or remarks participants make about using fuzzy cognitive mapping and the proposed FCM in particular, including comments about the choice of (symptom) concepts, model limitations and suggestions for improving the FCM.

FCMpy (Mkhitaryan et al., 2022), a library for the Python programming language, was used to combine the qualitative weights (ranging from Very, very low to Very, very high) provided by the experts in this study and to transform them into numeric FCM weights. As a preliminary evaluation of the resulting FCM decision algorithm, the obtained weights were tested using four fictitious user scenarios. Scenario testing has a long tradition in FCM research and is regarded as one of the most valuable applications of FCMs (Olazabal & Reckien, 2015). JFCM (Franciscis, 2014), a library for the Java programming language, was used to simulate the behavior of the FCM based on the four fictitious scenarios. The scenario experiments and their results are described in more detail in section 4.3.3.

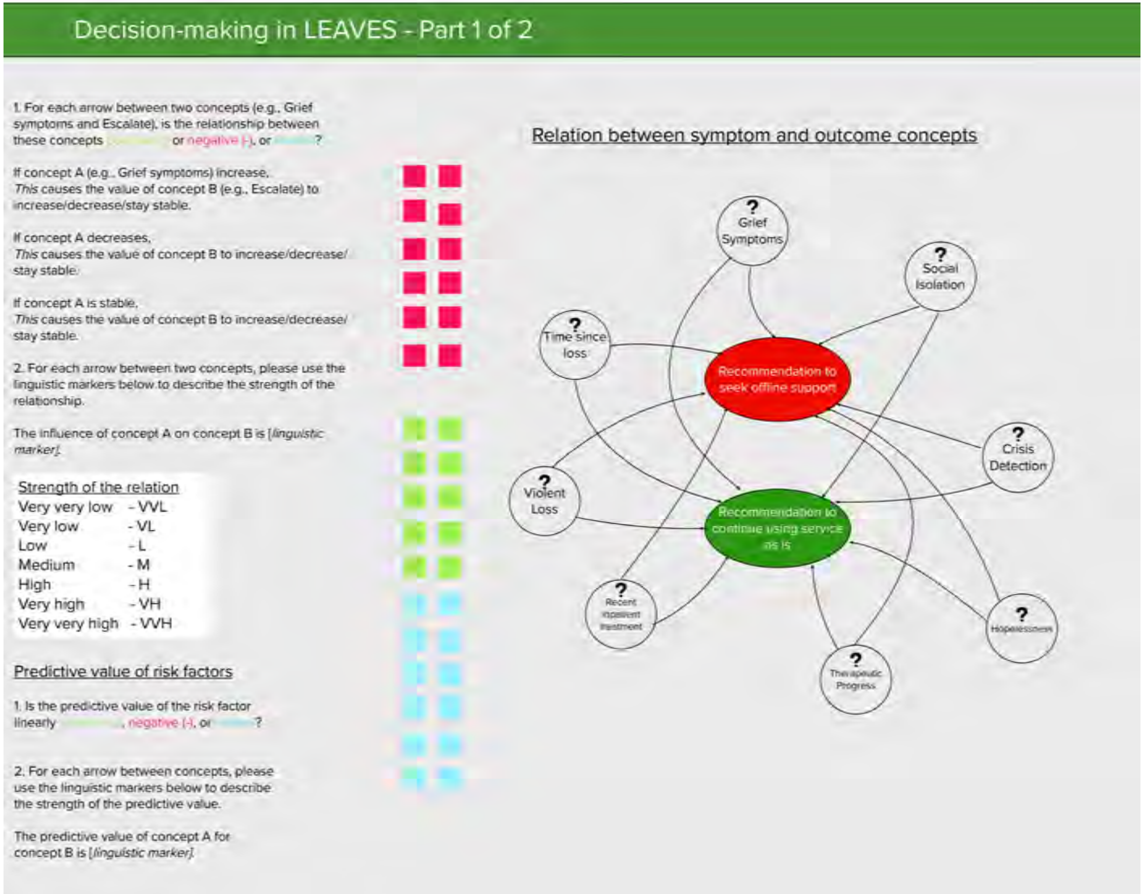


Figure 4.3: Screenshot of the study materials in Mural.

4

4.3 Results

4.3.1 Expert demographics

In total, eight clinical psychologists participated in the interview study, excluding the pilot session. Three experts were male. Their mean age was 46.5 years, with a standard deviation of 12.76 years. Table 4.2 summarizes experts' demographics. All experts had experience as grief coaches and most of them worked with older adults, with 12.5 years of experience on average and a standard deviation of 8.37 years.

4.3.2 Aggregated FCM weights

Each interview yielded an individual expert's account of the causal relations between symptom (e.g., Hopelessness) and outcome concepts (Escalate and Do not escalate) in the FCM in qualitative terms (e.g., Low, Medium, High). To develop a FCM, the process of aggregating the input of multiple experts involves four steps: 1) defining membership functions for the qualitative values, 2) applying a fuzzy implication rule, 3) combining the membership functions of individual experts and 4) defuzzifying the aggregated membership functions to derive a numerical weight for each relation between symptoms and outcomes in the FCM (Mkhitaryan et al., 2022). The theoretical and mathematical foundation of these steps are explained in more detail elsewhere (Papageorgiou, 2013).

In the first step, triangular membership functions were defined for the qualitative terms used by the clinicians in this study. Triangular membership functions are used in most applications (Frias et al., 2017; Papageorgiou, 2011b). Figure 4.4 shows the membership functions for the fourteen employed qualitative terms in this study: $+/-very, verylow; +/-verylow; +/-low; +/-medium; +/-high; +/-veryhigh; +/-very, veryhigh$. The na membership depicts missing values in individual expert's accounts for a specific relationship. Not all experts attributed a qualitative value to every relation because we asked them to only rate those relation between symptom concepts (e.g., Hopelessness and Social isolation) that they regarded as most important.

Likewise, we chose commonly applied methods for steps two to four; we applied Mamdani's fuzzy implication rule to determine the extent to which each membership function for a specific causal relation was activated (step 2) (Giabbanelli, 2014; Nandi, 2012), the algebraic SUM aggregation operation (step 3) (Giabbanelli, 2014; Papageorgiou, 2011a) for combining the qualitative weights provided by individual experts, and center of gravity (COG) defuzzification to derive a numeric weight for each causal relationship (step 4) (Papageorgiou et al., 2013; Stylios & Groumpos, 2004). In the resulting weight matrix, due to the competitive nature of this type of FCM which favours one outcome over the

Table 4.2: Expert demographics

	Age	Gender	Country	Expertise	Years of experience
Expert 1	29	Female	Switzerland	Psychotherapist (in training), older adults, grief, depression	4
Expert 2	33	Male	The Netherlands	Clinical psychologist, psychotherapist, Researcher, Grief, Depression	6
Expert 3	41	Female	The Netherlands	Grief coach, lecturer	7
Expert 4	56	Female	The Netherlands	Grief coach, social psychotherapist (in training), older adults	8
Expert 5	56	Female	The Netherlands	Grief coach, psychotherapist	10
Expert 6	39	Female	Slovenia	Psychotherapist, researcher, suicide	15
Expert 7	53	Male	Switzerland	Clinical psychologist, psychotherapist, lecturer	25
Expert 8	65	Male	The Netherlands	Minister, grief coach	25

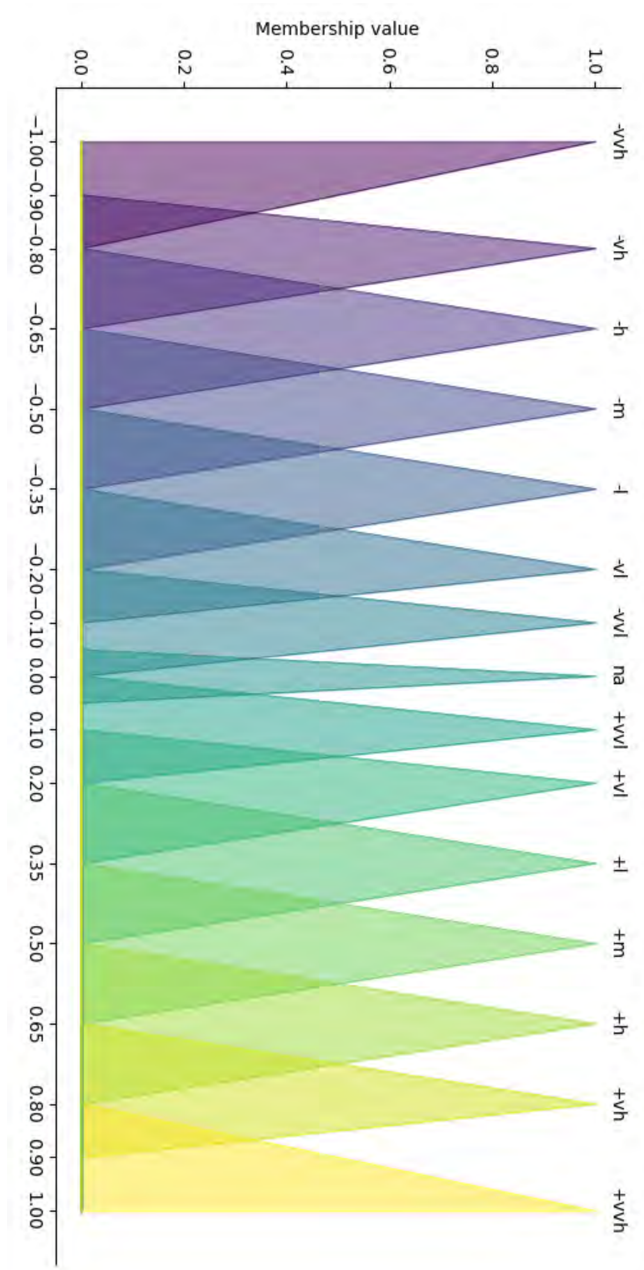
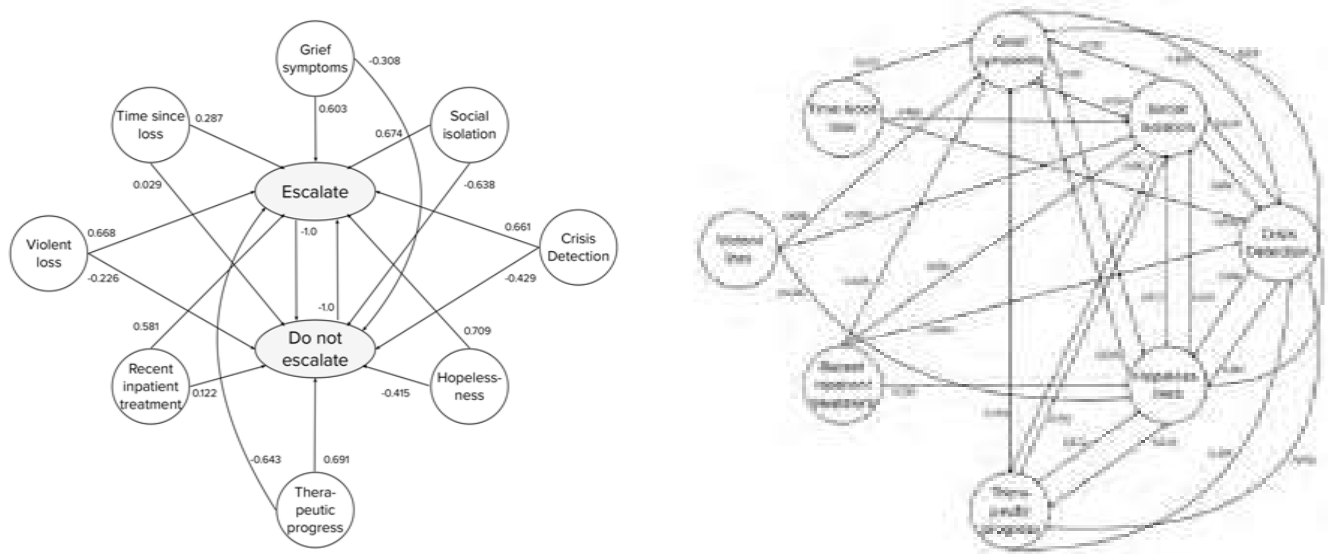


Figure 4.4: Triangular membership functions of the qualitative terms used by experts to describe the causal relationships in the fuzzy cognitive map between symptoms and outcomes.

other, the relations between the two outcomes Escalate and Do not Escalate were set to -1. At the same time, the outcomes do not exert any influence on the symptom concepts in the FCM. Hence, all outgoing weights from Escalate and Do not Escalate were set to 0. Incoming weights from other concepts to Violent loss, Recent inpatient treatment, and Time since loss were also set to 0. There is no logical influence from any of the other symptom concepts on these concepts. The aggregated numeric FCM weights are summarized in the matrix W on the next page and displayed in Figure 4.5. From the combined weights, a mutually strengthening symptom cluster arises between Crisis detection, Grief symptoms, Hopelessness, and Social isolation as indicated by their high and positive relations to each other. They also have strong positive weights towards the outcome Escalate. Therapeutic progress has a counterbalancing effect on them, indicated by its strong negative weights towards Crisis detection, Grief symptoms, Hopelessness, and Social isolation. The three concepts that are included from the initial risk assessment questionnaire, Violent loss, Recent inpatient treatment, and Time since loss all have positive relationships to the Escalate outcome. It is notable that experts did not always agree on the type of the relation, i.e. whether the sign of a weight between two symptom concepts or a symptom and an outcome concept was positive or negative. Some experts conceptualized the relation between the two outcome concepts Escalate and Do not escalate as mutually exclusive, while others did not. While these conceptualization differences did not affect the applied weight aggregation procedures in a mathematical sense (i.e., their principles still applied), this observation shows that FCMs can be used in multiple ways to model the decision-making for detecting risk situations in a grief eMental health service.

$$
W = \begin{array}{r} \\ \text{Crisis Detection (CD)} \\ \text{Hopelessness (H)} \\ \text{Grief Symptoms (GS)} \\ \text{Therapeutic Progress (TP)} \\ \text{Social Isolation (SI)} \\ \text{Recent Inpatient Treatment (RIP)} \\ \text{Violent Loss (VL)} \\ \text{Time since Loss (TSL)} \\ \text{Escalate} \\ \text{Do not Escalate} \end{array}
\begin{array}{c}
\begin{array}{cccccccccc} \text{CD} & \text{H} & \text{GS} & \text{TP} & \text{SI} & \text{RIP} & \text{VL} & \text{TSL} & \text{Escalate} & \text{Do not escalate} \end{array} \\
\left[\begin{array}{cccccccccc}
 & 0.766 & 0.631 & -0.65 & 0.624 & & & & 0.661 & -0.429 \\
0.816 & & 0.636 & -0.622 & 0.527 & & & & 0.709 & -0.415 \\
0.839 & 0.701 & & & 0.35 & & & & 0.603 & -0.308 \\
-0.705 & -0.672 & -0.704 & & -0.777 & & & & -0.643 & 0.691 \\
0.816 & 0.773 & 0.732 & -0.912 & & & & & 0.674 & -0.638 \\
0.65 & 0.727 & 0.601 & & 0.601 & & & & 0.581 & 0.122 \\
 & 0.65 & 0.669 & & 0.65 & & & & 0.668 & -0.226 \\
-0.65 & & -0.575 & & -0.168 & & & & 0.287 & 0.029 \\
 & & & & & & & & & -1 \\
 & & & & & & & & -1 &
\end{array}\right]
\end{array}
$$



(a) (b)

Figure 4.5: Visualization of the aggregated weights from symptom to outcome concepts (a) and (b) between sympyom concepts.

4.3.3 Scenario experiments with aggregated weights

As a preliminary evaluation of the resulting FCM algorithm, the obtained weights were tested using four fictitious scenarios. As an initial assessment of the utility of the obtained FCM weights, a clinical expert generated four scenarios, each representing the responses of a fictitious user case. Scenario 1 focused on a user that clearly needs support; scenario 2 on a user that can clearly continue using the service; scenario 3 on a user who should seek support, but it does not show obviously from the mourner's questionnaire responses; in scenario 4, the fictitious user should get the advice to continue using the service, but less obviously than in scenario 2.

Before the FCM could be simulated based on the fictitious cases, the monitoring questionnaire responses had to be transformed into numeric values so that they could be used as inputs for the FCM. For each monitoring parameter, a total score was calculated and mapped on the interval $[-1, 1]$. The continuous risk assessment questionnaire inquires about the frequency of symptoms in the past two weeks. Due to the high prevalence of positive weights towards the outcome Escalate, we decided to model the absence of symptoms or very infrequent symptoms as supportive for the outcome Do not escalate, hence we mapped the lowest response options to a negative input value. Violent loss and Recent inpatient treatment were mapped on either -1 or 1, depending on whether the user endorsed the question. The questionnaire total scores per dimension, corresponding to the four fictitious scenarios, are summarized in Table 4.3. Since Suicidality was removed from the FCM based on the feedback that we received during the pilot, no Suicidality scores were included in the scenarios. Instead, the final monitoring algorithm has a rule-based extension that checks users' suicidality responses for risk situations outside the FCM.

Table 4.3: Summary of fictitious user response scenarios using the aggregated expert-based FCM weights.

Monitoring item	**Scenario 1**	**Scenario 2**	**Scenario 3**	**Scenario 4**
Grief Symptoms	5	1	5	4
Social Isolation	3	1	2	2
Crisis Detection	6	2	5	4
Hopelessness	6	1	2	3
Therapeutic Progress	1	5	2	2
Violent Loss	0	0	0	1
Recent Inpatient Treatment	1	0	0	0
Time Since Loss	1	3	3	1

Scenarios were evaluated using equation 4.1, Kosko's modified inference equation for FCMs (Papageorgiou, 2011b), which iteratively calculates the values of the concepts in the FCM based on an initial baseline until the model converges.

In competitive FCMs (CFCMs), the outcome concept with the highest value is regarded as the model's outcome. The outcome concepts in this case were Escalate and Do not Escalate.

$$A_i^{(k+1)} = f(A_i^{(k)} + \sum_{j \neq i;\ j=1)}^{N} A_j^{(k)} w_{ji}) \tag{4.1}$$

A represents the initial state vector of the FCM. A is a $Nx1$ matrix where each of the N rows contains the initial value of a concept C in the FCM. In this study, the initial concept values are the transformed, fictitious user responses we use to construct the scenarios. $A_i^{(k+1)}$ is the value of concept C_i at timestep $k+1$; $A_i^{(k)}$ is the value of concept C_i at timestep k; $A_j^{(k)}$ is the value of concept C_j at time step k; and w_{ji} is the influence of concept C_j on concept C_i, expressed as a weight. f is a threshold function that transforms the content of the function. For all but the three risk assessment parameters, we used one of the most common threshold functions used in FCMs (Stylios & Groumpos, 2004), the hyperbolic tangent function that transforms the content into the interval $[-1, 1]$:

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{4.2}$$

To prevent the values of the three risk assessment parameters, Violent loss, Recent inpatient treatment, and Time since loss to change between simulation iterations, we chose the linear activation function for Time since loss instead and a bivalent activation function for the two boolean parameters Violent loss and Recent inpatient treatment.

$$Linear f(x) = x \tag{4.3}$$

$$Bivalent f(x) = \begin{cases} -1, & if x \leq 0, \\ 1, & if x > 0, \end{cases} \tag{4.4}$$

For the first scenario which represented a user that clearly needs support outside the service, the algorithm's outcome was Escalate. The algorithm's outcome for the second scenario which represented a user that could clearly continue using the service by themselves was Do not Escalate. For the third scenario, a fictitious user who should seek support, but it shows less obviously than in scenario 1 in their questionnaire responses, the algorithm's decision was

Do not Escalate. And for the final scenario, a fictitious user who can continue using the service, but it shows less obviously than in scenario 2, the algorithm advised Escalate.

4.3.4 FCM model limitations and suggestions for improvement

Next to configuring the weights of the FCM using the input of clinical experts, the interview data was analyzed with regard to the participants' appraisal of the FCM as a decision-making algorithm for detecting risk situations in a grief eMental health service. While filling in the FCM weights, the experts participating in this study identified limitations when using the FCM to determine whether or not someone using the online grief self-help service should seek offline support. They also provided suggestions to improve the FCM decision-making algorithm in the future. Most experts struggled with modelling the relation between Time since loss and the two outcomes of the FCM, as well as with modelling its influence on other symptom concepts linearly. They explained that the general expectation regarding Time since loss is that "time heals wounds". The more time has passed, the better one usually copes with a loss. However, if that is not the case, if the loss has occurred long ago and the person still suffers intensely from the loss, then this is an important indicator that someone is stuck in their grief process. In addition, it is not uncommon to see a rise in grief symptoms between the first and the second year after the loss, while generally, grief symptoms decrease as time passes by.

> "Especially if the loss is already quite a long time ago and they still have quite strong grieving symptoms, then I think 'Yes, now it is time to seek professional help'. To me, this would be one of the most important predictors or indicators for professional help." (Expert 7)

Another limitation of the current FCM is that its decisions are based on one measurement point instead of the history of a user's measurements. The experts in this study explained that to decide whether or not someone should seek offline support, how long symptoms have been elevated is an important decision criterion.

> "Regarding grief symptoms, if it gets worse several times, then I would look at that as an alarm bell. But if, for example, from the first [measurement] to the second they get worse I would perhaps not yet say 'Please seek professional help'. The self-help modules are very demanding and activate grief. From that point of view, it is also clear to me that they can temporarily deteriorate [the symptoms]." (Expert 1)

A discussion with expert 3 revealed that the influence of the three risk concepts Violent loss, Recent inpatient treatment and Time since loss is challenging to model using the FCM. In the FCM, relations between symptom concepts can strengthen or weaken the influence of any single concept on the outcome. However, there is no logical influence from the other symptom concepts in the FCM on these three risk concepts. For example, however high someone's grief symptoms are, there is no logical influence on how long ago the loss has occurred. Therefore, the risk concepts do have outgoing influences, but no incoming influences, increasing their overall influence on the model's outcome since the risk concepts cannot be relativized within the FCM. Several experts suggested how the FCM could be extended to better reflect how they make clinical decisions in practice. First, the current selection of FCM concepts overemphasizes a mourner's negative emotions and symptoms, while positive experiences such as reminiscing fond memories are also part of the grieving process.

> "You're actually only checking the negative sides of the person's experience; which is ok, of course, because we are worried and we want to know [about them], but it's not only negative. You remember the positive sides and, you know, there is laughter. Maybe the model indicates that it should be bad all the time, but that is not the case." (Expert 6)

In addition, individual factors such as the mourner's resilience and the significance of the lost relationship for their personal identity play a large role in how well one copes with the loss of a partner. When loss experiences accumulate, which is not uncommon in later life, a person's resilience is compromised. The cumulation of loss experiences is worthwhile to take into account, especially for older mourners:

> "There are unprocessed previous losses and that's what I often encounter with older people. [...] They are from a different generation. Regarding loss experiences, things have often remained unmentioned and those things are revived by the most recent loss, in this case, the loss of a partner. This resonates quite a bit." (Expert 5)

Other suggestions for additional symptom concepts include assessing how traumatic the loss experience has been for the mourner, in contrast to focusing on whether or not the loss has been violent and to complement Hopelessness with assessing the mourner's feeling of entrapment. A final suggestion was to incorporate different urgency levels when advising someone to seek offline support. Depending on the user's situation, it may be more or less urgent that they seek immediate support and this can be reflected in the recommendation the program gives.

4.4 Discussion

In this chapter, we present a monitoring module for an online self-help service for older mourners that recommends seeking offline (professional) support when the user's suffering becomes highly disruptive for their everyday life. We present a user profile specific to grief and we configure a fuzzy cognitive map (FCM) decision algorithm. The user profile consists of user parameters relevant for a mental eHealth service targeted at mourning older adults and a measurement protocol to monitor their situation regularly, reflecting Tielman et al. (2019)'s information gathering stage in their two-stage model for detecting risk situations in eMental health. The FCM decision algorithm recommends seeking offline support once a risk has been detected in the user's situation. The FCM represents Tielman et al.'s second phase, the decision-making. The current study demonstrates the configuration of the FCM decision algorithm based on expert-knowledge and investigates their performance in four fictitious user scenarios. The developed FCM consists of eight monitoring concepts, including Grief symptoms, Hopelessness, Social isolation, Therapeutic progress, Crisis detection, Violent loss, Recent inpatient treatment and Time since loss. The two outcomes of the FCM are either to Escalate, meaning that the service recommends the user to seek offline support, or Do not escalate, in which case the system encourages the user to continue using the service as is.

4.4.1 Preliminary FCM algorithm evaluation using scenarios

The four scenarios explored in this paper represent fictitious users and their responses to the monitoring questionnaires, formulated by a clinical expert in grief. The results of the tests show that while the current FCM algorithm distinguishes well between unambiguous cases, it can confuse borderline cases. A notable difference between the unambiguous and the borderline cases in this chapter are users' fictitious Time since loss responses, as well as whether or not the fictitious user has suffered a violent loss or not, meaning that they lost their spouse due to an accident, suicide, or another form of violent death. The borderline scenarios exemplify the difficulties of modelling the influence of Time since loss linearly as the current FCM does. As the experts explained in the interviews, Time since loss is an important indicator, but its influence on the model's outcome depends not only on its own value, but also on the value of other concepts in the model. According to the experts in this study, "time is generally expected to heal wounds", however, when this is not the case, it becomes an important decision criterion to decide whether a mourner should consider seeking offline (professional) support. It has a strong positive relation to the outcome Escalate and a strong negative relation to the outcome Do not escalate if and only if the rest of the symptom concepts indicate that the user is suffering emotionally. Otherwise, the signs of its weights are vice versa and the strength of the relation is lower. Unfortunately, the current FCM does not capture this information because it does not allow for conditional weight setting. It is likely that this led to the incorrect classification of the third scenario as no need to recommend the user to seek support, even though the scenario depicts a user who feels severely impaired in their daily life despite the loss having occurred more than 12 months ago. The second borderline case involves a fictitious user who has suffered a Violent loss. In the current weight matrix, Violent loss has a very high positive relation, not only to the outcome Escalate, but also to other parameters (e.g., Hopelessness and Grief symptoms) that have in turn strong and positive relationships to Escalate. In essence, it acts as a magnifier in the model for other symptoms of suffering, such as Grief symptoms. This in combination with the difficulties of modelling the conditional influence of Time since loss in the current FCM caused the algorithm to be led astray. The FCM decision algorithm is currently further evaluated as part of ongoing evaluation studies (Brodbeck et al., 2022) of the online grief service in which it is integrated and it will be interesting to see how it performs in a non-fictitious setting involving older users of the grief service.

4.4.2 Future FCM algorithm improvements

The above analysis of how the FCM algorithm performed in the scenario experiments exemplifies a strength of using FCMs for detecting risk situations in mental eHealth: its outcomes can be traced back and are thereby explainable. Transparency of decision making, also referred to as scrutability, is of utmost importance for decision support in healthcare where stakes are potentially high and expert users of decision support systems need to accept and rely on the system's advice. Scrutability is a key advantage of fuzzy systems (Balog et al., 2019; Nauck & Kruse, 1999; Papageorgiou et al., 2013) and it facilitates identifying approaches to improve the current FCM. For instance, to improve the accuracy of the current FCM in detecting risk situations in borderline cases, the conditional influence of Time since loss should be represented in the decision-making. This could be realized by introducing a rule-based extension that adjusts the weight from being positive towards the outcome Escalate to negative if the user's suffering, as measured by other symptom concepts in the model such as Hopelessness and Grief symptoms, stays below a to-be-determined threshold. This approach is similar to the approach taken by Papageorgiou et al. (2013) to account for the differing influences of the concepts psychomotor status and sleep disturbance in detecting depression in an elderly population, depending on the values that these concepts have. For example, in their FCM, psychomotor status has one of two values: psychomotor agitation or psychomotor retardation. Depending on which of the two values is activated, the outgoing weights from the concept psychomotor status to other concepts in the FCM change.

A second approach to improving the current FCM is to scrutinize the weight aggregation process. In the current study, not all relations between symptoms were rated by all experts because their task was to rate the relations that they considered most important. Consequently, some relations between symptoms were rated by the majority of the experts, while others were rated by one or two experts. As a result, individual expert ratings influenced the combined weight matrix to differing degrees. While there is no golden standard for managing expert disagreement during the weight aggregation process, some (Olazabal & Reckien, 2015; Stylios & Groumpos, 2004) argue that the effects of unequal numbers of ratings for each relation between symptoms as well as disagreement between experts can be mitigated by any rule-based or mathematical approach that is reasoned properly. For instance, Reckien et al. (2013) suggest a rule-based approach for aggregating highly divergent weights by excluding weights that diverge too much from the arithmetic mean. Another approach could be to eliminate weights from the aggregation that less than a pre-determined number of experts have rated. Stylios and Groumpos (2004) apply a credibility factor for regulating how much influence the weights provided by individual experts have in the combined weight matrix. Each expert starts with a credibility factor of 1 with which the suggested weights of the expert are multiplied. The expert is iteratively “penalized" if their suggested weights divert too much from the weights suggested by the rest of the sample.

A final approach to improving the current FCM heeds the recommendation of the experts that participated in this study and involves reconsidering the FCM architecture and introducing more symptom concepts to model the user’s resilience to outweigh the current focus on suffering in the FCM. Another way of emphasizing positivity instead of focusing on suffering is to enhance the impact of Therapeutic progress, for example by adding a rule that counteracts an escalation under the condition that the user feels that they are making progress, despite their suffering. In such a situation, the suffering measured in the continuous risk assessment may be less reason for concern and part of a healthy mourning process. Bourgani et al. (2013) furthermore stress the importance of including temporal information in medical decision making. The current monitoring module has no means to take the history of the user’s monitoring responses into account, while the experts in this study pointed out that this information is important to consider when deciding whether or not a user should seek offline (professional) support. Since including temporal information in FCMs is not straightforward (Olazabal & Reckien, 2015) information about the duration of the user’s suffering could be added via a rule-based extension that takes the user’s measurement history into account and formulates rules about when it indicates a risk situation.

The above approaches consider limitations of the current FCM and recommendations for future work to improve its capacity to reliably detect risk situations in an eMental health service to aid older mourners. Tielman et al. (2019) consider another factor that determines the effectiveness of monitoring modules in mental eHealth. Systems that ultimately rely on the user's initiative to seek professional support depend on their motivation to follow-up on the system's recommendation to seek offline support. The authors argue that depending on a) the user's initial stance on involving professional support (negative, doubtful, positive) and b) the severity of the risk situation (severe, negative, doubtful), the system should adopt different strategies to persuade the user. For example, a user that is doubtful about involving professional support and whose situation is severe needs to be persuaded. On the contrary, a user that is generally positive towards involving a professional and whose situation is severe requires facilitation of the self-referral process. This approach to modelling the user's motivation to accept and act on the system's recommendations is in line with a suggestion from an expert that participated in this study to design different urgency levels for recommendations, depending on the severity of the situation among other factors.

4.4.3 Limitations

The current study has limitations. First, we gather semi-quantitative data via individual interactive sessions with eight clinical experts. Since our approach to aggregating weights is quantitative in nature, the size of the expert sample becomes a point of discussion. There is, however, no standard to determine a sample size for fuzzy cognitive mapping studies. Studies that determine the very concepts of the FCM alongside their weights can use the extent to which the FCM concepts are saturated as indicator that the usefulness of including more experts has been exhausted. We chose the concepts of the FCM based on a previous study and hence, our concepts were fixed. Olazabal and Reckien (2015) acknowledge that sample sizes differ tremendously between individual studies, ranging from only a few participants (three in Papageorgiou et al. (2013)) to studies aggregating up to 376 individual maps (Reckien, 2014). The authors stress that developing meaningful and usable FCMs is a matter of careful selection of experts rather than sample size. The current study design, individual one-on-one interviews can be regarded as a second limitation. While fuzzy cognitive mapping studies use a wide variety of individual as well as group elicitation methods (Olazabal & Reckien, 2015), the experts participating in this study may have profited from a group setting to discuss the weights between FCM concepts and to exchange professional experiences. Experts indicated that determining exact weights for the relations between concepts was challenging. Discussing with peers could have reduced the perceived challenge of the task and limited the extent to which experts disagreed about weights, paving the way to a more coherent distribution of weights. Another argument in favour of a group design is that individual FCMs are inherently subjective with regard to the expert's professional experiences. They are also affected by recency biases, meaning that recent professional experiences may be more salient and therefore, appear more strongly related to FCM outcomes. Future group discussions could mitigate some of these limitations, while introducing challenges of their own, such as logistical challenges and fostering groupthink.

4.5 Conclusion

This chapter presents a monitoring module to detect risk situations in an eMental health self-help service to aid mourning older adults and to encourage users to seek offline support in times of need. The monitoring module uses an automatic fuzzy cognitive map (FCM) decision-making algorithm, configured with the help of eight clinical psychologists. FCMs are a powerful tool for modelling human reasoning due to their capacity to deal with vague definitions of symptoms provided by multiple discipline experts as well as the causal relations between them. Based on four fictitious scenario experiments, the resulting FCM appears to detect clear cases of risk and no-risk, but its accuracy in detecting less clear cases can be increased. To improve the current FCM algorithm, its assumption that symptoms are linearly related to the decision whether or not the user should seek offline support can be improved by introducing rules that regulate the influence of symptoms for which the linearity assumption does not hold, such as Time since loss. Another promising approach is to scrutinize the weight aggregation process and to actively deal with experts' disagreement regarding individual model weights. Due to their scrutability and independence of large amounts of privacy-sensitive patient mental health data for model configuration, FCMs hold much potential for automatic decision-making in eMental Health. However, lack of clear guidelines and golden standards for constructing FCMs require future work. The contribution of this research is showcasing the construction of a FCM decision algorithm for an eMental health service and in doing so, unravelling challenges when using FCMs for decision-making in eMental health applications.

4

Design implications for guiding older adults to offline support in eMental health

Based on:
Brandl, L., Jansen-Kosterink, S., Schokking, L., Siderakis, E., & Heylen, D. Design implications for transitioning from self-help to offline support in eMental health services for older adults. (*Submitted for publication*).

Prologue

Orphea stares at the screen of her computer. “We invite you to pause here for a moment. Were there situations in the past 2 weeks that were especially difficult? If so, looking back, what would you have needed? What would have made the situation easier, if just a little bit? We encourage you to confide in someone you trust about these thoughts.", Orphea reads from the screen. She had never given the mood check-up that the program prompts her to fill in every other week too much thought. Until now, the program had suggested that she was doing fine. What was different this time? Well, she had cancelled the vacation that she had looked forward to because she did not feel like going. And yes, she had made up an excuse not to go to Erika’s birthday last week. And there was the matter of the dancing classes. She had finally stopped attending. Even shortly after Justus’ death, she had continued dancing, but she hated being at the class without him. Still, she had continued going because she knew how much she had loved dancing and how much joy it had brought her in the past. “Difficult situations"? Were those “difficult situations"? Things came up, Orphea responded, best she could. End of story. If anything, she was trying to listen to what she wanted and needed, as the e-learning often suggested. Or was there more to this recommendation to...what? Talk to somebody about how she had been doing lately?

Abstract

Objective: Despite being at elevated risk to develop mental health problems, older adults hesitate to seek professional mental health support. Online self-help services can be a first, low-threshold step towards receiving mental health support. However, it remains unclear how older adults can effectively be guided from online self-help to offline support if their mental health deteriorates.

Method: The design of help-seeking recommendations based on regular mental health monitoring is explored via two focus groups with mourners who used an online grief service during a 10-week intervention, complemented with insights from data logs about how they used the monitoring, and the perspective of grief professionals on the design of stepping up online bereavement support.

Results and Conclusion: Effectively transitioning from online bereavement self-help to offline support affects the design of the entire service rather than singular monitoring features, should target older adults' perceived barriers, and should keep older adults in control about involving professional support.

5.1 Introduction

Despite having less stigma towards mental health conditions and help-seeking than younger adults, older adults are still less likely to seek professional mental health support (Kessler et al., 2015; Mackenzie et al., 2019). At the same time, the ageing process itself is associated with a growing risk of mental health decline due to common life stressors in later life (Hui Gan et al., 2020; Lyons et al., 2018), including the loss of loved ones. Older adults face barriers to seeking mental health support, including lack of perceived need for help (Meadows et al., 2002), poor mental health literacy (Farrer et al., 2008), and a lack of healthcare professionals specialized in treating older adults (Moye et al., 2019). In addition, older adults often wish not to burden those around them. This can motivate them to use digital technologies for mental health (Andrews et al., 2019; Peek et al., 2016). eMental health services have shown promising results for treating various mental health conditions in older adults, including depression, panic disorder (Carlbring et al., 2018), and prolonged grief (Brodbeck et al., 2017; Wagner et al., 2006). Some applications combine a web-based self-help program and minimal, but regular contact with healthcare professionals (Baumeister et al., 2014). Despite having reservations towards seeking professional support, it has been shown that older adults consider mental health support provided by professionals as superior compared to standalone online services (Andrews et al., 2019). Older adults also tend to adhere better to eMental health services when they are combined with professional support (Baumeister et al., 2014; van Zelst et al., 2021). To ensure that older adults receive the support they need, especially when digital services do not incorporate professional support by design, there is a need to detect when their mental health warrants professional intervention. eMental health services should thus be able to detect a deterioration in the user's mental health and communicate the implications of the detected mental health deterioration to the user (Tielman et al., 2019).

This implies a two-step approach to professional mental health support where the user first interacts with the self-help service that, among other things, provides information about the specific mental health condition the user is dealing with. The second step is a potential referral to offline professional support, either recommended by the system and initiated by the user, or automatically enforced by the system. This chapter discusses older adults' experiences with and use of an online grief service that implemented a two-step procedure that relies on the user's initiative. To complement users' perspectives, transitioning from online self-help to offline professional support in the Dutch healthcare system is further discussed with grief professionals, also exploiting their knowledge about older mourners' needs and perceived barriers to help-seeking. Ultimately, this chapter summarizes implications for designing a transition that is aligned with older adults' and professionals' viewpoints, thereby maximizing the chance that older adults will receive the mental health support they need.

5.1.1 Monitoring mental health in an online grief service for older adults

The online grief service used in this research has been developed for older mourners who have lost their spouse. It consists of ten content modules that contain psychoeducation about topics that are relevant for the grieving process (e.g., What are normal grief reactions?), cognitive-behavioral and reflective exercises that support the user in processing their loss (e.g., What are my current thoughts and feelings about the loss?) and self-care and activation strategies to support the user's restoration process (Brodbeck et al., 2022). The grief service is complemented by a monitoring system that determines in bi-weekly intervals whether it is advisable for the user to seek offline (professional) support. The monitoring system has two main components: a mental health user profile and a decision engine. The mental health user profile consists of two self-report monitoring questionnaires, an initial risk assessment (IRA) and a continuous risk assessment (CRA). The IRA identifies some characteristics of the loss, such as how long ago it occurred and whether it was violent (e.g., an accident), and represents an initial assessment of the user's grief symptoms and mood. The CRA updates the user's mental health profile every other week. Both questionnaires are included in the Appendix and their development is described in detail in Brandl et al. (2023). The decision engine consists of a set of rules that determines whether the user exceeds a suicidal threshold and a fuzzy cognitive map (FCM) decision algorithm. The suicidal threshold was determined by a team of clinical grief experts and by taking into account the assessment manual of the scale of suicide ideation (SSI) (Beck et al., 1979), a clinical research instrument designed to quantify and assess suicidal intention on which the IRA and CRA suicidality items are based. FCMs are a soft computing technique that model human reasoning as directed graphs and use fuzzy logic to represent causal influences between conceptual graph elements, such as the extent to which elevated grief symptoms influence the decision to recommend help-seeking, or how they influence a person's social withdrawal behavior (Stylios et al., 2008). The FCM algorithm runs on the mental health parameters in the user's profile whenever they are updated. It arrives at the decision to either display a recommendation to the user to seek offline support, shown in Figure 5.1, or an encouragement to continue using the grief service as is. For this, a competitive fuzzy cognitive map (CFCM) architecture (Stylios et al., 2008) was employed with two competing outputs: recommend help-seeking and recommend using the grief service as before. The development of both, the monitoring questionnaires (IRA and CRA) and the FCM decision algorithm are described extensively in Brandl et al. (2023). Filling in the CRA is not obligatory in the grief service, i.e. users can choose to ignore the service's reminders to fill in the monitoring questionnaire.

5.2 Method

The design implications for mental health monitoring and transitioning from online self-help to offline professional mental health support are based on a) an evaluation study that has been conducted between February and November 2022 in the Netherlands (Brodbeck et al., 2022) that included b) four focus groups, two with older adults who had previously participated in the evaluation study and two with grief professionals. The aim of the evaluation study was to investigate the technology acceptance of the above described online grief service and secondarily, the service's potential health effects. Its study protocol received an exemption from medical ethical approval from the medical ethical committee Oost-Nederland (file number 2021-13268). The study involved older adults (55+ years old) as end-users, recruited from the general population via the networks of two Dutch organizations, the National Foundation for the Elderly[1] and DELA[2], a funeral insurance company.

5.2.1 Participants

In total, 51 older mourners registered with the grief service. To be eligible for inclusion in the current research, participants had to

1. Fill in the initial risk assessment (IRA) ($n = 51$).
2. Complete the introduction to the online grief service. This included filling in an initial shorter version of the continuous risk assessment (CRA) ($n = 41$).
3. For comparisons between self-report and logged use of the monitoring questionnaires, participants had to fill in a self-report questionnaire five weeks after registration with the grief service ($n = 31$).

After applying the above inclusion criteria, the log data of 41 participants was included. Their mean age was 66.4 years ($SD = 8.2$ years) and 26 participants were female and 15 participants were male. Nine older adults participated in the focus groups with end-users, six in an online group and three in a face-to-face setting. They were invited via a telephone call or via e-mail to join the focus groups. In total, ten professionals participated in the focus groups, four in one group and six in the other. All professionals had experience working with (older) mourners and included grief coaches, psychotherapists, geriatric nurses, ministers, and volunteers at organizations that offer support for mourners. All participants signed an informed consent form.

[1] https://ouderenfonds.nl/
[2] https://www.dela.nl/

5.2.2 Data collection

Older adults' use of the monitoring questionnaires was assessed in two ways: via an online self-report questionnaire, administered halfway through the evaluation study via Qualtrics, and continuously via data logs. In the questionnaire, participants were asked whether they had filled in the CRA, how often they had filled it in (Never, Once, 1-3 times, 3-10 times), and if they had received a recommendation to seek offline support. If they had, they were asked about their adherence to the recommendation, about how they felt about the monitoring questions, and whether they had any privacy concerns. Participants could add an explanation to their response. The online grief service logged when users received an invitation to fill in a CRA, when they filled in the CRA (or if they did not fill it in), and whether the system had displayed a recommendation to seek offline support. It was not possible to reconstruct whether users had contacted a healthcare professional based on the log data.

During the focus groups, participants who reported to have filled in the CRA were asked whether they had received a recommendation to seek offline support. The follow-up questions focused on what participants' thoughts were about the monitoring questions and the system estimating whether they might need offline support. Participants who did not fill in the monitoring questionnaires were asked about their motivation not to use the functionality. Prior to the focus groups, the grief service was introduced via a demo video to the grief professionals. The video demonstrated the functionalities, structure, and look of the service. During the focus groups, the professionals were asked about their perspective on how older mourners can transition from online self-help to offline professional support and which role professionals could and should play in mental health support provided by an online grief service.

5.2.3 Data analyses

The use data was analyzed regarding the number of continuous risk assessments (CRAs) older adults filled in, depending on the progress of the grief intervention (measured in weeks). We compared users' perceived use of the monitoring system with their actual use by comparing how often they reported to have filled in the CRA with system logs. An inductive coding scheme was developed and applied to users' free text explanations in the self-report questionnaire and to what was discussed about the monitoring system during the focus groups. The coding scheme was developed and applied by one researcher and verified by a second researcher. Any discrepancies were discussed until agreement was reached. An example from the coding scheme is *Usability (and expectation management)* which summarized potential usability issues and unclarities surrounding the capacities of the monitoring system. The focus groups with grief professionals were analyzed following the same procedure.

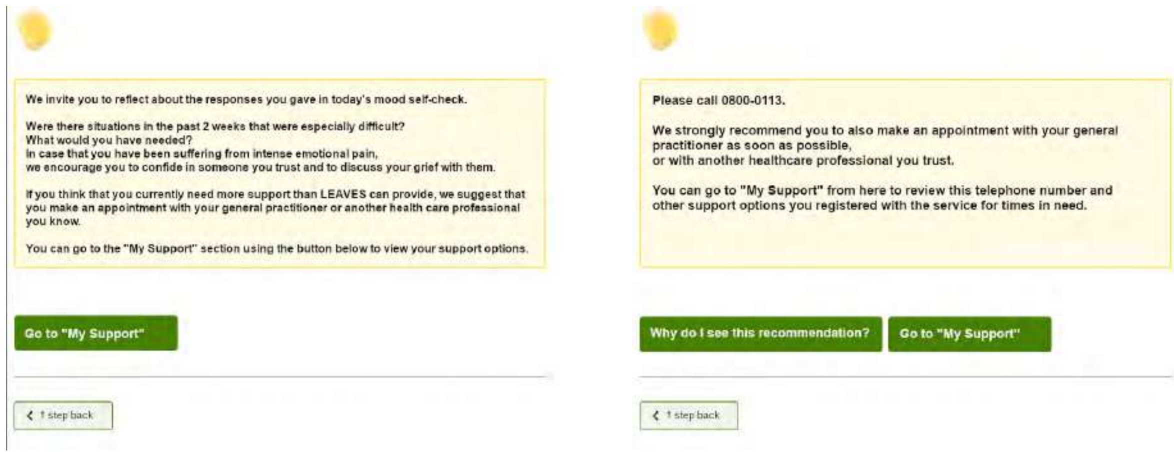
An example from the coding scheme is *Purpose of monitoring*, summarizing professionals' perspective on which goals a monitoring system should serve in an online grief service.

5.3 Results

5.3.1 Use of the monitoring questionnaires

Figure 5.2 shows for each participant in which intervention week they filled in a CRA questionnaire, based on log data. All participants completed the initial shortened version of the CRA during the program's introduction. After two weeks, 16 out of the remaining 27 participants who logged into the service filled in the CRA. From week six on, the number of participants that did not fill in the CRA exceeded those who completed the CRA. By week ten, only 14 of the initial 41 participants (34.15%) still logged into the grief service. Figure 5.2 shows that some participants responded to the CRA, but stopped at some point. For some, not filling in the CRA preceded non-use of the grief service (e.g., p 6 and p 32). Others did not respond to some CRAs but did respond to later instances again (e.g., p 7 and p 39).

Comparing how often participants filled in the CRA based on data logs compared to self-reports, seven participants reported not to have used the monitoring questionnaires at all. Out of those seven, however, log data shows that four did fill in the IRA. Three participants filled in one, or multiple instances of the CRA in addition to the IRA. Participants further tended to underestimate (18 out of 31) how often they used the monitoring, or to report correctly (11 out of 31). Only two participants overestimated how often they had filled in the monitoring questions.



(a) (b)

Figure 5.1: Screenshots of the help-seeking recommendations displayed in the grief service: a) Message displayed for suffering detected by the monitoring system, and b) message displayed when users' responses exceed a suicidal threshold.

5

5.3.2 Monitoring appreciation by older adults

Self-report questionnaire

Six participants reported to have received a recommendation to seek offline support. As per design of the questionnaire, these six participants received follow-up questions about their experience with the monitoring system. Regarding adherence to the recommendation to seek support, all six participants reported to have followed-up on the recommendation, to varying degrees. One participant explained they adhered to the recommendation because other people in their social circle had advised the same. Another participant explained that they adhered to the recommendation only sometimes because they preferred coping with their grief by themselves. One participant reported not to feel comfortable with the monitoring questions, the other five were comfortable with the questions, to varying degrees. The former explained that they had no interest in talking to a professional. Another participant explained they trusted that the people behind the grief service knew what they were doing by asking these questions. None of the participants voiced privacy concerns regarding the monitoring, though most found the questions to be privacy-sensitive and explained answering them can make you feel vulnerable.

Focus groups with older adults

Regarding usability and expectation management, the service lacks clarity in communicating to the user what the monitoring system offers and what not:

> “I could endorse or reject these [monitoring] questions accordingly, but then you are not being called by a psychologist. (...) So I wondered, what happens then? You should not get the idea that these questions are a real cry for help and that you can expect someone will approach you.” (Primary end-user focus group 2)

Regarding the involvement of professional support, older adults had diverging opinions about designing an it in such a way that it requires the user to take initiative.

> “It [the user taking initiative] is worrisome because when you say that you are not feeling well, the obstacle to actively call someone, the general practitioner or a hotline, is huge.” (Primary end-user focus group 2)

Another participant explained that they had taken initiative and found support in a churchman. Participants’ attitudes towards being actively called by a

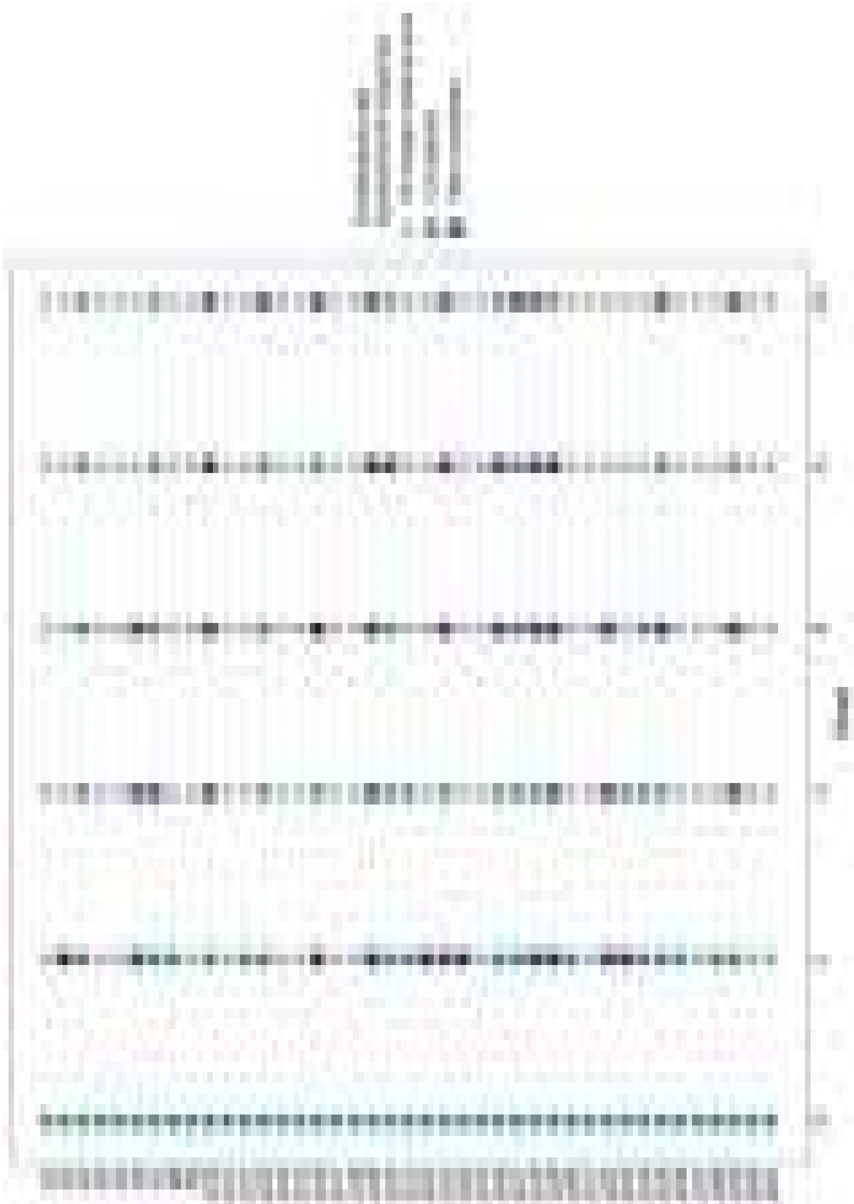


Figure 5.2: Users' use of the continuous risk assessment (CRA) questionnaire per intervention week. Week 0 represents the initial shortened version of the CRA during the program's introduction to the user.

5

professional were comparably divided. One participant said that they would appreciate any help, another rejected the option because they wanted to stay in control, while a third participant could agree with both options as long as they could choose either option at the beginning of the service. Something else to consider for monitoring the mental health of mourners is that sometimes, mourners simply do not know how to answer direct questions about how they are feeling. Being at a loss for answers can weigh emotionally on them and that can impact their motivation to respond to questions:

> “It is very personal and easy to interpret as self-pity by a third party I guess.” (Primary end-user in the evaluation study)

Older adults’ opinion about the monitoring system and its recommendations was influenced by how they thought about grief in general:

> “Emotionally, I came to the same conclusion [about why they felt comfortable with the recommendation to seek offline support].” (Primary end-user in the evaluation study)

> “You face it alone. No digital program can help and no therapist who tells you what to do. There is just one person who will get me through it and that is me.” (Primary end-user focus group 1)

5.3.3 Monitoring appreciation by grief professionals

Regarding the purpose of a monitoring system in an online grief service, professionals explained that any system feedback should ultimately encourage the user in their grief process.

> “I like the idea of people receiving a reminder ’How are you doing now?’, but I am not sure whether the outcome ’Please contact...’ has to be linked to it. You should get something positive out of it.” (Grief professional focus group 1)

When recommending users to seek offline support, it is crucial how the purpose of the monitoring and the recommendation itself are communicated in the service.

> “From the start, the goal should be clear: ’You are in a process, you are not a lost cause if you receive the recommendation to seek offline support’.” (Grief professional focus group 1)

The service, including the monitoring, must avoid coming across as patronizing and controlling.

> "As a user, (...) I want to face my grief, but now there is this 'Big Brother' that is assessing whether I am doing it well." (Grief professional focus group 2)

The professionals appreciated the reflective nature of the monitoring questions and suggested that the feature could be customizable for the user, so that they can choose to receive a reminder about the monitoring questions, or choose not to receive recommendations to seek support. Regarding transitioning from online self-help to offline support, professionals in both focus groups unanimously advocated relying on users' initiative for reaching out to offline support despite any barriers the user may perceive:

> "You could give care professionals rights to the grief service, that way they would get the user's contact details so that they can take action. But I am convinced that everyone is responsible for themselves. In practice, this [professionals taking initiative] would be very difficult, too." (Grief professional focus group 1)

However, the professionals did emphasize that an online service should facilitate the help seeking process as much as possible to reduce the perceived barriers surrounding seeking offline mental health support.

> "I [as a user] do not know the healthcare domain, I have no idea where to go and to whom I can turn. It would be very supportive to also lend a hand in this step." (Grief professional focus group 1)

The grief professionals pointed out that an online service should not reinvent the wheel, but connect to already existing support options, such as professionals in the vicinity of the user. Exploring local support offers via content offered in the grief service would help minimizing perceived barriers, as would normalizing feeling depressed and hopeless after a loss. Another way to reduce perceived barriers is to include low-threshold support options including volunteer organizations and peer groups, rather than focusing on professional support.

5.4 Discussion

This chapter explores design implications for eMental health services with the aim of guiding older users to offline (professional) support if their mental health deteriorates. The implications are based on older adults' use and appreciation of a monitoring system in an online grief service during a 10-week intervention, complemented with insights from grief professionals' appreciation of the same monitoring system. In particular, we explored their viewpoint on offering online self-help initially and encouraging self-referrals to offline (professional) support. In the focus groups, we discussed the transition from self-help to offline support and the impact of self-reported monitoring questions on elderly users.

5.4.1 Design Implications

Our main findings can be summarized in four design implications for transitioning from online self-help to offline support:

Implication 1

A monitoring system that recommends offline help-seeking should be conceptually ingrained in the entire eMental health service. The online service should target perceived barriers older adults experience towards seeking mental health support (Bui, 2018; Shear, 2015), such as increasing their mental health literacy (Kessler et al., 2015) about grief by normalizing emotional suffering after a loss and distinguishing it from pathological grief reactions. The service should make clear that help-seeking recommendations do not signify “failure” on the part of the mourner. This approach is in line with earlier research (Kessler et al., 2015) about older adults’ attitude towards seeking mental health services and fosters their psychological openness, the extent to which people acknowledge psychological problems. Incorporating lower-threshold offline support options in the application may contribute to increasing older adults’ perceived social support, another important determinant of their attitude towards seeking support. A monitoring system aimed at recommending offline professional support thereby goes beyond singular monitoring features and the formulation of recommendation messages. The eMental health service as a whole is faced with the task of conveying that “it is O.K." to ask for support when older adults face emotional pain that goes beyond what they can manage by themselves, and consequently, the pragmatics of seeking support.

Implication 2

Some users may be more inclined than others to fill in self-reported monitoring questionnaires and to receive recommendations about offline help-seeking. Our use data results show that some participants stopped filling in the monitoring questionnaires altogether, while others appear to have filled them in according to their needs. A distinction between features that are included in the online service by design, such as unobtrusively targeting older adults’ perceived barriers to help-seeking, and those that should be subjected to the user’s preferences will help designing a monitoring system that is acceptable and useful for a variety of users. For example, a user could choose to fill in the self-reported monitoring questions for reflective purposes, but prefer that the online service does not give advice about contacting a professional. Ideally, a monitoring system should be able to accommodate a pre-defined set of user preferences regarding how the systems manages stepping up mental health support options.

Implication 3

A monitoring system aimed at detecting a need for professional intervention should not (solely) rely on self-reported user data due to irregular and declining use patterns. Participants in this research tended to underestimate how often they had filled in the monitoring questionnaires, suggesting that filling in the continuous risk assessment was an acceptable burden. However, participants perceived the monitoring questions as confronting. Asking users directly how they are feeling can cause emotional distress, especially if they find it difficult to answer. There is no evidence indicating that asking people about their emotional states, including suicidal ideation, increases the risk of developing pathological mental states (Blades et al., 2018). Nevertheless, less obtrusive ways of monitoring users' mental health, such as automatic analysis of texts (possibly in response to more indirect monitoring questions) (Zhang et al., 2022) and log data should be considered as inputs for future mental health monitoring systems. In addition, awareness about the emotional impact of direct monitoring questions can guide service designers to combine monitoring assessment moments with content that promotes self-care.

Implication 4

While grief professionals advocate that users should stay in control while transitioning from self-help to offline support, the viewpoints of older adults are divided. The latter worry about reaching out to support in times of emotional distress which could be alleviated the earlier described design strategies. Grief professionals see a parallel between transitioning from online self-help to offline support and existing paths to mental healthcare where the client takes the first step. However, they stress that non-clinical, low-threshold support options for grief, such as mourning groups, may be a more feasible first step, thereby perpetuating the idea that transitioning is a step-wise process, involving familiarizing older adults with facilitators that can help them seek professional mental health support.

5.4.2 Future work

The monitoring system described in this chapter distinguishes between generic help-seeking recommendations and recommendations involving a suicidal threshold. Tielman et al. (2019) developed a complementary motivation model for self-referral in online mental health services based on theories of behavior change. The model determines one of three persuasion strategies (facilitating, persuading, accepting rejection) based on the severity of the risk situation and the user's initial stance towards involving professional support. The model also recommends accepting that users may reject professional support when their stance towards help seeking is negative and the situation is not severe. To ensure that users receive the mental healthcare they need, the persuasive capacities of monitoring systems should be considered in conjunction with the implications found in this chapter.

5.4.3 Limitations

The current research has limitations. First, we recruited older mourners from the general population and therefore, the proposed design implications may not hold for services targeted at clinical populations. Second, the design choice to only inquire qualitatively after users' monitoring appreciation in the self-report questionnaire if they indicated to have received a recommendation to seek offline support severly limited the number of qualitative responses collected. A final limitation of this research is the number of dropouts during the evaluation study. Dropout is a well-known issue in online health services (Eysenbach, 2005; Richards & Richardson, 2012). Unfortunately, participants' reasons for dropout were beyond the scope of this chapter and it is unclear to what extent the monitoring system may have contributed to their choice to discontinue with the grief service.

5.5 Conclusion

To ensure that older adults receive the mental health support they need, eMental health services should be equipped with strategies to guide users from online self-help to offline professional support whenever advisable. This chapter proposes four design implications for eMental health services that initially offer online self-help and rely on self-referrals to professional offline support encouraged by system recommendations. First, the transition from online self-help to offline support should be deeply ingrained in the entire service and target older adults' perceived barriers to seeking mental health support. Second, self-reported monitoring of users' mental health should be subjected to the user's preferences and take into account that some users do not wish any involvement of healthcare professionals. Third, alternatives to self-reported mental health monitoring should be considered to limit missing data and unintended emotional burden for users. Finally, regarding the allocation of help-seeking responsibility, any specific eMental health service designed for the general population should identify the barriers its users face to help-seeking and empower them to take initiative by targeting these barriers.

Towards meaningful evaluations of monitoring in eMental health: the case of an online grief service for older mourners

Based on:
Brandl, L., Jansen-Kosterink, S., Brodbeck, J., Jacinto, S., Mooser, B., & Heylen, D. Moving Towards Meaningful Evaluations of Monitoring in eMental Health based on the Case of a Web-Based Grief Service for Older Mourners: Mixed Methods Study. *JMIR Formative Research*; 8:e63262. doi:10.2196/63262.

Prologue

Orphea switches on the light as she pulls off her boots in the hallway, returning from an appointment with Dr. Epione, her general practitioner. She had doubted whether the appointment had been the right call and surely, she had not expected the appointment to end with her crying her eyes out over how she had stopped doing things. Not since Justus was gone, but even before his death. She had stopped making vacation plans because Justus had been very good at planning trips and doing the research on sights and accommodations. Not only had he been good at it, he had *loved* doing it. So naturally, over the years, Orphea's contribution to making vacation plans had boiled down to finding a pet-sitter for their cat and to packing bags. A few years ago, Orphea quit the walking club because it became too difficult to schedule with dancing classes twice a week, looking after their granddaughter once a week, the garden...Orphea did not realize that she stopped seeing the acquaintances she had made at the club soon after. She did not notice that the "friends" Justus and she had often invited for dinner were almost exclusively friends Justus had made in college. Her own college friends had been living far away, too far to keep in touch regularly. She did not realize that somewhere between the kids moving out, moving from one side of the country to the other, and making a new home together with her husband, she had somehow stopped being "Orphea" and had become a half of something whole that no longer exists. And tonight, in Dr. Epione's office, the realization sank in. And the tears followed.

Abstract

Background: Artificial intelligence (AI) tools hold much promise for mental healthcare by increasing the scalability and accessibility of care. However, current development and evaluation practices of AI tools limit their meaningfulness for healthcare contexts and thereby, the practical usefulness of such tools for professionals and clients alike.

Objective: To move towards meaningful evaluation of AI tools in eMental health, this chapter demonstrates the evaluation of an AI monitoring tool that detects the need for more intensive care in an online grief intervention for older mourners who have lost their spouse.

Method: We leverage the insights from three evaluation approaches: (1) the F1-metric evaluates the tool's capacity to classify user monitoring parameters as (a) in need of more intensive support, or (b) recommendable to continue using the online grief intervention as is; (2) we use linear regression to assess the predictive value of users' monitoring parameters for clinical changes in grief, depression, and loneliness over the course of a 10-week intervention. Finally, (3) we collect qualitative experience data from eCoaches (N=4) who incorporated the monitoring in their weekly e-mail guidance during the 10-week intervention.

Results: (1) Based on N=174 binary recommendation decisions, the F1-score of the monitoring tool was 0.91. (2) Due to minimal change in depression and loneliness scores after the 10-week intervention, only one linear regression was conducted. The difference score in grief before and after the intervention was included as dependent variable. Participants' (N=21) mean score on the self-report monitoring, the estimated slope of individually fitted growth curves, and its standard error (i.e. participants' response pattern to the monitoring questions) were used as predictors. Only the mean monitoring score exhibited predictive value for the observed change in grief ($R^2 = 1.19, SE = 0.33, t(df) = 3.58(16)$, $P = .002$). (3) The eCoaches appreciated the monitoring tool as an opportunity a) to confirm their initial impression about intervention participants, b) for personalizing their e-mail guidance and c) to detect when participants' mental health deteriorated during the intervention.

Conclusion: The monitoring tool evaluated in this chapter identifies a need for more intensive support reasonably well in a non-clinical sample of older mourners, has some predictive value for the change in grief symptoms during a 10-week intervention, and is appreciated as an additional source of mental health information by eCoaches who supported mourners during the intervention. Each evaluation approach in this chapter comes with its own set of limitations, including (a) skewed class distributions in prediction tasks based on real-life health data and (b) choosing meaningful statistical analyses based on clinical trial designs that are not targeted at evaluating AI tools. However, combining multiple evaluation methods facilitates drawing meaningful conclusions about

6

the clinical value of AI monitoring tools for their intended mental health context.

6.1 Introduction

Artificial intelligence (AI) tools hold much promise for mental healthcare by increasing the scalability and accessibility of care (Teachman et al., 2022). They have the potential to identify warning signs of serious mental health problems earlier than current mental healthcare systems allow and to timely deliver (digital) mental care, potentially preventing the full onset of mental health disorders, or limiting the severity with which they impair people's lives (Muñoz, 2022; Teachman et al., 2022). For example, Sakal et al. (2022) describe the development and evaluation of an AI-based screening tool for geriatric depression in elderly Chinese people. Taking into account cultural response biases to traditional depression screening tools, the tool focused on less emotionally sensitive demographic and quality of life predictors such as health status compared to 3 years ago, hearing status, income, and average hours of sleep per night in the previous month. The tool was found to perform well during validation and the authors explain the importance of the non-sensitive nature of the questions employed by the screening tool for early detection of geriatric depression in the Chinese aging population. The tool represents a means for Chinese public health officials to fight the growing mental health treatment gap in the country. Likewise, Zhang et al. (2024) leveraged AI to extensively analyze behavior-related and physiological risk factors for suicide in middle- and older-aged individuals who participated in the UK Biobank population-based cohort that was recruited between 2006 and 2010. The use of AI and advanced statistical tools enabled the authors to systematically identify and rank 246 behavior-related and 200 physiological factors and identified 58 robust predictors for suicide risk. The authors explain that the gained insights unravel new potential avenues for targeted suicide prevention.

Despite such promising examples of how AI tools can contribute to increasing the scalability, accessibility, and effectiveness of mental healthcare, AI tools are currently still considered to be in a proof-of-concept stage rather than creating clinical impact for mental healthcare (Thieme et al., 2023). Tornero-Costa et al. (2023) describe a mismatch between clinical trial designs that are common in mental healthcare and desired data qualities for AI development which are often difficult to reconcile in terms of time, money, and human resources. AI tools and clinical trials have fundamentally diverged data sampling considerations, specifically concerning exclusion criteria that are common in clinical trials to limit the influence of confounders on tested clinical outcomes, or due to safety considerations. However, given large enough sample sizes, confounders improve the generalizability of AI models which makes them a necessary element in any representative data set. In mental healthcare, AI tools are currently often developed in retrospection as secondary outcomes of clinical trials and are based on clinical data collected for other purposes than model development (Tornero-Costa et al., 2023).

In addition, current AI model engineering approaches for mental health are criticized for their focus on perfecting model performance without practical clinical value (Cabitza & Campagner, 2021; Whiting & Fazel, 2019). Whiting and Fazel (2019) explain in their recent clinical meta review on the accuracy of prediction models for detecting suicide risk that only a few models are developed with independent clinical validation or piloting in mind. Model developers tend to neglect the clinical meaning of the association between predictors and model outcomes and are not transparent in their decision-making process leading to the selection of model parameters (Tornero-Costa et al., 2023). Furthermore, current practices favor model evaluation metrics such as predictive accuracy without explaining how they are linked to a clinical decision. In the specific context of suicide risk detection, the authors advocate that prediction models should be compared to unstructured clinical assessments of suicide risk to investigate the incremental benefit of these tools supporting clinician decision-making. Ultimately, suicide prediction, as any mental health prediction task, is challenging for both, data-driven prediction and clinical practitioners. To build AI tools in mental healthcare with clinical impact, we need to start developing and evaluating models whose outcomes can be clearly linked to clinical decision-making and their roles in clinical practice should be well-defined. In this chapter, we evaluate a mental health monitoring tool in an eMental health service for older mourners by combining the insights from three evaluation approaches. We encounter some challenges that are common in AI evaluation studies and showcase how these affect the clinical meaningfulness of our obtained results. We thereby exemplify the need for AI tools in mental healthcare to go beyond classical AI evaluation metrics and statistical approaches in clinical research to have an impact. The next section briefly introduces the monitoring tool and the eMental health service in which it is embedded before describing our evaluation approach in more detail. The eMental health service for which the monitoring tool was developed supports older mourners in processing the loss of their spouse. We conclude with a discussion of the encountered evaluation challenges and some suggestions how to move the development of impactful AI tools in mental healthcare forward.

6.2 Method

6.2.1 Background: the monitoring tool

The monitoring tool that we evaluate in this paper is implemented in an online grief service for older mourners who have lost their spouse. The grief service consists of 10 content modules (e.g., unravelling myths and truths about grief) and exercises and activity suggestions that help the mourner process the loss and foster positive mental and physical well-being (e.g., writing a farewell letter to the deceased spouse, reconnecting with one's hobbies) (Brodbeck et al., 2022). The monitoring tool complements the service with a bi-weekly mental health self-check and by analyzing whether it is advisable for the user to seek offline (professional) support. It has two components: a mental health user profile and a decision-making component. The mental health user profile consists of two self-report questionnaires, an initial risk assessment (IRA) and a continuous risk assessment (CRA). The IRA represents an initial assessment of the user's affective state and grief symptoms when they start using the online grief service and controls for risk factors such as whether the loss has been violent (e.g., their partner committed suicide). The CRA assesses the extent to which the mourner experiences psychological suffering. The decision-making component consists of a set of rules that determines whether the user exceeds a suicidal threshold and a fuzzy cognitive map (FCM) decision algorithm. It arrives at the decision to either display a recommendation to seek offline support, or an encouragement to continue using the grief service as is. Filling in the CRA is optional, the grief program can be used without it. The development of the monitoring tool, including its parameter selection, the construction of the two monitoring questionnaires (IRA and CRA), and an initial error analysis based on fictitious scenarios, is described in detail in Brandl et al. (2023).

6.2.2 Evaluation context: randomized controlled trial and eCoach focus group

The current evaluation of the mental health monitoring tool is based on an ongoing randomized controlled trial (RCT) that started in March 2022 in Switzerland (Brodbeck et al., 2022). The aim of the RCT is to investigate the clinical efficacy of the above described online grief service and secondarily, to examine which delivery format of the online grief service (standardized vs. self-tailored) is associated with better clinical outcomes. At the time of writing this chapter, the RCT is ongoing until the needed sample size to test the two delivery formats of the service (standardized self-tailored) is achieved.

Our evaluation approach uses the data of older mourners (60+ years old) who participated in the RCT. Participants were recruited from the general population and had experienced the loss of their partner at least 1 month before the RCT. A more extensive list of inclusion and exclusion criteria can be found in the dedicated study protocol of the RCT (Brodbeck et al., 2022). During the RCT, four eCoaches provided guidance in the form of a weekly e-mail with short, personalized feedback and support. The eCoaches were encouraged to include participants' self-reported mood and therapeutic progress and the outcome of the monitoring tool in the guidance that they provided.

6.2.3 Evaluation approach

To evaluate the monitoring tool, we: (1) assess the classification performance of the monitoring decision algorithm using the F_1-metric, (2) investigate the predictive value of participants' monitoring responses for their clinical change in grief, depression, and loneliness after the 10-week RCT, and (3) collect qualitative user experience data to explore the tool's suitability for clinical practice from trained eCoaches who used the monitoring for their work during the RCT.

For (1), the classification performance was assessed using ground truth classification labels provided by the eCoaches for the tool's binary outcome (recommendation to seek support vs. encouraging to continue using the grief service as is). The eCoaches determined the ground truth labels based on their professional assessment given the participant's progress in the eMental health service, their bi-weekly monitoring responses, weekly e-mail exchanges with the participant, and a clinical interview at the beginning of the RCT. The monitoring's suggested classification was visible to the eCoaches alongside participants' raw monitoring responses to facilitate the eCoach's understanding of the classification. If the eCoach's assessment diverged from the outcome of the monitoring tool, they provided a brief textual explanation about their rationale. The ground truth labels for the monitoring predictions were provided by the eCoaches upon request at the time of conducting this analysis. (2) explores the predictive value of the CRA by relating CRA scores to the difference in clinical measurements before and after the RCT. (3) focuses on the eCoaches' experiences with the monitoring tool during the RCT. An online focus group was conducted in which the four eCoaches discussed how they used the monitoring tool in their role as eCoaches, how they experienced having the tool at their disposal, and how they think such a tool could be most useful for mourners who use the online grief service and for eCoaches such as themselves.

Measures

The continuous risk assessment (CRA) is a multi-dimensional scale that measures Hopelessness, Grief Symptoms, Social Isolation, and Psychological Crisis with two items each on a 4-point Likert scale. The items assess the frequency of emotional suffering in the past two weeks ranging from 0 (Not at all) to 3 (Every day). The CRA also measures Therapeutic Progress on a 4-point Likert scale ranging from 0 (Strongly disagree) to 3 (Strongly agree). The CRA serves as input for the fuzzy cognitive map (FCM) algorithm as part of the decision-making in the monitoring. Its development is described in more detail in Brandl et al. (2023) and a copy of the tool is included in the Appendix. For this study, the CRA scores of RCT participants were retrieved from data logs of the grief program. The three clinical measures (grief, loneliness, depression) were assessed at three measurement moments during the RCT via an online surveying tool: (1)

prior to starting the online grief program (t_0), (2) after completing the 10-week intervention (t_1), and (3) 20 weeks after starting the intervention program (t_2). For the current evaluation of the monitoring tool, we only take the first two measurement moments into account. The clinical measures include an assessment of the mourner's (1) grief symptoms using the Texas Revised Inventory of Grief (TRIG) (Futterman et al., 2010), (2) depressive symptoms using the Patient Health Questionnaire-9 (PHQ-9) (Martin et al., 2006), and (3) loneliness via the de Jong Gierveld Loneliness Scale (De Jong Gierveld & Van Tilburg, 2010; de Jong-Gierveld, 1987).

Data inclusion

For the evaluation of the classification performance, our first approach, we included any monitoring decision for which the eCoaches provided a ground truth label. For assessing the predictive value of the CRA for clinical change during the RCT, however, we only included participants who had (1) completed the 10-week intervention and (2) filled in the clinical measurements (depression, grief, loneliness) at baseline and 10 weeks after starting the intervention. We did not expect the delivery format of the grief program (self-tailored, standardized) nor that participants in the waitlist control condition received access to the intervention only after 12 weeks to impact the decisions of the monitoring algorithm. Likewise, we did not expect the delivery format or the waitlist control condition to affect the relation between how participants filled in the CRA and the clinical outcomes. Therefore, we included participants from all arms of the RCT in this analysis. Specifically, we included CRA scores from RCT weeks 2 to 10. In week 2, the CRA was administered for the first time.

Ethical considerations

The RCT based on which we evaluate the current AI monitoring tool of an online grief service received medical ethical approval by the Medical Ethical Committee of Northwestern and Central Switzerland (Business Administration System for Ethics Committees number 2021-02221) and is registered at ClinicalTrials.gov (NCT0528004)[1]. All older mourners who participated in the RCT signed an informed consent form that has been approved by the Medical Ethical Committee of Northwestern and Central Switzerland, allowing the secondary analysis of their monitoring data for the purpose of evaluating the online grief service with no further consent required. In addition, the eCoaches provided written informed consent prior to participation in the focus group. All analyses involving data of RCT subjects were conducted on an anonymized data set where each participant is represented by an arbitrary code that is not

[1]https://clinicaltrials.gov/ct2/show/NCT05280041

related to their identity. The eCoach focus group was recorded and automatically transcribed using Microsoft Teams. After checking the correctness of the automatic transcription, the recording was deleted, and the transcription was de-identified. All subsequent analyses were conducted using the de-identified focus group transcription. Participants did not receive (financial) compensation for participating in this research. We refer the reader to the dedicated RCT study protocol for more detailed information about its ethical review process (Brodbeck et al., 2022).

6.2.4 Data analyses

Analysis I: Classification evaluation

Our data was sampled from a non-clinical population. Therefore, we expect few (true) help-seeking recommendations in the sample. This has implications for choosing an appropriate classification evaluation metric (Paula Branco & Ribeiro, 2016; Sun et al., 2009). Regarding terminology, in the binary classification problem at hand (recommending to seek offline support vs. recommending to continue using the service as is), we chose the less frequent class, recommending help-seeking, as the positive outcome class and recommendations to continue using the service as the negative class, as recommended for imbalanced classification problems (Paula Branco & Ribeiro, 2016). The F-measure is used when there is no clear preference for either minimizing false positives (someone receives an unjustified recommendation to seek support) or false negatives (someone who needs support does not receive a recommendation to seek support) because both are regarded as equally important for determining the classifier's performance. The F_1-score is the harmonic mean between the true positive rate (TPR; recall) of a classifier and its precision:

$$recall = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

$$precision = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

$$F_1 = 2 * \frac{\text{recall} * \text{precision}}{\text{recall} + \text{precision}}$$

The F_1-score is bounded to the interval $[0, 1]$, where 1 represents maximum precision and recall and 0 represents zero precision and/or recall. All calculations were performed in Python 3.11 (Van Rossum & Drake, 2009).

Analysis II: Predicting clinical change using monitoring measurements

The reliability of the continuous risk assessment (CRA) questionnaire is assessed using Spearman-Brown ρ_{SP} coefficients for its five two-item sub-scales: Psychological Crisis, Hopelessness, Grief Symptoms, Social Isolation, and Therapeutic Progress (Eisinga et al., 2013). As pointed out by Tavakol and Dennick (Tavakol & Dennick, 2011), if a scale measures several constructs, it is recommended that reliability is assessed separately for each construct. Since each CRA construct is measured using two items, Spearman-Brown coefficients were deemed the most appropriate method for assessing reliability (Eisinga et al., 2013). To investigate the relation between the CRA and the clinical outcomes of the RCT, we conduct three linear regression analyses with the difference in clinical outcomes before and after the grief intervention as dependent variable and parameters of an individually fitted linear growth curve (the estimate of the linear coefficient and its standard error, i.e. the slope of the linear curve) and mean CRA scores as independent variables, as suggested by Welten et al. (Welten et al., 2018) for repeatedly measured predictors. Analyses were performed using Python 3.11 and R (R Core Team, 2021).

Analysis III: eCoaches' experience with the monitoring tool

An inductive coding scheme was developed and applied in ATLAS.ti (ATLAS.ti Scientific Software Development GmbH, 2023) to the transcript of the eCoach focus group about their experiences with the monitoring tool during the RCT. The coding scheme was developed and applied by one researcher and verified by a second researcher. Any discrepancies were discussed until agreement was reached. An exemplary code is *monitoringExp* which summarizes experiences and thoughts that the eCoaches had about having the monitoring at their disposal.

6

6.3 Results

6.3.1 Analysis I: Classification evaluation

Participants

The data of 44 RCT participants was included in the assessment of the classification performance of the monitoring module. On average, these 44 participants filled in 4.02 (SD 2.2) out of the five bi-weekly CRA questionnaires during the 10-week intervention, amounting to 174 monitoring decisions that were labelled by hand by the eCoaches.

Table 6.1: Confusion matrix of the monitoring decision algorithm that either recommends help-seeking or to continue using the mental health service with no change.

		Predicted		
		Positive (Help-seeking recommendation)	**Negative** (No help-seeking recommendation)	Total
True lables	**Positive**	5	1	6
	Negative	0	168	168
	Total	5	169	174

Confusion matrix and F_1-score

Table 6.1 shows the confusion matrix for the monitoring algorithm's decision-making. Most labelled monitoring decisions (n = 168) were true negatives (TNs), reflecting that detecting the need for professional intervention in an online grief service is an extremely imbalanced classification problem. Taking a closer look at the only false negative (FN) classification, the eCoach explains that they disagreed with not recommending additional support because the participant indicated an exacerbation of psycho-somatic symptoms (e.g., heart pounding) in their e-mail exchange with the eCoach as well as a lack of future perspective. The FCM does not include psycho-somatic symptoms in its decision-making, but it does include a measure of "lack of future perspective": hopelessness. Four out of the five true positives (TP) occurred in the initial monitoring assessment where additional risk factors are assessed such as a recent inpatient treatment for a psychological condition. Only two of the three participants did exhibit such risk factors and those were also named by the eCoach as reasons why they agreed with the recommendation to seek additional support. The remaining TPs represented moments of elevated emotional suffering as reflected by the participants' monitoring responses in the CRA. The monitoring algorithm's F_1-score was 0.91.

6.3.2 Analysis II: Predictive value CRA for clinical change

Participants

Since we only included participants who had (1) completed the 10-week intervention and (2) filled in the clinical measurements (depression, grief, loneliness) at baseline and ten weeks after starting the intervention, 21 participants were included in the analysis that assesses the predictive value of the CRA. Participants' mean age was 60.1 years (SD 11.4 years). 18 participants were female and 3 were male. Using Spearman-Brown coefficients ρ_{SP} to assess the reliability of the five CRA sub-scales resulted in $\rho_{SP} = 0.74$ for the Hopelessness sub-scale, $\rho_{SP} = 0.70$ for the Grief Symptoms and Therapeutic Progress sub-scales, $\rho_{SP} = 0.66$ for the Psychological Crisis construct, and $\rho_{SP} = 0.14$ for the Social Isolation sub-scale. ρ_{SP} scores between 0.5 and 0.7 are considered fair and scores between 0.7 and 0.9 are considered good (Nutley et al., 2023). Further looking into the low Spearman-Brown coefficient ρ_{SP} for the Social Isolation construct revealed that the two items in the sub-scale correlated poorly (Pearson's $r = 0.08$).

Linear regression analysis

Table 6.2: Descriptives of the independent and dependent variables in the regression analysis. *(n=84, 21/105, 20% missing values)

	Parameter	Mean (SD)	Median [Min,Max]	Scale range
	CRA total*	8.68 (4.16)	9.0 [1.0, 21.0]	0-24
$t1 - t0$	Grief	-2.9 (5.37)	-2.0 [-12.0, 6.0]	5-80
difference	Depression	-0.76 (2.83)	0.0 [-7.0, 5.0]	0-27
(N=21)	Loneliness	-0.29 (1.27)	0.0 [-3.0, 3.0]	0-6

Table 6.2 shows the descriptives of the dependent and independent variables in the regression analysis. Depression and loneliness measurements before the RCT (t_0) and after the RCT (t_1) differed little, making it difficult to reliably fit a model using either as dependent variable. We therefore decided to only conduct one regression analysis with the difference in grief scores before and after the RCT as dependent variable. Not everyone filled in the CRA regularly, resulting in $n = 84$ CRA measurements that were included in the analysis. Figure 6.1 shows a subset of the fitted individual growth curves, the entire set is included in the Appendix. Table 6.3 summarizes the results of the linear regression analysis with difference in grief before and after the RCT as dependent variable and individual CRA growth curves and CRA means as predictors. Overall, the regression model fit the observed data well ($R^2 = 0.45$, $F(df) = 4.42(3, 16)$, $P = .019$). Neither the slope of the individually fitted CRA curves ($R^2 = -1.18$, $SE = 2.21$, $t(df) = -0.54(16)$, $P = .60$), nor their standard error ($R^2 = -4.45$, $SE = 3.03$, $t(df) = -1.47(16)$, $P = .16$) predict how grief symptoms change during the RCT. CRA mean scores do have predictive value for how mourners' grief scores change during the intervention ($R^2 = 1.19$, $SE = 0.33$, $t(df) = 3.58(16)$, $P = .002$). We checked statistical assumptions visually, including normality and homoscedasticity of residuals, and found none to be violated.

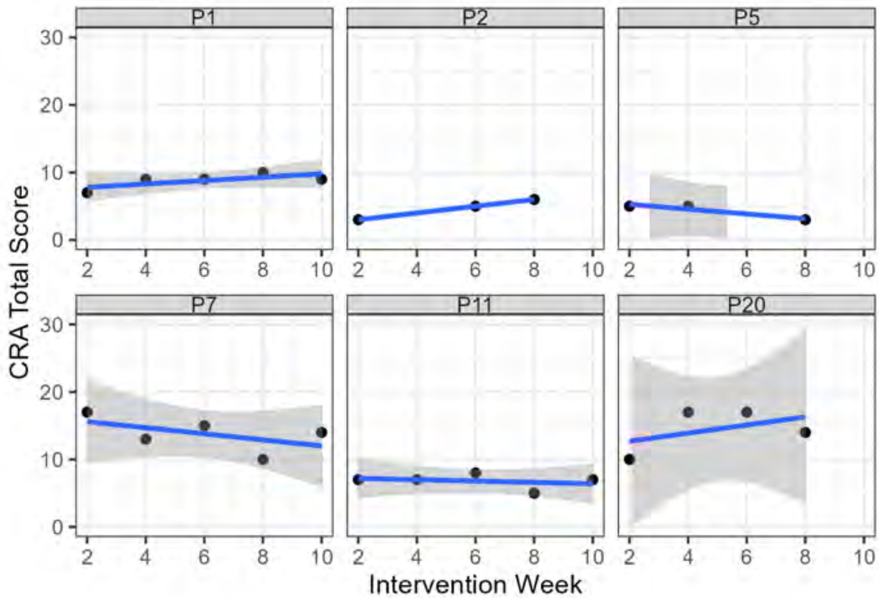


Figure 6.1: Subset of fitted linear individual growth curves that serve as predictor variables in the regression analysis.

Table 6.3: Summary of the linear regression analysis with individually fitted growth curve parameters and CRA mean scores as independent and difference in grief before and after the 10-week online grief intervention as dependent variables.

	R^2	**95% CI**	**t(*df*)**	**P value**
Intercept	-11.37	$[-16.82, -5.92]$	$-4.42(16)$	$< .001$
Slope of the growth curve	1.18	$[-5.87, 3.5]$	$-0.54(16)$	$.60$
Standard error of the growth curve slope	-4.45	$[-10.88, 1.98]$	$-1.47(16)$	$.16$
CRA mean	1.19	$[0.48, 1.89]$	$3.58(16)$	$.002$

6.3.3 Analysis III: eCoaches' experience with the monitoring tool

All eCoaches that provided guidance during the RCT (N=4) participated in the online focus group to discuss their experience with the monitoring tool. The eCoach team consisted of one trained psychological therapist and three final year clinical psychology students who partook in the e-coaching as part of their training. Their mean age was 26.9 years (SD 2.69 years). All eCoaches were female. The final year students were trained to provide e-mail guidance and where closely supervised by a trained psychotherapist.

Before the start of the RCT, the eCoaches discussed how they would use the monitoring tool to check on participants' health regularly. To determine a deterioration in a participant's mental health, the eCoaches took into account the mourner's CRA responses, recommendations suggested by the monitoring algorithm, their impression about the mourner from the clinical interview (e.g., knowledge about the death anniversary date of the deceased), and the mourner's weekly e-mail communication, if available. The eCoaches weighted recent CRA responses most and whether there was a pattern in the response behavior. One eCoach explained that they incorporated the monitoring responses into their weekly guidance e-mails for unresponsive participants to personalize their contact with them. All eCoaches confirmed that they used the monitoring to confirm their existing impressions of participants:

> "I think we regarded it as a kind of a safety option, to check how people are feeling and how it aligns with our impression of the person and the remaining contact we have with them. And for us to reflect, did we overlook anything or forget to ask anything?" (eCoach 1)

The eCoaches unanimously experienced having another source of information as helpful, especially for participants who otherwise communicated little with them during the RCT. The eCoaches experienced being able to monitor participants' progress with the grief service as supportive and reassuring:

> "Whether they [the mourner] made progress or deteriorated, a kind of support for recognizing if anything were to happen. Maybe if they [the mourner] did not tell us, to have another chance at detecting it." (eCoach 3)

Another eCoach agreed that they were curious about what participants responded in the monitoring, especially when mourners worked on intervention content that the eCoaches knew to be challenging for some mourners, such as writing a farewell letter to the deceased spouse.

For future versions of the monitoring tool, the eCoaches suggested providing feedback directly to the mourner, such as a regular written summaries and recommendations to seek additional support in times of crisis. This could support the mourner's reflection about their affective states and encourage them to seek offline support pro-actively instead of waiting until an eCoach advises them to seek support. In addition, the eCoaches expressed they would prefer to receive warning messages for example via SMS whenever the condition of a mourner deteriorates drastically to facilitate immediate intervention.

6.4 Discussion

The current study is situated in the rapidly emerging field of AI tools for mental healthcare and evaluates a monitoring module in an online grief intervention for older mourners with the aim of guiding them to offline support if their mental health deteriorates. We leveraged the insights from three evaluation approaches and encountered three main challenges when trying to come up with satisfactory and clear conclusions about "how well" a monitoring such as the one evaluated in this study performs.

First, many clinical classification problems are (extremely) imbalanced, meaning that the class for which correct classification is crucial (e.g., recommending help-seeking, detecting a malicious tumor, etc.) is underrepresented in real-life data sets (Sun et al., 2009). While there are evaluation metrics, including the F_1-metric (Apostolopoulos & Groumpos, 2023; Paula Branco & Ribeiro, 2016) that we used in the current paper, to mitigate class imbalance to some extent (Raeder et al., 2012), the clinical meaningfulness of obtained results should still be appraised critically. Complementing the evaluation with qualitative accounts from clinical practice is in line with Whiting and Fazel (2019)'s suggestion to consider the incremental benefit of AI tools in clinical practice and stress that any thorough evaluation of a monitoring tool should go beyond quantifiable accuracies and statistics. A monitoring tool that does not match the needs and preferences of its users and the clinical context in which it is used will ultimately not be used, regardless of its classification performance (Sendak et al., 2020). The qualitative evaluation of the tool revealed that the eCoaches envisioned the tool not only as a regular mental health check, but also an emergency detection tool for short-term psychological crisis. A more appropriate approach to evaluating the latter would be to investigate CRA measures around an episode of psychological crisis. However, the low prevalence of psychological crisis in our data makes any evaluation targeted at detecting emergencies impossible. The tool should be evaluated in a clinical sample in which short-term psychological crisis is expected to arise more frequently to investigate its suitability as an emergency detection tool.

Finding appropriate statistical approaches to evaluate AI tools for clinical practice using real-life mental health data represents a second challenge. Despite the mixed results obtained in this study, we argue that statistical approaches that allow for the explicit modelling of individual differences should receive more attention in future evaluations of AI tools in mental healthcare. Individual growth curve predictors are recommended when distinct developmental patterns are expected across outcome groups, in our case, the level of the individual mourner (Welten et al., 2018). Grief and its experienced intensity are inherently individual (Mancini & Bonanno, 2011), suggesting from a clinical point of view that individually optimised growth curves are a suitable means of analysis. In this chapter, individually fitted CRA growth curves captured participants' response patterns variably well. Participants' response patterns may require more complex functions (e.g., quadratic, cubic, etc.) than linearly fitted curves. Another reason why the estimated growth curves fit participants' response patterns variably well are missing values (Welten et al., 2018). The reliability of the underlying measurement tool likewise impacts the fit of estimated growth curves. The Social Isolation sub-scale needs revision since its two items were poorly correlated. The two items capture two different dimensions of being socially isolated: the feeling of being a burden to others and active social withdrawal behavior. It is difficult to reliably assess two dimensions of a construct using only two items. The construction of two-item scales is generally discouraged in terms of reliability (Eisinga et al., 2013), however, to limit the burden of filling in mental health checks regularly as part of a digital mental health service, short self-assessment tools are needed. In this context, incorporating less obtrusive assessment methods in digital mental health services, including sensing technologies (Abdullah & Choudhury, 2018) and natural language processing (Malgaroli et al., 2023), to complement self-report monitoring of clients' mental states should be considered in the future.

To move towards well developed monitoring systems in eMental health, we recommend clear and early decision-making about a) the responsibilities of the monitoring tool in the eMental health application (and which responsibilities the tool does not have) and b) what it takes to evaluate the tool in a satisfactory way so that it can live up to these responsibilities and contribute in a meaningful way to clinical practice. Currently, AI tools are often developed as secondary goals to the development of a new eMental health application (Tornero-Costa et al., 2023), which represents the third identified challenge since it limits time and effort invested into their development and evaluation. Extracting clinically meaningful results using common methods for evaluating AI tools is complex. Hence, such tools cannot afford ambiguities regarding their capabilities and responsibilities that further complicate the evaluation process.

6.4.1 Limitations

The current study has limitations. First, the eCoaches that provided the ground truth labels for assessing the classification performance of the monitoring tool had access to the tool's suggested decisions at the time of labelling participants' monitoring response patterns as either "advisable to seek support" vs. "fine to continue using the grief service as is". Having access to the tool's suggestions may have biased the eCoaches' ground truth labels in favor of the monitoring tool. However, the ground truth labels were provided by the coaches upon request at the time of conducting the analyses in this study, after RCT participants that were included in this study had completed the 10-week grief intervention. Therefore, in practice, the eCoaches revisited participants' monitoring responses retrospectively, as well as their own initial decision-making during participants' participation in the RCT. The retrospective nature of the labelling task likely limited the potentially introduced bias since the eCoaches had the knowledge of the participants' trial outcome at their disposal, solidifying the truth of the provided labels. With regard to providing clinically meaningful insights, a second limitation of the current study is the small sample size in the regression analysis and the small number of psychological crisis events in the classification evaluation analysis that we mentioned earlier. Any (clinical) conclusions based on the obtained results should be drawn with caution.

6.5 Conclusion

In recent years, the demand for high-quality and accessible mental healthcare has been increasing. Digital mental health self-help services have potential to support today's healthcare systems in meeting care demands and their potential is further increased by leveraging the benefits of emerging AI tools, including monitoring tools that track users' affective states and guide them towards offline support if their mental health warrants professional intervention. Such AI tools come with challenges that must be addressed systematically before they have an impact in clinical practice. These challenges include finding meaningful evaluation approaches in the face of (1) (extremely) imbalanced real-life clinical data sets, (2) ambiguous demands and expectations regarding the capabilities and responsibilities of such tools in eMental health and (3) priority misalignments between evaluation approaches for AI tools and the overarching goals of clinical trials in which their evaluation is usually embedded. We hope to contribute to an enhanced awareness about these challenges and thereby to developing evaluation approaches for AI tools in eMental health that facilitate their introduction in clinical practice.

General discussion

Prologue

Orphea puts on an instrumental piece of music that she has become fond of in the past months. Two or three modules into the e-learning, Orphea had made a habit of listening to motion picture soundtracks of movies that she and Justus had watched together in the past. This will be a difficult final exercise and Orphea welcomes the familiarity of this simple, but surprisingly soothing little ritual while she starts writing on the white sheet of paper in front of her: “Dear Justus, my love, ...farewell.".

In the past chapters of this dissertation, we have summarized the research involved in developing a system to monitor mourners' mental health with the aim of detecting pathological tendencies in the context of an online self-help grief service for older adults who have lost their spouse. The aim of such a system is to inform decision-making about stepping up to more intensive bereavement support if necessary. This dissertation informed the development of such a system in five research steps:

1. The selection of reliable indicators for people's mental health and changes thereof in the context of an online self-help grief service (chapters 2 and 4).
2. The definition of a measurement protocol including decisions about the frequency of measurements, assessment tools, and the processing of collected monitoring data for decision-making about recommending mourners to seek more intensive bereavement support (chapter 4).
3. The selection and configuration of an automatic decision-making component (chapters 3 and 4).
4. Exploring a strategy for communicating the results of the monitoring to mourners and/or care professionals (chapter 5).
5. A thorough evaluation of the developed monitoring system with regard to user acceptance and experience, accuracy of algorithmic decision-making and clinical relevance (chapters 5 and 6).

This dissertation thereby provides a blueprint for the development of monitoring systems in eMental health and exemplifies a research path to inform the development of such systems. The steps that we have taken as well as the lessons we have learned along the way are by no means exhaustive. However, this dissertation can serve fellow researchers that aim to develop similar monitoring systems as a starting point. In this general discussion, we will give an overview of the key results and insights that we gained regarding each of the above mentioned steps and we will put our findings into a broader context, discussing some limitations, challenges and opportunities for future work.

In chapter 2, we consulted an expert-panel in grief and eHealth using the Delphi method with two goals: (1) arriving at a set of user characteristics that we can monitor to inform decision-making about a) whether the mourner should consider seeking more intensive bereavement support even before starting the online grief service and b) to detect if the user's mental health deteriorates to the extent that professional intervention is warranted once they start using the service. (2) The second aim of the research documented in chapter 2 was to explore a personalization strategy for adapting the content of the online grief service to the needs of the mourner. This chapter contributes a ranked set of monitoring parameters for monitoring grief symptoms, and the aggregation of several ways of adapting the content of eMental health services in a conceptual personalization model. We also critically reflect on the use of the Delphi method for early personalization research in eHealth.

Chapter 3 introduces fuzzy cognitive mapping (FCM), a soft computing technique that can be used to build explainable decision models based on expert knowledge. The chapter explains the theoretical and technical foundations of the decision-making algorithm presented in chapter 4. The chapter does not represent a new technical contribution to the field of FCM. However, it explores a new application area for FCMs while providing fellow researchers with theoretical and practical guidelines for using the method for researching decision-making systems in the field of eMental health. The transparency, interpretability, ability to handle uncertainty, and independence of large training data sets to arrive at an initial model configuration make FCM a promising methodology for developing algorithms for automatic decision-making in mental healthcare. By critically reflecting on the limitations of FCM for the mental health domain, including the lack of a straightforward way to include a client's clinical history in the model and important implications of the mathematical foundations of the technique, we also identify areas of future work for applying FCM to the mental health domain.

Chapter 4 continues the applied research towards the development of a monitoring system to inform decision-making about recommending more intensive bereavement support in the LEAVES online grief service. Based on the results from chapter 2, it presents the measurement protocol, assessment tools, and a FCM-based decision algorithm, alongside a first scenario-based evaluation of the developed monitoring module.

In chapter 5, the perspectives of grief professionals and older mourners about the LEAVES monitoring system and recommendations that the system gives about help-seeking are integrated into design principles for help-seeking recommendations targeted at older adults. The main insight of this chapter is that to make support on-demand work it should be a core feature of the entire online service, including preparing older mourners practically for help-seeking should the need arise and increasing their mental health literacy, including fostering a positive attitude towards seeking mental health support. This requires that the service prepares users to take action in times of need well before the need arises by informing them about support options in their physical vicinity, clinical or otherwise, and encouraging them to prepare actionable steps they can easily follow.

Finally, in chapter 6, based on a three-fold approach, we evaluate the LEAVES monitoring system by assessing a) the accuracy of the decisions made by the FCM model using the F_1-metric, b) how well its assessment tools capture clinical changes in grief over the course of a 10-week intervention, and c) the system's perceived usefulness by eCoaches who used it to inform the guidance they provided to mourners during the 10-week intervention. We find that the developed monitoring system performs reasonably well using standard machine learning evaluation metrics, but we critically question the meaningfulness of commonly used evaluation metrics and strategies in machine learning and clinical trials for evaluating algorithmic decision-making in eMental health. Chapter 6 rounds off the research involved in the development of the monitoring system for the LEAVES online grief service.

7.1 Stepping up in preventative eMental health for grief

The research in this dissertation has been applied for the development of a monitoring system that facilitates support-on demand, a form of stepped care that recommends help-seeking based on mental health monitoring, but does not enforce stepping up care unless the client takes the initiative. Stepped care has been investigated in various forms in mental healthcare for years (Cornish, 2020; O'Donohue & Draper, 2010). As a response to the increasing crisis of access in mental healthcare where the demand for mental health support exceeds treatment capacity, stepped care addresses the inefficient and unsustainable allocation of treatment resources including people-intensive and expensive psychotherapy and psychopharmacological treatment. In stepped care, higher intensity treatment usually involves professional care. With the rise of eHealth, digital self-help interventions have been added to the lower intensity repertoire of stepped care and their effectiveness for treating mild to moderately severe mental health problems has been documented (O'Donohue & Draper, 2010). For example, for major depressive disorder, O'Donohue and Draper (2010) suggest 1. group therapy, 2. individual outpatient therapy, 3. medication, and 4. inpatient therapy as higher intensity treatment options *after* self-guided therapy using eHealth has been exhausted.

The online grief service presented in this dissertation represents a low-intensity support option for mourners, specifically targeted at preventing the exacerbation of grief symptoms. An insight we gained in this dissertation is that support options outside the clinical healthcare system can be viable, low-threshold alternatives for stepping up support from online self-help. For some clients and mental health problems, including non-clinical higher intensity support options may even be preferable. Specifically, the idea of facilitating and encouraging connecting with bereavement support offered by individuals, communities, and services in the mourner's physical vicinity was suggested as a step up option in chapter 4 of this dissertation. The advantage of broadening step-up options to include non-clinical support is that often, alternative bereavement support is already available, such as in the form of peer support groups and volunteer organizations. Grief support groups may vary in whether they are lead by a peer or a professional and the extent to which they embrace a certain philosophical or spiritual viewpoint about death and grief (O'Donohue & Draper, 2010). Finding a 'good fit' can be challenging and therefore, a valuable addition to the content of an online grief service to facilitate the expansion of the mourner's support network and finding a safe space where they can exchange experiences and find comfort. Ultimately, returning to Breen et al. (2022)'s notion of 'grief-denying' societies, including non-clinical steps in stepped care for grief can contribute to the deinstitutionalization and deprofessionalization of bereavement care, paving the way towards a more grief-literate society.

The mental healthcare landscape, including the way it implements stepped care, needs to evolve with current societal changes. Part of the necessity for change relates to a cornerstone assumption of stepped care. In its original form, stepped care assumes that high intensity treatment is the best solution for clients with more severe symptoms whereas low intensity treatment does not work for them. One consequence of this specific assumption is that stepped care policies have focused on saving treatment options that involve expensive professional care for higher intensity levels. However, recently, this assumption has been critically questioned (Cornish, 2020) as evidence has accumulated that low intensity digital mental health support benefits everybody, regardless of symptom severity (Aoun et al., 2018; Bower et al., 2013; Breen et al., 2022). An understanding has emerged that allocating treatment based on symptom severity and complexity alone is unjustified as other factors are equally important to consider when it comes to choosing optimal treatment for clients. Cornish (2020) explains that in recent years, stepped care has moved from being "client-centered" to being "client-centric", meaning that mental healthcare is gradually moving towards allocating mental health support based on client preferences, goals and strengths. "Client-centric" healthcare focuses on enabling clients to make informed decisions about their own treatment, rather than solely and unknowingly relying on decision-making by experts about what is best for them. We have seen the same trend towards incorporating client preferences and strengths in this dissertation where mourners requested to remain in control about involving a professional in their grief process (chapter 5). In addition, grief professionals pointed out that the LEAVES fuzzy cognitive map decision model focuses too much on negative symptoms, while neglecting the influence of other factors, such as positive reminiscing, on decision-making about stepping up bereavement care (chapter 6). These two examples from this dissertation, as well as the earlier mentioned idea of including non-clinical options in stepped care applied to bereavement support, reflect a societal movement towards "client-centric" mental healthcare that at the same time requires and is fueled by an increase in mental health literacy in the general population.

7.2 Algorithmic decision-making in eMental health

Decision-making by non-human actors, including algorithms, is an essential item on the research agenda as we move towards client-centricity in stepped care. Employing preventative mental health support via stand-alone digital services has implications for guaranteeing that this form of mental health-care remains self-corrective (Cornish, 2020). In the context of stepped care, “self-corrective" means that higher intensity treatment in terms of stakeholder involvement will be offered to those who need and want it based on regular check-ups and progress assessments. The research in this dissertation works towards techniques that enable stepping up from preventative support to active treatment when no healthcare professional is involved in the decision-making. Instead, we are moving towards digital services that provide the necessary information for clients to make informed decisions about their treatment options. This is a plausible scenario for a preventative online service. When we use algorithms and other artificial intelligence tools to inform decision-making in this context, it is of paramount importance that they are not only reasonably accurate, but indeed supporting the client in their decision-making.

There has been a rapid growth in studies that explore artificial intelligence (AI) in the field of mental health (Thieme et al., 2023). AI has been applied (1) for mental health diagnosis, symptom or risk detection (Nobles et al., 2018), often using sensor and text data, (2) to improve treatment access and delivery, including chatbots and other conversation-based approaches (Fitzpatrick et al., 2017), and (3) for AI-based decision-support for treatment selection (Chikersal et al., 2020; Delgadillo et al., 2022). This dissertation contributes a thorough initial exploration of a specific type of automatic decision model, fuzzy cognitive mapping (FCM), to the ongoing research field of algorithmic decision-making in mental health. On top of that, by integrating a form of automatic decision-making in an online mental health service where potentially no human professional is involved in decision-making, our research represents a novel application of AI-based decision-support for stepping up mental health support. Whereas AI-based mental health decision-support for healthcare professionals has attracted some interest in recent years (e.g., Delgadillo et al. (2022)), decision-support that informs client decision-making about treatment options represents a relatively uncharted field of research. The trend towards client-centric approaches, where clients take more responsibility for their own treatment, raises new questions about how digital mental health services can provide clients with the necessary knowledge and skills to make good decisions. Part of providing clients with the necessary information is to give the client an overview of their treatment options and to point out costs, requirements in terms of engagement and time, and outcome prospective (Cornish, 2020). But how exactly does mental health decision-support targeted at clients differ from decision-support targeted at healthcare professionals? In client decision-support, people face difficult and impactful decisions about their health while potentially being mentally and emotionally at their worst. Unfortunately, we still know little about how specific mental health issues and emotional vulnerability influence people's capacity to make good health decisions (Bishop & Gagne, 2018; Hindmarch et al., 2013). Consequently, more research is needed about how clients' mental and emotional vulnerability factors into the design of client decision-support, specifically about stepping up care in preventative self-help online mental health services.

Another interesting question about designing client decision-support is in what way client-centric systems differ regarding requirements about transparency and explainability. Given clients' emotional vulnerability, we anticipate a more pronounced need for emotional and practical guidance regarding stepping up mental health support, whereas technical details about how a decision model arrives at a recommendation may be less relevant to disclose in detail. Nevertheless, transparency is essential for the sake of persuading the client to take the initiative in seeking mental health support (Tielman et al., 2019), and to empower a client's inquiry for higher intensity treatment when they reach out to a mental healthcare professional. For example, by providing comprehensive empiric evidence gathered by regular mental health check-ups or continuous monitoring of relevant physiological symptoms, such as measurements of sleep quality and self-report tracking of other somatic symptoms, the self-help mental health service can support a client's case for stepping up mental health treatment. In sum, to advance the development of algorithmic decision-making in preventative online mental health services, we must look beyond the often isolated design and evaluation of AI-based decision models and take a closer look at the interactions between client characteristics, including their vulnerabilities, strengths and preferences, and the mental healthcare system as a whole to develop services that support people in making well-informed mental health decisions, and to support clients in taking the initiative in seeking mental health support (chapter 5 of this dissertation).

7.3 Limitations and future work

The research path that we have paved in this dissertation for the development of a monitoring system for stepping up care in an online grief service has limitations and some of those we would like to acknowledge and highlight at this point. Often times, fortunately, limitations of past research efforts are simultaneously a source for inspiration and a starting point for future work.

First, while applying the traditional form of fuzzy cognitive mapping (FCM) to the LEAVES monitoring system, methodological advances in the field of FCM continued in parallel that we did not take into account while developing our own FCM-based decision algorithm (chapter 4). We mention these recent advances in our introduction to FCMs (chapter 3), in the hope that the interested reader and fellow researcher who wishes to apply FCM in their own research will look beyond this dissertation to inform themselves about FCMs. In hindsight and in light of the recent methodological advances, there were alternatives for the configuration (chapter 4) and evaluation (chapter 5 and 6) of the LEAVES FCM that we did not mention explicitly in the respective chapters of this dissertation. These include the exploration of alternative reasoning rules that are independent of sigmoid and hyperbolic tangent activation and a thorough simulation evaluation of the model's dynamic behavior as suggested by Nápoles et al. (2024). Based on the evaluation that we did conduct using scenarios (chapter 4) and the qualitative and quantitative data of real mourners (chapter 5 and 6), we have been able to offer some insight into the utility and accuracy of the LEAVES decision-making algorithm. However, as we point out in chapter 6, it remains challenging to draw clinically meaningful conclusions using common evaluation approaches in the field of artificial intelligence, or in mental healthcare. This is also why eventually, we opted for combining several evaluation approaches. In particular, our chosen evaluation approach for the LEAVES FCM, including scenarios and exploring F_1 scores, offers some confidence that the model does not converge to a unique fixed point attractor, in which case no real decision-making would take place. Still, more insight into the model's dynamic behavior would be desirable for future applications of the algorithm described in this thesis. By extending our evaluation approach with an investigation of the FCM's dynamic behavior, we could have added one more criterion to base our evaluation on. It remains for future work, however, to examine to what extent this FCM-specific evaluation technique will contribute to generating clinically meaningful evaluation metrics for FCMs in the mental health domain.

The discontinuation of our research about personalization strategies for online grief self-help services for older adults represents a second limitation of this dissertation. Chapter 2 of this dissertation had two goals, (1) determining monitoring parameters for decision-making about recommending help-seeking in an online grief service and (2) defining a personalization strategy for adapting such an online service to the needs of the user. While we met both aims in chapter 2, due to practical considerations relating to the overarching goals and requirements of the research project in which our research was situated, we discontinued further investigation and validation of the conceptual personalization model in the remainder of this dissertation. There is some conceptual overlap between these two initial research endeavours (monitoring and personalization) as adapting treatment options based on symptom severity represents one form of personalising an online self-help service to the user. The conceptual distinction between initial and dynamic adaptations likewise resurfaces in our later distinction between an initial and a continuous risk assessment for monitoring mental health (chapter 4). One can further argue that regular monitoring is a pre-requisite for dynamic tailoring of therapeutic content and that trends towards client-centric decision-making in stepped care also apply to personalization, where client strengths and goals can be a starting point for personalizing therapeutic content (Burley et al., 2020). To be more specific, we consider our insights about the dynamics of user parameters that form the foundation for decision-making, automatic decision-making models and their challenges, and the importance of including the client in decision-making equally relevant for future exploration of the conceptual personalization model that we proposed in chapter 2. In addition, we encourage other researchers to explore the suitability of fuzzy cognitive mapping for building personalization algorithms and to put our insights regarding the challenges of drawing (clinically) meaningful conclusions about algorithmic decision-making in eMental health into practice, be it for monitoring or for personalization purposes.

7.4 Conclusion

At some point, life confronts most people with the loss of an acquaintance, a friend, or a family member and when it does, we are struck with loss and grief. Despite being a common occurrence, we appear to live in a “grief-denying" society where people tend to feel overwhelmed by comforting colleagues, friends, and family members and advice about how to be compassionate in the face of another’s suffering is scarce. Unfortunately, this can isolate mourners socially and emotionally. Where one’s support network and own coping strategies are insufficient, grief can develop into a serious mental health problem. eMental health applications can help to prevent the exacerbation of grief symptoms into mental illness by a) offering information about the grief process, b) supporting mourners in the process of adapting to a life without the deceased, and c) detecting maladaptive grief coping and the exacerbation of grief symptoms into pathological tendencies that warrant more intensive forms of bereavement support. The research summarized in this dissertation has informed the development of an automatic risk detection system to inform client decision-making about stepping up bereavement support from online self-help to offline (professional) support. The results of this dissertation contribute to the ongoing research field of algorithmic decision-making in eMental health, offering insights about the suitability of a specific form of decision-model, fuzzy cognitive maps, and practical advice about how to implement it, while raising questions and unravelling challenges linked to the integration of algorithmic decision-making into online mental health self-help services.

Epilogue

"...so there is pain, so much pain. But the pain starts to make space for... gratitude. Gratitude for the good years that we had, the difficult times that we somehow managed to get through, gratitude for the space and support that you gave me throughout the years. I can feel that slowly, very, very slowly gratitude is sneaking into tiny little bits of space that were burned out by pain and grief.". Orphea puts down her pen and sits back. The motion picture soundtrack she had put on finished playing hours ago. As she turns towards the window she notices the first rays of sunlight breaking through the dark of the morning. It seems to be a chilly, but sunny December day. "This is not the end, this is not farewell, not really.", she thinks as she pushes back from the desk. She turns off the computer and walks across the room while she muses "Something is different, I am different. I will always miss having you by my side, always. Being able to exchange thoughts, laugh together, to hug and touch and kiss. Somehow - it is difficult to put into words - it feels as if we have started a new conversation, you and I. And I have never shared this kind of conversation with anyone else before."

&

Appendix: Monitoring questionnaires

Initial risk assessment questionnaire (IRA)

Ratings in brackets

Time passed since spousal loss

1. How much time has passed since you lost your partner?

☐ Less than a month (0)
☐ Between 1 to 6 months (1)
☐ Between 7 to 12 months (2)
☐ More than 12 months ago (3)

Characteristics of the loss

2. Did the circumstances of your loss involve any of the following
   - Accident
   - Suicide
   - Violence

☐ Yes (1)
☐ No (0)

Inpatient treatment

3. Have you received treatment for a psychological condition in an inpatient clinic in the past?

☐ Yes, in the past year (1)
☐ Yes, more than a year ago (0)
☐ No (0)

Technical skills

4. How would you describe your skills with a computer and the internet? (*Ratings in parentheses*)

☐ Very poorly, I avoid using the computer as much as possible. (0)
☐ Poor, I often need assistance to find my way around the computer and the internet. (1)
☐ Fair, I can perform standard actions on the computer *without* assistance, for example sending and reading e-mails, printing, or managing my bank account online. (2)
☐ Good, I perform standard actions regularly on the computer and occasionally use it for other purposes (e.g., playing games, booking a vacation, online shopping. (3)
☐ Very good, I find my way around the computer and the internet very easily and use both regularly for standard business and other purposes. (4)

Suicidality

1. In the past 2 weeks, I have contemplated to commit suicide.

   ☐ Yes (1)
   ☐ No (0)

   If yes, adapted Scale for Suicide Ideation (SSI) items 12 through 16 assess plans, preparations, and resolve to make a suicide attempt:

   2. In the past 2 weeks, I have considered a plan for a suicide attempt.

      ☐ No, I have not worked out a plan. (0)
      ☐ Yes, but the details of the plan are not worked out. (1)
      ☐ Yes, and the details of the plan are well worked out. (2)

   3. In the past 2 weeks, I felt that there was an opening for a suicide attempt or a suicide plan that I thought about was available.

      ☐ The contemplated plan is not available; there is no opportunity. (0)
      ☐ The contemplated plan would take time and/or effort and it is not readily available. I don't see an opportunity. (1)

Initial risk assessment questionnaire (IRA)

Ratings in brackets

Time passed since spousal loss

1. How much time has passed since you lost your partner?

☐ Less than a month (0)
☐ Between 1 to 6 months (1)
☐ Between 7 to 12 months (2)
☐ More than 12 months ago (3)

Characteristics of the loss

2. Did the circumstances of your loss involve any of the following
    - Accident
    - Suicide
    - Violence

☐ Yes (1)
☐ No (0)

Inpatient treatment

3. Have you received treatment for a psychological condition in an inpatient clinic in the past?

☐ Yes, in the past year (1)
☐ Yes, more than a year ago (0)
☐ No (0)

Technical skills

4. How would you describe your skills with a computer and the internet? (*Ratings in parentheses*)

☐ Very poorly, I avoid using the computer as much as possible. (0)
☐ Poor, I often need assistance to find my way around the computer and the internet. (1)
☐ Fair, I can perform standard actions on the computer *without* assistance, for example sending and reading e-mails, printing, or managing my bank account online. (2)
☐ Good, I perform standard actions regularly on the computer and occasionally use it for other purposes (e.g., playing games, booking a vacation, online shopping. (3)
☐ Very good, I find my way around the computer and the internet very easily and use both regularly for standard business and other purposes. (4)

Suicidality

1. In the past 2 weeks, I have contemplated to commit suicide.

    ☐ Yes (1)
    ☐ No (0)

    If yes, adapted Scale for Suicide Ideation (SSI) items 12 through 16 assess plans, preparations, and resolve to make a suicide attempt:

    2. In the past 2 weeks, I have considered a plan for a suicide attempt.

        ☐ No, I have not worked out a plan. (0)
        ☐ Yes, but the details of the plan are not worked out. (1)
        ☐ Yes, and the details of the plan are well worked out. (2)

    3. In the past 2 weeks, I felt that there was an opening for a suicide attempt or a suicide plan that I thought about was available.

        ☐ The contemplated plan is not available; there is no opportunity. (0)
        ☐ The contemplated plan would take time and/or effort and it is not readily available. I don't see an opportunity. (1)

- □ The contemplated plan is available or there is another opportunity to commit suicide. (2)
- □ I anticipate that in the future, the contemplated plan will be available or there will be an opportunity. (2)

4. In the past 2 weeks, I have considered how capable I am to commit suicide.
   - □ I am certain that I have **no** courage and/or **no** competence to commit suicide. (0)
   - □ I am uncertain of my courage and/or competence to commit suicide. (1)
   - □ I am certain that I have courage and/or competence to commit suicide. (2)

5. In the past 2 weeks, I have engaged in actual preparation for the contemplated suicide plan.
   - □ No, I did no preparation at all. (0)
   - □ I started to prepare the contemplated suicide plan. (2)
   - □ I have completed preparing the contemplated suicide plan. (2)

Continuous risk assessment questionnaire (CRA)

Crisis detection

1. In the past 2 weeks, how often have you felt impaired in your daily life or experienced severe emotional distress which made it difficult to think of other things?
   - ☐ Not at all
   - ☐ A few days
   - ☐ More than half the days
   - ☐ (Nearly) every day

2. In the past 2 weeks, how often have you felt that you need extra support to cope with everyday life or getting through the day?
   - ☐ Not at all
   - ☐ A few days
   - ☐ More than half the days
   - ☐ (Nearly) every day

Depressive symptom: Hopelessness

1. In the past 2 weeks, how often have you given up because your future feels dark, and seems to only get darker?
   - ☐ Not at all
   - ☐ A few days
   - ☐ More than half the days
   - ☐ (Nearly) every day

2. In the past 2 weeks, how often have you felt that all that awaits you in your future is emptiness, loneliness or suffering and that there is nothing you can do about it?
   - ☐ Not at all
   - ☐ A few days
   - ☐ More than half the days
   - ☐ (Nearly) every day

Grief symptoms

1. In the past 2 weeks, how often have you felt impaired in your daily activities by intense feelings of emotional pain or suffering related to your grief?
   - ☐ Not at all
   - ☐ A few days
   - ☐ More than half the days
   - ☐ (Nearly) every day

2. In the past 2 weeks, how often have you felt impaired in your daily activities by feeling stunned or dazed by your loss?
   - ☐ Not at all
   - ☐ A few days
   - ☐ More than half the days
   - ☐ (Nearly) every day

Suicidality

1. In the past 2 weeks, I have contemplated to commit suicide.
   - ☐ Yes
   - ☐ No

   If yes, adapted Scale for Suicide Ideation (SSI) items 12 through 16 assess plans, preparations, and resolve to make a suicide attempt:

2. In the past 2 weeks, I have considered a plan for a suicide attempt.
   - ☐ No, I have not worked out a plan.
   - ☐ Yes, but the details of the plan are not worked out.
   - ☐ Yes, and the details of the plan are well worked out.

3. In the past 2 weeks, I felt that there was an opening for a suicide attempt or a suicide plan that I thought about was available.
   - ☐ The contemplated plan is not available; there is no opportunity.
   - ☐ The contemplated plan would take time and/or effort, and it is not readily available. I don't see an opportunity.
   - ☐ The contemplated plan is available or there is another opportunity to commit suicide.
   - ☐ I anticipate that in the future, the contemplated plan will be available or there will be an opportunity.

4. In the past 2 weeks, I have considered how capable I am to commit suicide.
   - ☐ I am certain that I have **no** courage and/or **no** competence to commit suicide.
   - ☐ I am uncertain of my courage and/or competence to commit suicide.
   - ☐ I am certain that I have courage and/or competence to commit suicide.

5. In the past 2 weeks, I have engaged in actual preparation for the contemplated suicide plan.
   - ☐ No, I did no preparation at all.
   - ☐ I started to prepare the contemplated suicide plan.
   - ☐ I have completed preparing the contemplated suicide plan.

Social isolation

1. In the past 2 weeks, how often have you been thinking that the people in your life would be better off if you were gone?
   - ☐ Not at all
   - ☐ A few days
   - ☐ More than half the days
   - ☐ (Nearly) every day
2. In the past 2 weeks, I have avoided getting in touch with friends and family.
   - ☐ Not at all
   - ☐ A few days
   - ☐ More than half the days
   - ☐ (Nearly) every day

Therapeutic progress

1. In the past 2 weeks, I have had positive experiences in addressing my problems and/or I have gained important insights.
   - ☐ Very strongly disagree
   - ☐ Disagree
   - ☐ Agree
   - ☐ Very strongly agree

2. In the past 2 weeks, I generally have been feeling better.
   - ☐ Very strongly disagree
   - ☐ Disagree
   - ☐ Agree
   - ☐ Very strongly agree

Appendix: Linear regression individual growth curves (Chapter 6)

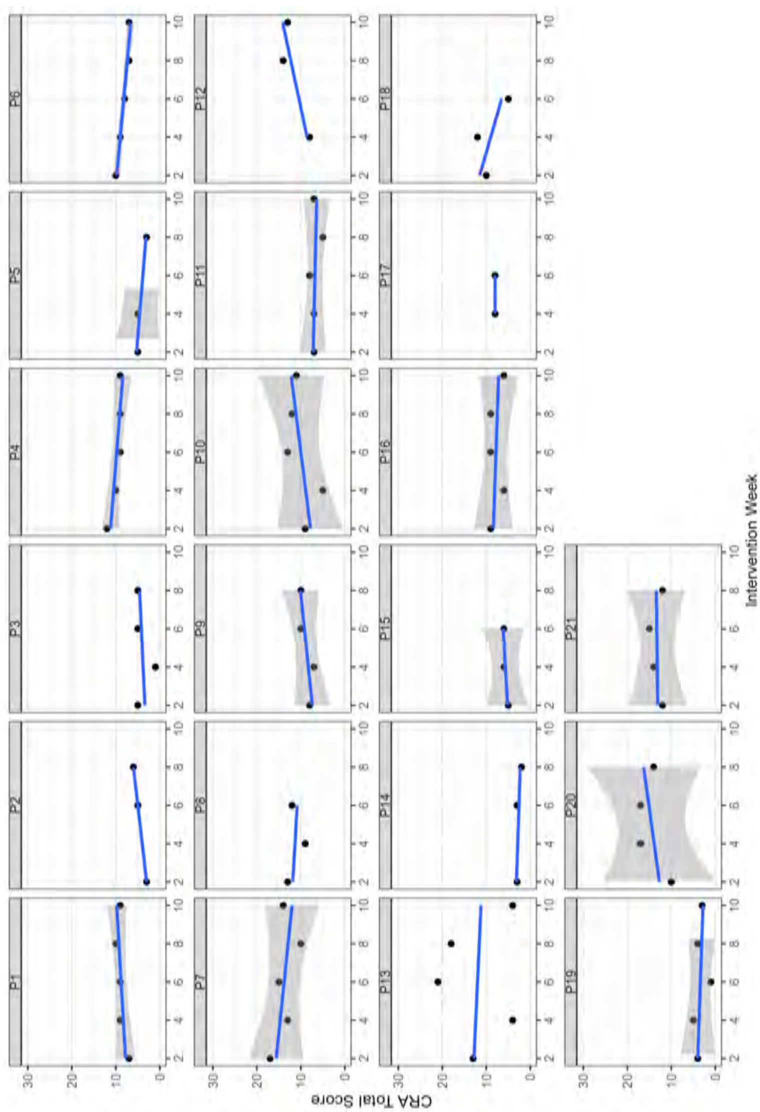


Figure 1: Fitted linear individual growth curves with the intervention week as predictor and continuous risk assessment (CRA) total scores as dependent variable. Growth curves serve as predictor variables in the regression analysis.

References

Chapter 1

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: Dsm-5* (Vol. 5). American psychiatric association Washington, DC.

Aoun, S. M., Breen, L. J., White, I., Rumbold, B., & Kellehear, A. (2018). What sources of bereavement support are perceived helpful by bereaved people and why? Empirical evidence for the compassionate communities approach. *Palliative Medicine*, *32*(8), 1378–1388. https://doi.org/https://doi.org/10.1177/0269216318774995

Axelrod, R. (2015). *Structure of decision: The cognitive maps of political elites*. Princeton University Press.

Borghouts, J., Eikey, E., Mark, G., De Leon, C., Schueller, S. M., Schneider, M., Stadnick, N., Zheng, K., Mukamel, D., & Sorkin, D. H. (2021). Barriers to and facilitators of user engagement with digital mental health interventions: Systematic review. *Journal of Medical Internet Research*, *23*(3), e24387.

Brandl, L., van Velsen, L., Brodbeck, J., Jacinto, S., Hofs, D., & Heylen, D. (2023). Developing an emental health monitoring module for older mourners using fuzzy cognitive maps. *Digital Health*, *9*, 20552076231183549.

Breen, L. J., Kawashima, D., Joy, K., Cadell, S., Roth, D., Chow, A., & Macdonald, M. E. (2022). Grief literacy: A call to action for compassionate communities. *Death Studies*, *46*(2), 425–433. https://doi.org/https://doi.org/10.1080/07481187.2020.1739780

Brodbeck, J., Berger, T., Biesold, N., Rockstroh, F., & Znoj, H. J. (2019). Evaluation of a guided internet-based self-help intervention for older adults after spousal bereavement or separation/divorce: A randomised controlled trial. *Journal of Affective Disorders*, *252*, 440–449. https://doi.org/10.1016/j.jad.2019.04.008

Brodbeck, J., Jacinto, S., Gouveia, A., Mendonça, N., Madörin, S., Brandl, L., Schokking, L., Rodrigues, A. M., Gonçalves, J., Mooser, B., et al. (2022). A web-based self-help intervention for coping with the loss of a partner: Protocol for randomized controlled trials in 3 countries. *JMIR research protocols*, *11*(11), e37827.

Carlbring, P., Maurin, L., Törngren, C., Linna, E., Eriksson, T., Sparthan, E., Strååt, M., von Hage, C. M., Bergman-Nordgren, L., & Andersson, G. (2011). Individually-tailored, internet-based treatment for anxiety disorders: A randomized controlled trial. *Behaviour Research and Therapy*, *49*(1), 18–24.

Christensen, H., Griffiths, K. M., Farrer, L., et al. (2009). Adherence in internet interventions for anxiety and depression: Systematic review. *Journal of Medical Internet Research*, *11*(2), e1194.

Clement, S., Schauman, O., Graham, T., Maggioni, F., Evans-Lacko, S., Bezborodovs, N., Morgan, C., Rüsch, N., Brown, J. S., & Thornicroft, G. (2015). What is the impact of mental health-related stigma on help-seeking? a systematic review of quantitative and qualitative studies. *Psychological Medicine*, *45*(1), 11–27.

Dahlin, M., Carlbring, P., Håkansson, A., & Andersson, G. (2020). Internet-based self-help using automatic messages and support on demand for generalized anxiety disorder: An open pilot study. *Digital Psychiatry*, *3*(1), 12–19.

Elshaikh, U., Sheik, R., Saeed, R. K. M., Chivese, T., & Alsayed Hassan, D. (2023). Barriers and facilitators of older adults for professional mental health help-seeking: A systematic review. *BMC Geriatrics*, *23*(1), 516.

Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review. *BMC Psychiatry*, *10*(1), 1–9.

Hurmuz, M. Z., Brandl, L., Siderakis, E., Schokking, L., Jacinto, S., Verdam, M., & Jansen-Kosterink, S. (2023). Can we support mourning older adults in their grief process with an e-mental health intervention? (*Preprint)*. *Research Square*. https://doi.org/https://doi.org/10.21203/rs.3.rs-3185131/v1

Jagayat, J. K., Kumar, A., Shao, Y., Pannu, A., Patel, C., Shirazi, A., Omrani, M., & Alavi, N. (2024). Incorporating a stepped care approach into internet-based cognitive behavioral therapy for depression: Randomized controlled trial. *JMIR Mental Health*, *11*(e51704), 1. https://doi.org/10.2196/51704

Leung, C., Pei, J., Hudec, K., Shams, F., Munthali, R., & Vigo, D. (2022). The effects of nonclinician guidance on effectiveness and process outcomes in digital mental health interventions: Systematic review and meta-analysis. *Journal of Medical Internet Research*, *24*(6), e36004.

Macdonald, M. E. (2019). The denial of grief: Reflections from a decade of anthropological research on parental bereavement and child death. In *Exploring grief* (pp. 125–139). Routledge.

Meurling, J., Rondung, E., Leiler, A., Wasteson, E., Andersson, G., Richards, D., Shahnavaz, S., & Bjärtå, A. (2023). An online tiered screening procedure to identify mental health problems among refugees. *BMC Psychiatry*, *23*(1), 7.

Mohr, D. C., Burns, M. N., Schueller, S. M., Clarke, G., & Klinkman, M. (2013). Behavioral intervention technologies: Evidence review and recommendations for future research in mental health. *General Hospital Psychiatry*, *35*(4), 332–338.

Oromendia, P., Orrego, J., Bonillo, A., & Molinuevo, B. (2016). Internet-based self-help treatment for panic disorder: A randomized controlled trial comparing mandatory versus optional complementary psychological support. *Cognitive Behaviour Therapy*, *45*(4), 270–286.

Pagliari, C. (2007). Design and evaluation in ehealth: Challenges and implications for an interdisciplinary field. *Journal of Medical Internet Research*, *9*(2), e614.

Richards, D., & Richardson, T. (2012). Computer-based psychological treatments for depression: A systematic review and meta-analysis. *Clinical Psychology Review*, *32*(4), 329–342.

Schröder, J., Berger, T., Westermann, S., Klein, J. P., & Moritz, S. (2016). Internet interventions for depression: New developments. *Dialogues in Clinical Neuroscience*, *18*(2), 203–212.

Schut, H., Margaret Stroebe. (1999). The dual process model of coping with bereavement: Rationale and description. *Death Studies*, *23*(3), 197–224.

Stroebe, M., & Schut, H. (2010). The dual process model of coping with bereavement: A decade on. *OMEGA-Journal of Death and Dying*, *61*(4), 273–289.

Tielman, M., Neerincx, M., Pagliari, C., Rizzo, A., & Brinkman, W. (2019). Considering patient safety in autonomous e-mental health systems–detecting risk situations and referring patients back to human care. *BMC Medical Informatics and Decision Making*, *19*, 1–16.

van Velsen, L., Cabrita, M., op den Akker, H., Brandl, L., Isaac, J., Suárez, M., Gouveia, A., Evans, N., de Sousa, R. D., Rodrigues, A. M., et al. (2020). Leaves (optimizing the mentaL health and resiliencE of older Adults that haVe lost thEir spouSe via blended, online therapy): Proposal for an online service development and evaluation. *JMIR Research Protocols*, *9*(9), e19344.

Wilson, D. M., Darko, E. M., Kusi-Appiah, E., Roh, S. J., Ramic, A., & Errasti-Ibarrondo, B. (2022). What exactly is "complicated" grief? A scoping research literature review to understand its risk factors and prevalence. *OMEGA-Journal of Death and Dying*, *86*(2), 471–487. https://doi.org/10.1177/0030222820977305

Worden, J. W. (2018). *Grief counseling and grief therapy: A handbook for the mental health practitioner* (Vol. 5). Springer Publishing Company, LLC.

World Health Organization. (2019). *International statistical classification of diseases and related health problems* (Vol. 11).

Zadeh, L. A., Klir, G. J., & Yuan, B. (1996). *Fuzzy sets, fuzzy logic, and fuzzy systems: Selected papers* (Vol. 6). World Scientific.

Zuelke, A. E., Luppa, M., Löbner, M., Pabst, A., Schlapke, C., Stein, J., & Riedel-Heller, S. G. (2021). Effectiveness and feasibility of internet-based interventions for grief after bereavement: Systematic review and meta-analysis. *JMIR Mental Health*, *8*(12), e29661.

&

Chapter 2

Alwin, D. F., & Krosnick, J. A. (1985). The measurement of values in surveys: A comparison of ratings and rankings. *Public Opinion Quarterly*, *49*(4), 535–552. https://doi.org/10.1086/268949

Baumeister, H., Reichler, L., Munzinger, M., & Lin, J. (2014). The impact of guidance on internet-based mental health interventions—a systematic review. *Internet Interventions*, *1*(4), 205–215. https://doi.org/10.1016/j.invent.2014.08.003

Berger, T., Boettcher, J., & Caspar, F. (2014). Internet-based guided self-help for several anxiety disorders: A randomized controlled trial comparing a tailored with a standardized disorder-specific approach. *Psychotherapy*, *51*(2), 207. https://doi.org/10.1037/a0032527

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brodbeck, J., Berger, T., Biesold, N., Rockstroh, F., & Znoj, H. J. (2019). Evaluation of a guided internet-based self-help intervention for older adults after spousal bereavement or separation/divorce: A randomised controlled trial. *Journal of Affective Disorders*, *252*, 440–449. https://doi.org/10.1016/j.jad.2019.04.008

Brodbeck, J., Berger, T., & Znoj, H. J. (2017). An internet-based self-help intervention for older adults after marital bereavement, separation or divorce: Study protocol for a randomized controlled trial. *Trials*, *18*(1), 21. https://doi.org/10.1186/s13063-016-1759-5

Burley, C., Anderson, D., Brownlee, A., Lafer, G., Luong, T., McGowan, M., Nguyen, J., Trotter, W., Wine, H., Baglione, A., et al. (2020). Increasing engagement in eHealth interventions using personalization and implementation intentions. *2020 Systems and Information Engineering Design Symposium (SIEDS)*, 1–5. https://doi.org/10.1109/SIEDS49339.2020.9106640

Carlbring, P., Andersson, G., Cuijpers, P., Riper, H., & Hedman-Lagerlöf, E. (2018). Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: An updated systematic review and meta-analysis. *Cognitive Behaviour Therapy*, *47*(1), 1–18. https://doi.org/10.1080/16506073.2017.1401115

Carlbring, P., Maurin, L., Törngren, C., Linna, E., Eriksson, T., Sparthan, E., Strååt, M., von Hage, C. M., Bergman-Nordgren, L., & Andersson, G. (2011). Individually-tailored, internet-based treatment for anxiety disorders: A randomized controlled trial. *Behaviour Research and Therapy*, *49*(1), 18–24.

Dahlin, M., Carlbring, P., Håkansson, A., & Andersson, G. (2020). Internet-based self-help using automatic messages and support on demand for generalized anxiety disorder: An open pilot study. *Digital Psychiatry*, *3*(1), 12–19.

Dajani, J. S., Sincoff, M. Z., & Talley, W. K. (1979). Stability and agreement criteria for the termination of Delphi studies. *Technological Forecasting and Social Change*, *13*(1), 83–90. https://doi.org/10.1016/0040-1625(79)90007-6

Eisma, M. C., Boelen, P. A., van den Bout, J., Stroebe, W., Schut, H. A., Lancee, J., & Stroebe, M. S. (2015). Internet-based exposure and behavioral activation for complicated grief and rumination: A randomized controlled trial. *Behavior Therapy*, *46*(6), 729–748. https://doi.org/10.1016/j.beth.2015.05.007

Eysenbach, G. (2005). The law of attrition. *Journal of Medical Internet Research*, *7*(1), e11. https://doi.org/10.2196/jmir.7.1.e11

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health*, *4*(2), e19. https://doi.org/10.2196/mental.7785

Gokalp, H., & Clarke, M. (2013). Monitoring activities of daily living of the elderly and the potential for its use in telecare and teleHearth: A review. *Telemedicine and eHealth*, *19*(12), 910–923. https://doi.org/10.1089/tmj.2013.0109

Green, B., Jones, M., Hughes, D., & Williams, A. (1999). Applying the Delphi technique in a study of GPs' information requirements. *Health & Social care in the Community*, *7*(3), 198–205.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, *32*(4), 1008–1015.

Johansson, R., Sjöberg, E., Sjögren, M., Johnsson, E., Carlbring, P., Andersson, T., Rousseau, A., & Andersson, G. (2012). Tailored vs. standardized internet-based cognitive behavior therapy for depression and comorbid symptoms: A randomized controlled trial. *PloS one*, *7*(5), e36905. https://doi.org/10.1371/journal.pone.0036905

Kip, H., Keizer, J., da Silva, M. C., Beerlage-de Jong, N., Köhle, N., Kelders, S. M., et al. (2022). Methods for human-centered eHealth development: Narrative scoping review. *Journal of Medical Internet Research*, *24*(1), e31858.

Krebs, P., Prochaska, J. O., & Rossi, J. S. (2010). A meta-analysis of computer-tailored interventions for health behavior change. *Preventive Medicine*, *51*(3-4), 214–221. https://doi.org/10.1016/j.ypmed.2010.06.004

Lustria, M. L. A., Noar, S. M., Cortese, J., Van Stee, S. K., Glueckauf, R. L., & Lee, J. (2013). A meta-analysis of web-delivered tailored health behavior change interventions. *Journal of Health Communication*, *18*(9), 1039–1069. https://doi.org/10.1080/10810730.2013.768727

Miaskiewicz, T., & Kozar, K. A. (2011). Personas and user-centered design: How can personas benefit product design processes? *Design Studies*, *32*(5), 417–430.

Molina, N., Viola, M., Rogers, M., Ouyang, D., Gang, J., Derry, H., & Prigerson, H. G. (2019). Suicidal ideation in bereavement: A systematic review. *Behavioral Sciences*, *9*(5), 53. https://doi.org/10.3390/bs9050053

Morrison, L. G. (2015). Theory-based strategies for enhancing the impact and usage of digital health behaviour change interventions: A review. *Digital Health*, *1*, 2055207615595335. https://doi.org/10.1177/2055207615595335

Noar, S. M., Grant Harrington, N., Van Stee, S. K., & Shemanski Aldrich, R. (2011). Tailored health communication to change lifestyle behaviors. *American Journal of Lifestyle Medicine*, *5*(2), 112–122. https://doi.org/10.1177/1559827610387255

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: An example, design considerations and applications. *Information & Management*, *42*(1), 15–29. https://doi.org/10.1016/j.im.2003.11.002

Oromendia, P., Orrego, J., Bonillo, A., & Molinuevo, B. (2016). Internet-based self-help treatment for panic disorder: A randomized controlled trial comparing mandatory versus optional complementary psychological support. *Cognitive Behaviour Therapy*, *45*(4), 270–286.

Pagliari, C. (2007). Design and evaluation in ehealth: Challenges and implications for an interdisciplinary field. *Journal of Medical Internet Research*, *9*(2), e614.

Ryan, K., Dockray, S., & Linehan, C. (2019). A systematic review of tailored eHealth interventions for weight loss. *Digital Health*, *5*, 2055207619826685.

Schmidt, R. C. (1997). Managing Delphi surveys using nonparametric statistical techniques. *Decision Sciences*, *28*(3), 763–774. https://doi.org/10.1111/j.1540-5915.1997.tb01330.x

Schröder, J., Berger, T., Westermann, S., Klein, J. P., & Moritz, S. (2016). Internet interventions for depression: New developments. *Dialogues in Clinical Neuroscience*, *18*(2), 203–212.

Schut, H., Margaret Stroebe. (1999). The dual process model of coping with bereavement: Rationale and description. *Death Studies*, *23*(3), 197–224.

Straat, A. C., Coenen, P., Smit, D. J., Hulsegge, G., Bouwsma, E. V., Huirne, J. A., van Geenen, R. C., Janssen, R., Boymans, T. A., Kerkhoffs, G. M., et al. (2020). Development of a personalized m/eHearth algorithm for the resumption of activities of daily life including work and sport after total and unicompartmental knee arthroplasty: A multidisciplinary Delphi study. *International Journal of Environmental Research and Public Health*, *17*(14), 4952. https://doi.org/10.3390/ijerph17144952

Stroebe, M., & Schut, H. (2010). The dual process model of coping with bereavement: A decade on. *OMEGA-Journal of Death and Dying*, *61*(4), 273–289.

van Dooren, M. M., Visch, V., Spijkerman, R., Goossens, R. H., & Hendriks, V. M. (2020). Mental health therapy protocols and eHearth design: Focus group study. *JMIR Formative Research*, *4*(5), e15568. https://doi.org/10.2196/15568

van Velsen, L., Cabrita, M., op den Akker, H., Brandl, L., Isaac, J., Suárez, M., Gouveia, A., Evans, N., de Sousa, R. D., Rodrigues, A. M., et al. (2020). Leaves (optimizing the mentaL health and resiliencE of older Adults that haVe lost thEir spouSe via blended, online therapy): Proposal for an online service development and evaluation. *JMIR Research Protocols*, *9*(9), e19344.

von der Gracht, H. A. (2012). Consensus measurement in Delphi studies: Review and implications for future quality assurance. *Technological Forecasting and Social Change*, *79*(8), 1525–1536. https://doi.org/10.1016/j.techfore.2012.04.013

&

Wagner, B., Knaevelsrud, C., & Maercker, A. (2006). Internet-based cognitive-behavioral therapy for complicated grief: A randomized controlled trial. *Death Studies*, *30*(5), 429–453. https://doi.org/10.1080/07481180600614385

West, A. (2011). Using the Delphi technique: Experience from the world of counselling and psychotherapy. *Counselling and Psychotherapy Research*, *11*(3), 237–242.

Wilson, D. M., Darko, E. M., Kusi-Appiah, E., Roh, S. J., Ramic, A., & Errasti-Ibarrondo, B. (2020). What exactly is "complicated" grief? a scoping research literature review to understand its risk factors and prevalence. *OMEGA-Journal of Death and Dying*. https://doi.org/10.1177/0030222820977305

Yap, M. B., Lawrence, K. A., Rapee, R. M., Cardamone-Breen, M. C., Green, J., & Jorm, A. F. (2017). Partners in parenting: A multi-level web-based approach to support parents in prevention and early intervention for adolescent depression and anxiety. *JMIR Mental Health*, *4*(4), e8492.

Chapter 3

Acampora, G., & Loia, V. (2011). On the temporal granularity in fuzzy cognitive maps. *IEEE Transactions on Fuzzy Systems*, *19*(6), 1040–1057. https://doi.org/10.1109/TFUZZ.2011.2159799

Amirkhani, A., Papageorgiou, E., Mohseni, A., & Mosavi, M. R. (2017). A review of fuzzy cognitive maps in medicine: Taxonomy, methods, and applications. *Computer Methods and Programs in Biomedicine*, *142*, 129–145.

Apostolopoulos, I. D., & Groumpos, P. P. (2023). Fuzzy cognitive maps: Their role in explainable artificial intelligence. *Applied Sciences*, *13*(6), 3412. https://doi.org/10.3390/app13063412

Axelrod, R. (2015). *Structure of decision: The cognitive maps of political elites*. Princeton University Press.

Baños, R. M., Herrero, R., & Vara, M. D. (2022). What is the current and future status of digital mental health interventions? *The Spanish Journal of Psychology*, *25*, e5.

Bennett, C. C., & Doub, T. W. (2016). Expert systems in mental health care: Ai applications in decision-making and consultation. In *Artificial intelligence in behavioral and mental health care* (pp. 27–51). Elsevier.

Billis, A., Papageorgiou, E., Frantzidis, C., Tsatali, M., Tsolaki, A., & Bamidis, P. D. (2014). A decision-support framework for promoting independent living and ageing well. *IEEE journal of biomedical and health informatics*, *19*(1), 199–209.

Bourgani, E., Stylios, C., Georgopoulos, V., & Manis, G. (2013). A study on fuzzy cognitive map structures for medical decision support systems. *8th conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-13)*, 784–791.

Brandl, L., van Velsen, L., Brodbeck, J., Jacinto, S., Hofs, D., & Heylen, D. (2023). Developing an emental health monitoring module for older mourners using fuzzy cognitive maps. *Digital Health*, *9*, 20552076231183549.

Bueno, S., & Salmeron, J. L. (2009). Benchmarking main activation functions in fuzzy cognitive maps. *Expert Systems with Applications*, *36*(3, Part 1), 5221–5229. https://doi.org/https://doi.org/10.1016/j.eswa.2008.06.072

Carvalho, J. P. (2013). On the semantics and the use of fuzzy cognitive maps and dynamic cognitive maps in social sciences. *Fuzzy Sets and Systems*, *214*, 6–19. https://doi.org/10.1016/j.fss.2011.12.009

Carvalho, J. P., & Tomè, J. A. (1999). Rule-based fuzzy cognitive maps and fuzzy cognitive maps - a comparative study. *18th International Conference of the North American Fuzzy Information Processing Society-NAFIPS (Cat. No. 99TH8397)*, 115–119. https://doi.org/10.1109/NAFIPS.1999.781665

Carvalho, J. P., & Tomè, J. A. (2000). Rule based fuzzy cognitive maps-qualitative systems dynamics. *PeachFuzz 2000. 19th International Conference of the North American Fuzzy Information Processing Society-NAFIPS (Cat. No. 00TH8500)*, 407–411.

Chattopadhyay, S. (2017). A neuro-fuzzy approach for the diagnosis of depression. *Applied computing and informatics*, *13*(1), 10–18.

De Franciscis, D. (2014). Jfcm: A java library for fuzzycognitive maps. In *Fuzzy cognitive maps for applied sciences and engineering: From fundamentals to extensions and learning algorithms* (pp. 199–220). Springer. https://doi.org/10.1007/978-3-642-39739-4_12

Farahani, H., Nápoles, G., & Azadfallah, P. (2021). Fuzzy cognitive maps for impact assessment in psychological research: Case study of psychological well-being. *3th International Conference on Modern Approach in Humanities and Social Sciences (ICMHS)*.

Felix, G., Nápoles, G., Falcon, R., Froelich, W., Vanhoof, K., & Bello, R. (2019). A review on methods and software for fuzzy cognitive maps. *Artificial intelligence review*, *52*, 1707–1737.

Georgopoulos, V. C., Malandraki, G. A., & Stylios, C. D. (2003). A fuzzy cognitive map approach to differential diagnosis of specific language impairment. *Artificial Intelligence in Medicine*, *29*(3), 261–278. https://doi.org/10.1016/S0933-3657(02)00076-3

Giabbanelli, P. J. (2014). Modelling the spatial and social dynamics of insurgency. *Security Informatics*, *3*, 1–15.

Giabbanelli, P. J., & Nápoles, G. (2024). *Fuzzy cognitive maps: Best practices and modern methods*. Springer Nature.

Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H.-C., & Jeste, D. V. (2019). Artificial intelligence for mental health and mental illnesses: An overview. *Current psychiatry reports*, *21*, 1–18.

Gray, S. A., Gray, S., Cox, L. J., & Henly-Shepard, S. (2013). Mental modeler: A fuzzy-logic cognitive mapping modeling tool for adaptive environmental management. *2013 46th Hawaii International Conference on System Sciences*, 965–973.

Higgins, O., Short, B. L., Chalup, S. K., & Wilson, R. L. (2023). Artificial intelligence (AI) and machine learning (ML) based decision support systems in mental health: An integrative review. *International Journal of Mental Health Nursing*. https://doi.org/https://doi.org/10.1111/inm.13114

Kosko, B. (1986). Fuzzy cognitive maps. *International journal of man-machine studies*, *24*(1), 65–75.

Kumar, A., Sangwan, S. R., Arora, A., & Menon, V. G. (2022). Depress-dcnf: A deep convolutional neuro-fuzzy model for detection of depression episodes using iomt. *Applied Soft Computing*, *122*, 108863.

Leichsenring, F., Steinert, C., & Ioannidis, J. P. (2019). Toward a paradigm shift in treatment and research of mental disorders. *Psychological medicine*, *49*(13), 2111–2117.

Luxton, D. D. (2016). An introduction to artificial intelligence in behavioral and mental health care. In *Artificial intelligence in behavioral and mental health care* (pp. 1–26). Elsevier.

Masri, R. Y., & Jani, H. M. (2012). Employing artificial intelligence techniques in mental health diagnostic expert system. *2012 International Conference on Computer & Information Science (ICCIS)*, *1*, 495–499. https://doi.org/10.1109/ICCISci.2012.6297296

Mkhitaryan, S., Giabbanelli, P., Wozniak, M. K., Nápoles, G., De Vries, N., & Crutzen, R. (2022). Fcmpy: A python module for constructing and analyzing fuzzy cognitive maps. *PeerJ Computer Science*, *8*, e1078.

Nair, A., Reckien, D., & van Maarseveen, M. F. (2020). Generalised fuzzy cognitive maps: Considering the time dynamics between a cause and an effect. *Applied Soft Computing*, *92*, 106309. https://doi.org/10.1016/j.asoc.2020.106309

Nandi, A. K. (2012). Ga-fuzzy approaches: Application to modeling of manufacturing process. *Statistical and computational techniques in manufacturing*, 145–185.

Nápoles, G., Grau, I., & Salgueiro, Y. (2024). A revised cognitive mapping methodology for modeling and simulation. *Knowledge-Based Systems*, *299*, 112089. https://doi.org/https://doi.org/10.1016/j.knosys.2024.112089

Nápoles, G., Espinosa, M. L., Grau, I., & Vanhoof, K. (2018). FCM expert: Software tool for scenario analysis and pattern classification based on fuzzy cognitive maps. *International Journal on Artificial Intelligence Tools*, *27*(07), 1860010.

Nápoles, G., Salmeron, J. L., Froelich, W., Falcon, R., Leon Espinosa, M., Vanhoenshoven, F., Bello, R., & Vanhoof, K. (2020). Fuzzy cognitive modeling: Theoretical and practical considerations. *Intelligent Decision Technologies 2019: Proceedings of the 11th KES International Conference on Intelligent Decision Technologies (KES-IDT 2019), Volume 1*, 77–87.

Olazabal, M., & Reckien, D. (2015). Fuzzy cognitive mapping: Applications to urban environmental decision-making. In *Handbook of research methods and applications in environmental studies* (pp. 148–176). Edward Elgar Publishing.

Özesmi, U., & Özesmi, S. L. (2004). Ecological models based on people's knowledge: A multi-step fuzzy cognitive mapping approach. *Ecological Modelling*, *176*(1-2), 43–64. https://doi.org/10.1016/j.ecolmodel.2003.10.027

Papageorgiou, E. (2011). A new methodology for decisions in medical informatics using fuzzy cognitive maps based on fuzzy rule-extraction techniques. *Applied Soft Computing*, *11*(1), 500–513.

Papageorgiou, E. (2013). *Fuzzy cognitive maps for applied sciences and engineering: From fundamentals to extensions and learning algorithms* (Vol. 54). Springer Science & Business Media.

Papageorgiou, E., Billis, A., Frantzidis, C., Konstantinidis, E., & Bamidis, P. D. (2013). A preliminary fuzzy cognitive map-based desicion support tool for geriatric depression assessment. *2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 1–8.

Papakostas, G., & Koulouriotis, D. (2010). Classifying patterns using fuzzy cognitive maps. *Fuzzy cognitive maps: Advances in theory, methodologies, tools and applications*, 291–306.

Rahimi, N. (2018). *Developing a mixed-methods method to model elderly health technology adoption with fuzzy cognitive map, and its application in adoption of remote health monitoring technologies by elderly women* [Doctoral dissertation, Portland State University].

Reckien, D. (2014). Weather extremes and street life in india—implications of fuzzy cognitive mapping as a new tool for semi-quantitative impact assessment and ranking of adaptation measures. *Global Environmental Change*, *26*, 1–13.

Sendak, M., Elish, M. C., Gao, M., Futoma, J., Ratliff, W., Nichols, M., Bedoya, A., Balu, S., & O'Brien, C. (2020). "The human body is a black box": Supporting clinical decision-making with deep learning. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 99–109.

Stylios, C. D., & Groumpos, P. P. (2004). Modeling complex systems using fuzzy cognitive maps. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *34*(1), 155–162.

Szwed, P. (2021). Classification and feature transformation with fuzzy cognitive maps. *Applied Soft Computing*, *105*, 107271.

Teachman, B. A., Silverman, A. L., & Werntz, A. (2022). Digital mental health services: Moving from promise to results. *Cognitive and Behavioral Practice*, *29*(1), 97–104.

Thieme, A., Hanratty, M., Lyons, M., Palacios, J., Marques, R. F., Morrison, C., & Doherty, G. (2023). Designing human-centered AI for mental health: Developing clinically relevant applications for online CBT treatment. *ACM Transactions on Computer-Human Interaction*, *30*(2), 1–50.

Tornero-Costa, R., Martinez-Millana, A., Azzopardi-Muscat, N., Lazeri, L., Traver, V., Novillo-Ortiz, D., et al. (2023). Methodological and quality flaws in the use of artificial intelligence in mental health research: Systematic review. *JMIR Mental Health*, *10*(1), e42045.

Vergini, E., & Groumpos, P. (2017). New concerns on fuzzy cognitive maps equation and sigmoid function. *2017 25th Mediterranean Conference on Control and Automation (MED)*, 1113–1118.

Zadeh, L. A., Klir, G. J., & Yuan, B. (1996). *Fuzzy sets, fuzzy logic, and fuzzy systems: Selected papers* (Vol. 6). World Scientific.

Zhong, H., Miao, C., Shen, Z., & Feng, Y. (2008). Temporal fuzzy cognitive maps. *2008 IEEE International Conference on Fuzzy Systems (IEEE World Congress on Computational Intelligence)*, 1831–1840. https://doi.org/10.1109/FUZZY.2008.4630619

Chapter 4

Aoun, S. M., Breen, L. J., Howting, D. A., Rumbold, B., McNamara, B., & Hegney, D. (2015). Who needs bereavement support? a population based survey of bereavement risk and support need. *PloS ONE.*, *10*(3).

Balog, K., Radlinski, F., & Arakelyan, S. (2019). Transparent, scrutable and explainable user models for personalized recommendation. *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval; 2019*, 265–274.

Baumeister, H., Reichler, L., Munzinger, M., & Lin, J. (2014). The impact of guidance on internet-based mental health interventions—a systematic review. *Internet Interventions*, *1*(4), 205–215. https://doi.org/10.1016/j.invent.2014.08.003

Beck, A. T., Kovacs, M., & Weissman, A. (1979). Assessment of suicidal intention: The scale for suicide ideation. *Journal of consulting and clinical psychology.*, *47*(2), 343.

Belnap, B., Schulberg, H., He, F., Mazumdar, S., Reynolds, I. C., & Rollman, B. (2015). Electronic protocol for suicide risk management in research participants. *Journal of Psychosomatic Research.*, *78*(4), 340–345.

Bourgani, E., Stylios, C., Georgopoulos, V., & Manis, G. (2013). A study on fuzzy cognitive map structures for medical decision support systems. *8th conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-13)*, 784–791.

Brandl, L., Cabrita, M., Brodbeck, J., Heylen, D., & van Velsen, L. (2022). Consulting the oracle: A delphi study for determining parameters for a mental health user profile and personalization strategy for an online service to aid grieving older adults. *Internet interventions*, *28*, 100534.

Brodbeck, J., Jacinto, S., Gouveia, A., Mendonça, N., Madörin, S., Brandl, L., et al. (2022). A web-based self-help intervention for coping with the loss of a partner: Protocol for randomized controlled trials in 3 countries. *JMIR Research Protocols.*, *11*(11).

Brodbeck, J., Berger, T., Biesold, N., Rockstroh, F., & Znoj, H. J. (2019). Evaluation of a guided internet-based self-help intervention for older adults after spousal bereavement or separation/divorce: A randomised controlled trial. *Journal of Affective Disorders*, *252*, 440–449. https://doi.org/10.1016/j.jad.2019.04.008

Dahlin, M., Carlbring, P., Håkansson, A., & Andersson, G. (2020). Internet-based self-help using automatic messages and support on demand for generalized anxiety disorder: An open pilot study. *Digital Psychiatry.*, *3*(1), 12–19.

Eisma, M. C., Boelen, P. A., van den Bout, J., Stroebe, H., Schut, A., Lancee, J., et al. (2015). Internet-based exposure and behavioral activation for complicated grief and rumination: A rondomized controlled trial. *Behavior Therapy.*, *46*(6), 729–748.

Franciscis, D. D. (2014). Jfcm: A java library for fuzzy cognitive maps. In *Fuzzy cognitive maps for applied sciences and engineering. berlin* (pp. 199–220). Springer.

Frias, M., Filiberto, Y., Nápoles, G., García-Socarrás, Y., Vanhoof, K., & Bello, R. (2017). Fuzzy cognitive maps reasoning with words based on triangular fuzzy numbers. *Advances in Soft Computing: 16th Mexican International Conference on Artificial Intelligence; 2017; City*, 197–207.

Giabbanelli, P. J. (2014). Modelling the spatial and social dynamics of insurgency. *Security Informatics*, *3*, 1–15.

Knight, K., Kenny, A., & Endacott, R. (2015). Assessing clinical urgency via telephone in rural australia. *Nursing and Health Sciences.*, *17*(2), 201–207.

Kosko, B. (1986). Fuzzy cognitive maps. *International journal of man-machine studies*, *24*(1), 65–75.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The phq-9: Validity of a brief depression severity measure. *Journal of general internal medicine.*, *16*(9), 606–613.

Lundorff, M., Holmgren, H., Zachariae, R., Farver-Vestergaard, I., & O'Connor, M. (2017). Prevalence of prolonged grief disorder in adult bereavement: A systematic review and meta-analysis. *Journal of Affective Disorders. Journal of Affective Disorders.*, *212*, 138–149.

McCall, W. V., Porter, B., Pate, A. R., Bolstad, C. J., Drapeau, C. W., Krystal, A. D., et al. (2021). Examining suicide assessment measures for research use: Using item response theory to optimize psychometric assessment for research on suicidal ideation in major depressive disorder. *Suicide and Life-Threatening Behavior.*, *51*(6), 1086–1094.

Mkhitaryan, S., Giabbanelli, P., Wozniak, M. K., Nápoles, G., De Vries, N., & Crutzen, R. (2022). Fcmpy: A python module for constructing and analyzing fuzzy cognitive maps. *PeerJ Computer Science*, *8*, e1078.

Molina, N., Viola, M., Rogers, M., Ouyang, D., Gang, J., Derry, H., & Prigerson, H. G. (2019). Suicidal ideation in bereavement: A systematic review. *Behavioral Sciences*, *9*(5), 53. https://doi.org/10.3390/bs9050053

Mural.co. (2022). *Mural* [Available at https://www.mural.co/about-us [cited 2022 September 13]].

Nandi, A. K. (2012). Ga-fuzzy approaches: Application to modeling of manufacturing process. In *Statistical and computational techniques in manufacturing. berlin* (pp. 145–185). Springer.

Nauck, D., & Kruse, R. (1999). Obtaining interpretable fuzzy classification rules from medical data. *Artificial Intelligence in Medicine.*, *16*(2), 149–169.

Olazabal, M., & Reckien, D. (2015). Fuzzy cognitive mapping: Applications to urban environmental decision-making. In *Handbook of research methods and applications in environmental studies* (pp. 148–176). Edward Elgar Publishing.

Oromendia, P., Orrego, J., Bonillo, A., & Molinuevo, B. (2016). Internet-based self-help treatment for panic disorder: A randomized controlled trial comparing mandatory versus optional complementary psychological support. *Cognitive Behavioral Therapy.*, *45*(4), 270–286.

Papageorgiou, E. (2011a). Fuzzy inference map approach to cope with uncertainty in modeling medical knowledge and making decisions. *Intelligent Decision Techniques.*, *5*(3), 219–235.

Papageorgiou, E. (2011b). A new methodology for decisions in medical informatics using fuzzy cognitive maps based on fuzzy rule-extraction techniques. *Applied Soft Computing*, *11*(1), 500–513.

Papageorgiou, E. (2013). *Fuzzy cognitive maps for applied sciences and engineering: From fundamentals to extensions and learning algorithms* (Vol. 54). Springer Science & Business Media.

Papageorgiou, E., Billis, A., Frantzidis, C., Konstantinidis, E., & Bamidis, P. D. (2013). A preliminary fuzzy cognitive map-based desicion support tool for geriatric depression assessment. *2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 1–8.

Reckien, D., Wildenberg, M., & Bachhofer, M. (2013). Subjective realities of climate change: How mental maps of impacts deliver socially sensible adaptation options. *Sustainability science.*, *8*, 159–172.

Reckien, D. (2014). Weather extremes and street life in india—implications of fuzzy cognitive mapping as a new tool for semi-quantitative impact assessment and ranking of adaptation measures. *Global Environmental Change*, *26*, 1–13.

Robinson, J., Hetrick, S., Cox, G., Bendall, S., Yung, A., & Pirkis, J. (2015). The safety and acceptability of delivering an online intervention to secondary students at risk of suicide: Findings from a pilot study. *Early intervention in psychiatry.*, *9*(6), 498–506.

Sands, N., Elsom, S., Gerdtz, M., Henderson, K., Keppich-Arnold, S., Droste, N., et al. (2013). Identifying the core competencies of mental health telephone triage. *Journal of Clinical Nursing.*, *22*(21-22), 3203–3216.

Sokero, T. P., Melartin, T. K., Rytsälä, H. J., Leskelä, U. S., & Isometsä, E. T. (2003). Suicidal ideation and attempts among psychiatric patients with major depressive disorder. *The Journal of clinical psychiatry.*, *64*(9), 1094–1100.

Spahni, S., Morselli, D., Perrig-Chiello, P., & Bennett, K. M. (2015). Patterns of psychological adaptation to spousal bereavement in old age. *Gerontology.*, *61*(5), 456–468.

Spence, J., Titov, N., Johnston, L., Jones, M. P., Dear, B. F., & Solley, K. (2014). Internet-based trauma-focused cognitive behavioural therapy for ptsd with and without exposure components: A randomised controlled trial. *Journal of Affective Disorders.*, *162*, 73–80.

Stroebe, M., & Schut, H. (1999). The dual process model of coping with bereavement: Rationale and description. *Death Studies.*, *23*(3), 197–224.

Stroebe, M., & Schut, H. (2010). The dual process model of coping with bereavement: A decade on. *OMEGA-Journal of Death and Dying*, *61*(4), 273–289.

Stylios, C. D., Georgopoulos, V. C., Malandraki, G. A., & Chouliara, S. (2008). Fuzzy cognitive map architectures for medical decision support systems. *Applied Soft Computing.*, *8*(3), 1243–1251.

Stylios, C. D., & Groumpos, P. P. (2004). Modeling complex systems using fuzzy cognitive maps. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *34*(1), 155–162.

Tielman, M., Neerincx, M., Pagliari, C., Rizzo, A., & Brinkman, W. (2019). Considering patient safety in autonomous e-mental health systems–detecting risk situations and referring patients back to human care. *BMC Medical Informatics and Decision Making*, *19*, 1–16.

Tielman, M., Neerincx, M. A., Bidarra, R., Kybartas, B., & Brinkman, W. (2017). A therapy system for post-traumatic stress disorder using a virtual agent and virtual storytelling to reconstruct traumatic memories. *Journal of Medical Systems.*, *41*, 1–10.

van Velsen, L., Cabrita, M., op den Akker, H., Brandl, L., Isaac, J., Suárez, M., Gouveia, A., Evans, N., de Sousa, R. D., Rodrigues, A. M., et al. (2020). Leaves (optimizing the mentaL health and resiliencE of older Adults that haVe lost thEir spouSe via blended, online therapy): Proposal for an online service development and evaluation. *JMIR Research Protocols*, *9*(9), e19344.

Wagner, B., Knaevelsrud, C., & Maercker, A. (2006). Internet-based cognitive-behavioral therapy for complicated grief: A randomized controlled trial. *Death Studies*, *30*(5), 429–453. https://doi.org/10.1080/07481180600614385

Wolters, M. K., Martínez-Miranda, J., Estevez, S., Hastie, H. F., & Matheson, C. (2013). Managing data in help4mood. *EAI Endorsed Transactions on Ambient Systems.*, *1*(2).

Zadeh, L. A., Klir, G. J., & Yuan, B. (1996). *Fuzzy sets, fuzzy logic, and fuzzy systems: Selected papers* (Vol. 6). World Scientific.

Chapter 5

Andrews, J. A., Brown, L. J., Hawley, M. S., & Astell, A. J. (2019). Older adults' perspectives on using digital technology to maintain good mental health: Interactive group study. *Journal of Medical Internet Research*, *21*(2), e11694.

Baumeister, H., Reichler, L., Munzinger, M., & Lin, J. (2014). The impact of guidance on internet-based mental health interventions—a systematic review. *Internet Interventions*, *1*(4), 205–215. https://doi.org/10.1016/j.invent.2014.08.003

Beck, A. T., Kovacs, M., & Weissman, A. (1979). Assessment of suicidal intention: The scale for suicide ideation. *Journal of consulting and clinical psychology.*, *47*(2), 343.

Blades, C. A., Stritzke, W. G., Page, A. C., & Brown, J. D. (2018). The benefits and risks of asking research participants about suicide: A meta-analysis of the impact of exposure to suicide-related content. *Clinical psychology review*, *64*, 1–12.

Brandl, L., van Velsen, L., Brodbeck, J., Jacinto, S., Hofs, D., & Heylen, D. (2023). Developing an emental health monitoring module for older mourners using fuzzy cognitive maps. *Digital Health*, *9*, 20552076231183549.

Brodbeck, J., Jacinto, S., Gouveia, A., Mendonça, N., S, M., Brandl, L., Schokking, L., Rodrigues, A., Gonçalves, J., Mooser, B., Marques, M., Isaac, J., Nogueira, V., Matos Pires, A., & van Velsen, L. (2022). A web-based self-help intervention for coping with the loss of a partner: Protocol for randomized controlled trials in 3 countries. *JMIR Research Protocols*, *11*(11).

Brodbeck, J., Berger, T., & Znoj, H. J. (2017). An internet-based self-help intervention for older adults after marital bereavement, separation or divorce: Study protocol for a randomized controlled trial. *Trials*, *18*(1), 21. https://doi.org/10.1186/s13063-016-1759-5

Bui, E. (2018). Grief: From normal to pathological reactions. *Clinical handbook of bereavement and grief reactions*, 85–101.

Carlbring, P., Andersson, G., Cuijpers, P., Riper, H., & Hedman-Lagerlöf, E. (2018). Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: An updated systematic review and meta-analysis. *Cognitive Behaviour Therapy*, *47*(1), 1–18. https://doi.org/10.1080/16506073.2017.1401115

Eysenbach, G. (2005). The law of attrition. *Journal of Medical Internet Research*, *7*(1), e11. https://doi.org/10.2196/jmir.7.1.e11

Farrer, L., Leach, L., Griffiths, K. M., Christensen, H., & Jorm, A. F. (2008). Age differences in mental health literacy. *BMC Public Health*, *8*(1), 1–8.

Hui Gan, G. Z., Hill, A.-M., Yeung, P., Keesing, S., & Netto, J. A. (2020). Pet ownership and its influence on mental health in older adults. *Aging & mental health*, *24*(10), 1605–1612.

Kessler, E.-M., Agines, S., & Bowen, C. E. (2015). Attitudes towards seeking mental health services among older adults: Personal and contextual correlates. *Aging & mental health*, *19*(2), 182–191.
Lyons, A., Alba, B., Heywood, W., Fileborn, B., Minichiello, V., Barrett, C., Hinchliff, S., Malta, S., & Dow, B. (2018). Experiences of ageism and the mental health of older adults. *Aging & mental health*, *22*(11), 1456–1464.
Mackenzie, C. S., Heath, P. J., Vogel, D. L., & Chekay, R. (2019). Age differences in public stigma, self-stigma, and attitudes toward seeking help: A moderated mediation model. *Journal of Clinical Psychology*, *75*(12), 2259–2272.
Meadows, G., Burgess, P., Bobevski, I., Fossey, E., Harvey, C., & Liaw, S.-T. (2002). Perceived need for mental health care: Influences of diagnosis, demography and disability. *Psychological Medicine*, *32*(2), 299–309. https://doi.org/10.1017/S0033291701004913
Moye, J., Karel, M. J., Stamm, K. E., Qualls, S. H., Segal, D. L., Tazeau, Y. N., & DiGilio, D. A. (2019). Workforce analysis of psychological practice with older adults: Growing crisis requires urgent action. *Training and Education in Professional Psychology*, *13*(1), 46.
Peek, S. T., Luijkx, K. G., Rijnaard, M. D., Nieboer, M. E., Van Der Voort, C. S., Aarts, S., Van Hoof, J., Vrijhoef, H. J., & Wouters, E. J. (2016). Older adults' reasons for using technology while aging in place. *Gerontology*, *62*(2), 226–237.
Richards, D., & Richardson, T. (2012). Computer-based psychological treatments for depression: A systematic review and meta-analysis. *Clinical Psychology Review*, *32*(4), 329–342.
Shear, M. K. (2015). Complicated grief. *New England Journal of Medicine*, *372*(2), 153–160.
Stylios, C. D., Georgopoulos, V. C., Malandraki, G. A., & Chouliara, S. (2008). Fuzzy cognitive map architectures for medical decision support systems. *Applied Soft Computing.*, *8*(3), 1243–1251.
Tielman, M., Neerincx, M., Pagliari, C., Rizzo, A., & Brinkman, W. (2019). Considering patient safety in autonomous e-mental health systems–detecting risk situations and referring patients back to human care. *BMC Medical Informatics and Decision Making*, *19*, 1–16.
van Zelst, C. M., Kasteleyn, M. J., van Noort, E. M., Rutten-van Molken, M. P., Braunstahl, G.-J., Chavannes, N. H., et al. (2021). The impact of the involvement of a healthcare professional on the usage of an ehealth platform: A retrospective observational copd study. *Respiratory research*, *22*(1), 1–8.
Wagner, B., Knaevelsrud, C., & Maercker, A. (2006). Internet-based cognitive-behavioral therapy for complicated grief: A randomized controlled trial. *Death Studies*, *30*(5), 429–453. https://doi.org/10.1080/07481180600614385
Zhang, T., Schoene, A. M., Ji, S., & Ananiadou, S. (2022). Natural language processing applied to mental illness detection: A narrative review. *NPJ digital medicine*, *5*(1), 1–13.

Chapter 6

Abdullah, S., & Choudhury, T. (2018). Sensing technologies for monitoring serious mental illnesses. *IEEE MultiMedia*, *25*(1), 61–75. https://doi.org/10.1109/MMUL.2018.011921236
Apostolopoulos, I. D., & Groumpos, P. P. (2023). Fuzzy cognitive maps: Their role in explainable artificial intelligence. *Applied Sciences*, *13*(6), 3412. https://doi.org/10.3390/app13063412
ATLAS.ti Scientific Software Development GmbH. (2023). Atlas.ti Windows (version 23.2.2.27458). https://atlasti.com/
Brandl, L., van Velsen, L., Brodbeck, J., Jacinto, S., Hofs, D., & Heylen, D. (2023). Developing an emental health monitoring module for older mourners using fuzzy cognitive maps. *Digital Health*, *9*, 20552076231183549.
Brodbeck, J., Jacinto, S., Gouveia, A., Mendonça, N., S, M., Brandl, L., Schokking, L., Rodrigues, A., Gonçalves, J., Mooser, B., Marques, M., Isaac, J., Nogueira, V., Matos Pires, A., & van Velsen, L. (2022). A web-based self-help intervention for coping with the loss of a partner: Protocol for randomized controlled trials in 3 countries. *JMIR Research Protocols*, *11*(11).

Cabitza, F., & Campagner, A. (2021). The need to separate the wheat from the chaff in medical informatics: Introducing a comprehensive checklist for the (self)-assessment of medical AI studies. *International Journal of Medical Informatics*, *153*, 104510.

De Jong Gierveld, J., & Van Tilburg, T. (2010). The de jong gierveld short scales for emotional and social loneliness: Tested on data from 7 countries in the un generations and gender surveys. *European journal of ageing*, *7*, 121–130.

de Jong-Gierveld, J. (1987). Developing and testing a model of loneliness. *Journal of personality and social psychology*, *53*(1), 119.

Eisinga, R., Grotenhuis, M. t., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International journal of public health*, *58*, 637–642. https://doi.org/https://doi.org/10.1007/s00038-012-0416-3

Futterman, A., Holland, J. M., Brown, P. J., Thompson, L. W., & Gallagher-Thompson, D. (2010). Factorial validity of the texas revised inventory of grief—present scale among bereaved older adults. *Psychological assessment*, *22*(3), 675.

Malgaroli, M., Hull, T. D., Zech, J. M., & Althoff, T. (2023). Natural language processing for mental health interventions: A systematic review and research framework. *Translational Psychiatry*, *13*(1), 309. https://doi.org/https://doi.org/10.1038/s41398-023-02592-2

Mancini, A. D., & Bonanno, G. A. (2011). Loss and grief: The role of individual differences. *Resilience and mental health: Challenges across the lifespan*, 189–199.

Martin, A., Rief, W., Klaiberg, A., & Braehler, E. (2006). Validity of the brief patient health questionnaire mood scale (PHQ-9) in the general population. *General Hospital Psychiatry*, *28*(1), 71–77.

Muñoz, R. F. (2022). Harnessing psychology and technology to contribute to making health care a universal human right. *Cognitive and Behavioral Practice*, *29*(1), 4–14.

Nutley, S. B., Burén, J., & Thorell, L. B. (2023). Covid-19 restrictions resulted in both positive and negative effects on digital media use, mental health, and lifestyle habits. *International Journal of Environmental Research and Public Health*, *20*(16), 6583.

Paula Branco, L., & Ribeiro, R. (2016). A survey of predictive modelling under imbalanced distributions. *ACM Computing Surveys*, *49*, 31–48. https://doi.org/https://doi.org/10.1145/2907070

R Core Team. (2021). R: A language and environment for statistical computing (version 4.3.1). *R Foundation for Statistical Computing*. https://www.R-project.org/

Raeder, T., Forman, G., & Chawla, N. V. (2012). Learning from imbalanced data: Evaluation matters. *Data Mining: Foundations and Intelligent Paradigms: Volume 1: Clustering, Association and Classification*, 315–331.

Sakal, C., Li, J., Xiang, Y.-T., & Li, X. (2022). Development and validation of the chinese geriatric depression risk calculator (CGD-risk): A screening tool to identify elderly chinese with depression. *Journal of Affective Disorders*, *319*, 428–436. https://doi.org/10.1016/j.jad.2022.09.034

Sendak, M., Elish, M. C., Gao, M., Futoma, J., Ratliff, W., Nichols, M., Bedoya, A., Balu, S., & O'Brien, C. (2020). "The human body is a black box": Supporting clinical decision-making with deep learning. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 99–109.

Sun, Y., Wong, A. K., & Kamel, M. S. (2009). Classification of imbalanced data: A review. *International journal of pattern recognition and artificial intelligence*, *23*(04), 687–719.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of Medical Education*, *2*, 53.

Teachman, B. A., Silverman, A. L., & Werntz, A. (2022). Digital mental health services: Moving from promise to results. *Cognitive and Behavioral Practice*, *29*(1), 97–104.

Thieme, A., Hanratty, M., Lyons, M., Palacios, J., Marques, R. F., Morrison, C., & Doherty, G. (2023). Designing human-centered AI for mental health: Developing clinically relevant applications for online CBT treatment. *ACM Transactions on Computer-Human Interaction*, *30*(2), 1–50.

Tornero-Costa, R., Martinez-Millana, A., Azzopardi-Muscat, N., Lazeri, L., Traver, V., Novillo-Ortiz, D., et al. (2023). Methodological and quality flaws in the use of artificial intelligence in mental health research: Systematic review. *JMIR Mental Health*, *10*(1), e42045.

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

Welten, M., de Kroon, M. L., Renders, C. M., Steyerberg, E. W., Raat, H., Twisk, J. W., & Heymans, M. W. (2018). Repeatedly measured predictors: A comparison of methods for prediction modeling. *Diagnostic and Prognostic Research*, *2*(1), 1–10.

Whiting, D., & Fazel, S. (2019). How accurate are suicide risk prediction models? Asking the right questions for clinical practice. *BMJ Mental Health*, *22*(3), 125–128. https://doi.org/https://doi.org/10.1136/ebmental-2019-300102

Zhang, B., You, J., Rolls, E. T., Wang, X., Kang, J., Li, Y., Zhang, R., Zhang, W., Wang, H., Xiang, S., et al. (2024). Identifying behaviour-related and physiological risk factors for suicide attempts in the uk biobank. *Nature human behaviour*, 1–14. https://doi.org/10.1038/s41562-024-01903-x

Chapter 7

Aoun, S. M., Breen, L. J., White, I., Rumbold, B., & Kellehear, A. (2018). What sources of bereavement support are perceived helpful by bereaved people and why? Empirical evidence for the compassionate communities approach. *Palliative Medicine*, *32*(8), 1378–1388. https://doi.org/https://doi.org/10.1177/0269216318774995

Bishop, S. J., & Gagne, C. (2018). Anxiety, depression, and decision making: a computational perspective. *Annual Review of Neuroscience*, *41*(1), 371–388. https://doi.org/https://doi.org/10.1146/annurev-neuro-080317-062007

Bower, P., Kontopantelis, E., Sutton, A., Kendrick, T., Richards, D. A., Gilbody, S., Knowles, S., Cuijpers, P., Andersson, G., Christensen, H., et al. (2013). Influence of initial severity of depression on effectiveness of low intensity interventions: Meta-analysis of individual patient data. *BMJ*, *346*. https://doi.org/10.1136/bmj.f540

Breen, L. J., Kawashima, D., Joy, K., Cadell, S., Roth, D., Chow, A., & Macdonald, M. E. (2022). Grief literacy: A call to action for compassionate communities. *Death Studies*, *46*(2), 425–433. https://doi.org/https://doi.org/10.1080/07481187.2020.1739780

Burley, C., Anderson, D., Brownlee, A., Lafer, G., Luong, T., McGowan, M., Nguyen, J., Trotter, W., Wine, H., Baglione, A., & Barnes, L. E. (2020). Increasing engagement in eHealth interventions using personalization and implementation intentions. *2020 Systems and Information Engineering Design Symposium (SIEDS)*, 1–5. https://doi.org/10.1109/SIEDS49339.2020.9106640

Chikersal, P., Belgrave, D., Doherty, G., Enrique, A., Palacios, J. E., Richards, D., & Thieme, A. (2020). Understanding client support strategies to improve clinical outcomes in an online mental health intervention. *Proceedings of the 2020 CHI conference on Human Factors in Computing Systems*, 1–16. https://doi.org/https://doi.org/10.1145/3313831.33763

Cornish, P. (2020). *Stepped care 2.0: A paradigm shift in mental health*. Springer. https://doi.org/10.1007/978-3-030-48055-4

Delgadillo, J., Ali, S., Fleck, K., Agnew, C., Southgate, A., Parkhouse, L., Cohen, Z. D., DeRubeis, R. J., & Barkham, M. (2022). Stratified care vs stepped care for depression: A cluster randomized clinical trial. *JAMA psychiatry*, *79*(2), 101–108. https://doi.org/10.1001/jamapsychiatry.2021.3539

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health*, *4*(2), e19. https://doi.org/10.2196/mental.7785

Hindmarch, T., Hotopf, M., & Owen, G. S. (2013). Depression and decision-making capacity for treatment or research: a systematic review. *BMC Medical Ethics*, *14*, 1–10. https://doi.org/https://doi.org/10.1186/1472-6939-14-54

Nápoles, G., Grau, I., & Salgueiro, Y. (2024). A revised cognitive mapping methodology for modeling and simulation. *Knowledge-Based Systems*, *299*, 112089. https://doi.org/https://doi.org/10.1016/j.knosys.2024.112089

Nobles, A. L., Glenn, J. J., Kowsari, K., Teachman, B. A., & Barnes, L. E. (2018). Identification of imminent suicide risk among young adults using text messages. *Proceedings of the 2018 CHI conference on human factors in computing systems*, 1–11. https://doi.org/https://doi.org/10.1145/3173574.3173987

O'Donohue, W., & Draper, C. (2010). *Stepped care and e-Health*. Springer. https://doi.org/10.1007/978-1-4419-6510-3

Thieme, A., Hanratty, M., Lyons, M., Palacios, J., Marques, R. F., Morrison, C., & Doherty, G. (2023). Designing human-centered AI for mental health: Developing clinically relevant applications for online CBT treatment. *ACM Transactions on Computer-Human Interaction*, *30*(2), 1–50.

Tielman, M., Neerincx, M., Pagliari, C., Rizzo, A., & Brinkman, W. (2019). Considering patient safety in autonomous e-mental health systems–detecting risk situations and referring patients back to human care. *BMC Medical Informatics and Decision Making*, *19*, 1–16.

Summary

The loss of a loved one is a common occurrence, especially in later life. Despite the universality of loss experiences, many people who have recently lost someone have difficulty making their wishes and needs explicit, for themselves and for friends and family that surround them. At the same time, comforting bereaved friends, family and loved ones is challenging, as there are no clear guidelines for providing solace. Grieving is a very personal matter and only recently, the authority on mental health disorders, the American Psychological Association (APA), has acknowledged that grief can exacerbate into serious mental health problems if maladaptive coping is left unattended. In line with recent trends in mental healthcare that move towards giving individuals more responsibility and control over the mental health support they receive, online self-guided psychological interventions can help prevent the exacerbation of grief into more serious mental health problems. In the context of such an online self-guided grief service, called LEAVES, this dissertation summarizes the research involved in developing a system to monitor the mental health of (older) mourners who have lost their spouse. The aim of such a monitoring system is to detect whether the individual's mental health deteriorates to the extent that professional intervention is advisable and consequently, to inform the mourner's decision-making about stepping up to more intensive bereavement support. This dissertation informed the development of such a system in five research steps that correspond to the chapters in this dissertation:

1. The selection of reliable indicators for people's mental health and changes thereof, in the context the LEAVES online self-guided grief service (chapters 2 and 4).

2. The definition of a measurement protocol including decisions about the frequency of measurements, assessment tools, and the processing of collected monitoring data for decision-making (chapter 4).

3. The selection and configuration of an automatic decision-making component that determines whether it is advisable to recommend help-seeking to mourners (chapters 3 and 4).

4. Exploring a strategy for communicating the results of the mental health monitoring to mourners and/or care professionals (chapter 5).

5. A thorough evaluation of the developed monitoring system with regard to user acceptance and experience, accuracy of algorithmic decision-making and clinical relevance (chapters 5 and 6).

After introducing the LEAVES online grief service in the introduction of this dissertation, we approach the first research step in chapter 2. To be more

specific, we consult a group of experts in grief and eHealth to arrive at a set of user characteristics that we can monitor to inform decision-making about a) whether the mourner should consider seeking more intensive bereavement support even before starting the online grief service, and b) to detect if the user's mental health deteriorates to the extent that professional intervention is warranted once they start using the service. In addition, we explore personalization strategies for adapting the content of the online grief service to the needs of the mourner, including adaptations to the service that can be applied based on preferences and characteristics of the mourner before they even start working on the content of the grief service, and adaptations that can be dynamically adjusted depending on the mourner's day-to-day, or week-to-week needs and preferences.

Chapter 3 introduces fuzzy cognitive mapping (FCM), a soft computing technique that can be used to build automatic decision models based on the input of experts, such as psychotherapists and grief professionals. The chapter explains the theoretical and technical foundations of the decision-making algorithm presented in chapter 4. It explores a new application area for FCMs while providing fellow researchers with theoretical and practical guidelines to use the method for researching decision-making systems in eMental health.

Chapter 4 continues the research towards the development of a monitoring system in the LEAVES online grief service. Based on the results from chapter 2, it presents the measurement protocol, assessment tools, and a FCM-based decision algorithm, alongside a first scenario-based evaluation of the developed monitoring system.

In chapter 5, the perspectives of grief professionals and older mourners about the developed monitoring system and the recommendations it gives about help-seeking are integrated into design principles for support on-demand targeted at older adults, a form of stepped care that relies on the initiative of the user of an online mental health service for transitioning to higher intensity support.

In chapter 6, we evaluate the LEAVES monitoring system by assessing a) the accuracy of the decisions made by the FCM model, b) how well its assessment tools capture clinical changes in grief over the course of a 10-week intervention, and c) the system's perceived usefulness by eCoaches who used it to inform the guidance they provided to mourners during the 10-week intervention. This chapter rounds off the research involved in the development of the monitoring system for the LEAVES online grief service.

Finally, in the discussion of this dissertation, we summarize the main findings of each chapter and situate our research in a broader context, including its contributions to the research field of stepped care models, the application of FCMs to a novel research domain and the opportunities and challenges of employing algorithmic decision-making in digital mental health services, such as the LEAVES online grief service.

Samenvatting

Het verlies van een dierbare is, vooral op latere leeftijd, een veelvoorkomende gebeurtenis. Ondanks dat vrijwel iedereen te maken krijgt met verlieservaringen, hebben veel rouwenden moeite met het in kaart brengen van hun emoties, behoeften en wensen met betrekking tot rouwverwerking. Hierdoor is het voor hen ook lastiger over het verlies met nabestaanden te spreken. Tegelijkertijd zijn er geen duidelijke richtlijnen voor het bieden van troost, waardoor er voor naasten geen houvast is. Rouwen is zeer persoonlijk, wat de één prettig vindt, kan een ander wellicht niet waarderen. Pas onlangs heeft de American Psychological Association (APA), de autoriteit op het gebied van psychische stoornissen, erkend dat rouw kan verergeren tot ernstige psychische problemen. Gecompliceerde rouw werd pas 2022 als psychische stoornis opgenomen in de vijfde editie van het Diagnostic and Statistical Manual of Psychological Disorders (DSM-5). Online psychologische interventies kunnen preventief voorkomen dat rouw verergert tot ernstigere psychische problemen door bijvoorbeeld ongezonde coping-strategieën te identificeren en de rouwende te ondersteunen deze te veranderen.

In het kader van een online rouwinterventie, beschrijft dit proefschrift een reeks onderzoeken die de basis vormen voor de ontwikkeling van een monitoringsysteem met als doel de overstap naar offline professionele hulp te vergemakkelijken. Dit doet het monitoringsysteem door te detecteren of de mentale gezondheid van de rouwende zodanig verslechtert dat professionele interventie wenselijk is en dit met de rouwende eindgebruiker te delen.

Het onderzoek in dit proefschrift is onderverdeeld in vijf onderzoekstappen die behandeld worden in hoofdstuk 2 tot en met 6:

1. De selectie van betrouwbare indicatoren voor de mentale gezondheid van rouwenden en veranderingen daarin (hoofdstukken 2 en 4).
2. De definitie van een meetprotocol, inclusief beslissingen over de frequentie van metingen, meetinstrumenten en de verwerking van de verzamelde monitoringsdata als input voor een beslissingsalgoritme (hoofdstuk 4).
3. De selectie en configuratie van een beslissingsalgoritme dat bepaalt of het raadzaam is de rouwende eindgebruiker aan te bevelen professionele hulp in te schakelen (hoofdstukken 3 en 4).
4. Het verkennen van een effectieve strategie voor het communiceren van de resultaten van de mentale gezondheidsmonitoring aan rouwenden en/of zorgprofessionals (hoofdstuk 5).
5. Een grondige evaluatie van het ontwikkelde monitoringsysteem met betrekking tot gebruikersacceptatie en de nauwkeurigheid waarmee het

beslissingsalgoritme bepaalt of professionele hulp verstandig is. Verder werd gekeken naar de klinische relevantie van de meetinstrumenten die als input dienen voor het algoritme. De klinische relevantie werd bepaald aan de hand van empirische data over hoe rouwsymptomen gedurende een interventie van tien weken veranderen en in hoeverre de ontwikkelde meetinstrumenten deze veranderingen kunnen meten (hoofdstukken 5 en 6).

Na de introductie van de LEAVES online rouwservice in de inleiding van dit proefschrift, raadplegen we in hoofdstuk 2 een groep experts op het gebied van rouw en eHealth om indicatoren die geschikt zijn de mentale gezondheid van rouwende eindgebruikers van LEAVES te monitoren en a) te bepalen of de rouwende zou moeten overwegen om intensievere rouwondersteuning te zoeken en b) om te detecteren of de mentale gezondheid van de gebruiker zodanig verslechtert dat professionele interventie noodzakelijk is. Daarnaast verkennen we personalisatiestrategieën voor het aanpassen van de inhoud van de online rouwservice aan de behoeften van de rouwende, waaronder aanpassingen die eenmalig aan de service worden toegepast op basis van de voorkeuren en kenmerken van de rouwende en aanpassingen die dynamisch kunnen worden aangepast afhankelijk van de dagelijkse of wekelijkse behoeften en voorkeuren van de rouwende.

Hoofdstuk 3 introduceert fuzzy cognitive mapping (FCM), een soft computing techniek die kan worden gebruikt om automatische beslissingsmodellen te maken op basis van de input van experts, zoals psychotherapeuten en rouwprofessionals. Het hoofdstuk legt de theoretische en technische basis uit van het in hoofdstuk 4 gepresenteerde beslissingsalgoritme. Door FCM toe te passen op eMental health, verkennen we een nieuw toepassingsgebied voor de methode en bieden we andere onderzoekers theoretische en praktische richtlijnen voor het gebruik van de methode bij het onderzoek naar beslissingssystemen in eMental health.

Na het theoretische uitstapje in hoofdstuk 3, zet hoofdstuk 4 het onderzoek voort naar de ontwikkeling van een monitoringsysteem voor de LEAVES online rouwservice. Op basis van de resultaten uit hoofdstuk 2 presenteren wij het meetprotocol, de beoordelingsinstrumenten en het FCM-gebaseerde LEAVES beslissingsalgoritme, samen met een eerste scenario-gebaseerde evaluatie van het ontwikkelde monitoringsysteem.

In hoofdstuk 5 wordt nader onderzocht hoe twee belangrijke stakeholders van online rouwservices, rouwprofessionals en oudere rouwenden, over het ontwikkelde monitoringsysteem denken. Specifiek waren wij geïnteresseerd hoe professionals en eindgebruikers kijken naar de aanbevelingen met betrekking tot het zoeken naar psychologische hulp die geïntegreerd zijn in de LEAVES rouwservice. Overstappen van online zelfhulp naar offline support door een professional is een vorm van stepped care. In het geval van de LEAVES service wordt daarnaast op het eigen initiatief van de rouwende eindgebruiker gerekend om over te gaan naar intensievere ondersteuning en de juiste configuratie van deze functionaliteit vraagt nauwe samenwerking met alle stakeholders, zoals rouwprofessionals en rouwenden.

In hoofdstuk 6 evalueren we het LEAVES-monitoringsysteem door ten eerste de nauwkeurigheid van de beslissingen gemaakt door het FCM-model te beoordelen. Ten tweede onderzoeken wij hoe goed onze ontwikkelde meetinstrumenten klinische veranderingen in rouw vastleggen gedurende een interventie van 10 weken. Ten slotte waren wij benieuwd naar de toegevoegde waarde van het monitoringsysteem voor het werk van eCoaches die er gedurende een interventie van 10 weken gebruik van gemaakt hebben. Dit hoofdstuk sluit het onderzoek naar de ontwikkeling van het monitoringsysteem voor de LEAVES online rouwservice af.

De discussie van dit proefschrift plaatst ons onderzoek in een bredere context, inclusief de bijdragen ervan aan het onderzoeksveld van stepped care, de toepassing van FCM's op een nieuw onderzoeksgebied en de mogelijkheden en uitdagingen van het gebruik van algoritmische beslissingsmodellen in eMental health, waarvoor de LEAVES online rouwservice een voorbeeld is.

Zusammenfassung

Viele Menschen werden mit dem Verlust eines geliebten Menschen konfrontiert, insbesondere im fortgeschrittenen Alter. Trotz der Unvermeidlichkeit von Verlusterfahrungen im Leben, fällt es Trauernden oft schwer, ihre Wünsche und Bedürfnisse bei der Trauerverarbeitung zu erkennen und in Worte zu fassen – sowohl zum Zweck der Selbstreflektion als auch im Umgang mit Freunden und Verwandten. Gleichzeitig wird das Unterstützungsnetzwerk der Trauernden vor eine Herausforderung gestellt, denn es gibt keine Richtlinien, wie man Trost "richtig" spendet, und wie man Trauernde "richtig" unterstützt. Trauer ist eine sehr persönliche Angelegenheit und in vielen Gesellschaften immer noch ein Tabuthema. Und obwohl so gut wie jeder Mensch früher oder später mit Verlust und Trauer konfrontiert wird, Trauer und Trauerverarbeitung demnach ein allgegenwärtiges Thema sind, hat die American Psychological Association (APA) erst vor einigen Jahren anerkannt, dass Trauer zu ernsthaften psychischen Problemen führen kann, wenn ungesunde Bewältigungsstrategien nicht erkannt werden und man ihnen nicht gezielt entgegenwirkt.

In diesem Kontext können online psychologische Interventionen (eMental health applications) dazu beitragen, dass sich aus ungesunder Trauerbewältigung keine ernsthaften psychischen Probleme entwickeln, insbesondere wenn der Trauernde bereits psychisch vorbelastet ist. Eine solche Intervention kann Trauernde bei der Trauerverarbeitung unterstützen, indem sie die Selbstreflektion fördert und über wissenschaftlich erwiesene Fakten zum Trauerprozess und wirksame Verarbeitungsstrategien informiert. Somit kann eine Trauerintervention die Funktion eines roten Fadens für den Trauernden erfüllen, oder es ermöglichen, Freunde und Verwandte effektiver bei der Trauerbewältigung miteinzubeziehen.

Die vorliegende Dissertation trägt die Forschungserkenntnisse aus fünf Studien zusammen, die über mehrere Jahre an der Entwicklung einer solchen Trauerintervention beigetragen haben. Im Speziellen beschäftigt sich die vorliegende Dissertation mit der Entwicklung eines Systems zur Überwachung des psychischen Wohlbefindens von Nutzern einer solchen Trauerintervention, im Fachbereich auch Monitoring genannt. Ziel eines solchen Monitorings ist es, festzustellen, ob sich das psychische Wohlbefinden des Trauernden so weit verschlechtert, dass es ratsam ist, dass ein Experte, wie z.B. eine*n Trauercoach*in die Trauerverarbeitung professionell begleitet. Das entwickelte Monitoring soll dem trauernden Nutzer bei der Entscheidungsfindung helfen und dem Trauernden verdeutlichen, ob es ratsam ist, in seiner oder ihrer individuellen Situation professionelle Unterstützung bei der Trauerverarbeitung einzuschalten.

In dieser Dissertation wird ein Konzept für die Entwicklung von Überwachungssystemen im Bereich eMental Health vorgestellt und in fünf Forschungsschritten dargelegt, wie ein solches System entwickelt wird und welche Überlegungen bei

&

jedem Schritt berücksichtigt werden müssen. Diese fünf Forschungsschritte entsprechen den fünf Hauptkapiteln dieser Dissertation:

1. Die Auswahl zuverlässiger Indikatoren für das psychische Wohlbefinden von Nutzern einer online Trauerintervention, die es ermöglichen sorgliche Veränderungen des Wohlbefindens während der Trauerverarbeitung zu erfassen (Kapitel 2 und 4).
2. Die Ausarbeitung eines Messprotokolls, das unter anderem festlegt, wie häufig die im ersten Schritt definierten Indikatoren gemessen werden, mit welchen Messinstrumenten und wie Messungen weiterverarbeitet werden sollen, sodass sie als Entscheidungsbasis für einen automatischen Algorithmus geeignet sind (Kapitel 4).
3. Die Auswahl und Konfiguration eines automatischen Algorithmus, der bestimmt, ob die Trauerintervention dem trauernden Nutzer die Inanspruchnahme von professioneller Hilfe bei der Trauerbewältigung empfiehlt (Kapitel 3 und 4).
4. Die Definition von Designrichtlinien für psychologische Onlineinterventionen, die Monitoring nutzen, um ihre Nutzer, speziell Erwachsene im fortgeschrittenen Alter, darauf hinzuweisen, wenn deren psychische Gesundheit sich so weit verschlechtert, dass es ratsam ist, professionelle Unterstützung zu suchen (Kapitel 5).
5. Eine umfassende Auswertung des entwickelten Monitorings in Bezug auf die Akzeptanz durch trauernde Nutzer, der Korrektheit und die klinische Relevanz der durch den Entscheidungsalgorithmus gegebenen Empfehlungen (Kapitel 5 und 6).

Nach kurzer Einleitug der Thematik der vorliegenden Dissertation im ersten Kapitel, wurden im zweiten Kapitel (messbare) Eigenschaften von Trauernden definiert, die als Entscheidungsbasis für das Monitoring geeignet sind. Unabängig vom Nutzen der Trauerintervention kann professionelle Hilfe bei der Trauerverarbeitung ratsam sein, wenn der Verstorbene einen gewaltsamen Tod erlitten hat, wie z.B. einen Unfall oder Tod durch Suizid. Des Weiteren sollten diese Eigenschaften es dem Monitoring ermöglichen, den psychischen Zustand des Nutzers zu überwachen und zu erkennen, ob sich dieser während der Nutzung der Intervention ernsthaft verschlechtert. Basierenden auf einer Delphi-Studie, an der 16 Experten im Bereich der Trauerverarbeitung und eHealth teilgenommen haben, identifizierten wir fünf für diesen Zweck geeignete Indikatoren. Dazu gehören 1) selbstzerstörerisches Verhalten, einschließlich Suizidalität, 2) eine vom Nutzer gewünschte Intensivierung der Betreuung, 3) Unerreichbarkeit des Nutzers, 4) (Verschlechterung) der Trauersymptome und 5) Gefühle der Hoffnungslosigkeit und/oder Aussichtlosigkeit.

Kapitel 3 ist Fuzzy Cognitive Mapping (FCM) gewidmet, eine Methode, die in dieser Dissertation verwendet wurde, um automatische Entscheidungsmodelle zu erstellen und zu konfigurieren, die basieren auf den Erfahrungen von Experten. In dieser Dissertation wurde FCM erstmals angewendet zur Erstellung eines Monitoring Modells für psychologische Onlineinterventionen und identifizierten wir Vorteile der Methode gegenüber anderen Entscheidungsmodellen, wie ihrer Unabhängigkeit von klinischen Trainingsdaten für eine Erstkonfiguration des Algorithmus. Gleichzeitig haben wir Herausforderungen bei der Anwendung von FCM im Bereich eMental health aufgedeckt. Dazu gehört beispielsweise das Fehlen einer einfachen Möglichkeit, historische Nutzerdaten in das Entscheidungsmodell einzubeziehen.

Auf Basis der Erkenntnisse aus den ersten zwei Kapiteln, stellten wir im vierten Kapitel das Monitoring der LEAVES Trauerintervention vor. Es besteht aus zwei Fragebögen, die bei Nutzung der Intervention jede zweite Woche die Indikatoren, die in Kapitel 1 ermittelt wurden, beim Nutzer erfragen. Die Nutzerantworten werden dann automatisch durch ein FCM Modell und einer regelbasierten Erweiterung des FCM Entscheidungsmoduls ausgewertet und das System entscheidet, ob es ratsam ist, dem Nutzer zu empfehlen, professionelle Hilfe bei der Trauerbewältigung zu suchen. Das Kapitel schließt mit einer ersten Szenario basierten Auswertung des LEAVES Monitoring Moduls ab. Wir konnten feststellen, dass der Algorithmus gut zwischen Trauernden, die eindeutig keine intensivere Betreuung benötigen, und Trauernden, die eindeutig professionelle Unterstützung in Betracht ziehen sollten, unterscheidet. Die Genauigkeit des Modells in Fällen, in denen die Entscheidung weniger eindeutig ist, musste jedoch noch gründlicher und idealerweise basierend auf empirischen Daten bewertet werden.

In Kapitel 5 tragen wir die Erfahrungen von Trauernden im fortgeschrittenen Alter zusammen, die die LEAVES Trauerintervention zehn Wochen lang benutzt haben, und kombinieren deren Erfahrungen mit der professionellen Perspektive von Trauerfachleuten, wie z.B. Trauercoachs und psychologische Psychotherapeuten und -therapeutinnen. Wir befragten Trauerfachleute bezüglich des Monitorings und der Ratgeberfunktion, die das Monitoring erfüllt. Darauf basierend haben wir Designrichtlinien für Monitoringsysteme erstellt, dessen Ziel es ist, Nutzern zu empfehlen, professionelle Unterstützung zu suchen, wenn die Onlineintervention unzureichend beiträgt an einer Verbesserung der psychischen Situation des Nutzers. Eine dieser Designempfehlungen ist beispielsweise die Konzipierung des Monitorings als integraler Bestandteil der Onlineintervention. Das heißt, dass Ideen des Monitorings wie etwa der Bewusstseinswandel dahingehend, dass die Inanspruchnahme professioneller Unterstützung normal und keineswegs ein Zeichen von Versagen ist, in den Inhalten der gesamten Trauerintervention integriert und nicht nur innerhalb eines einzelnen Moduls behandelt werden sollten. Dies erhöht die Chance, dass der trauernde Nutzer tatsächlich Unterstützung sucht, sollte diese notwendig werden.

Im sechsten Kapitel werten wir das LEAVES Monitoring empirisch aus, indem wir die Ergebnisse von drei Bewertungsmethoden miteinander kombinieren und zusammen interpretieren. 1. Die Korrektheit der durch das Monitoring getroffenen Entscheidungen wurde mit dem F-Maß bewertet, das Genauigkeit (precision) und Trefferquote (recall) mittels des gewichteten harmonischen Mittels kombiniert. 2. Wir prüfen an der Hand einer linearen Regressionsanalyse, wie gut die Messinstrumente, die die Grundlage für das LEAVES Monitoring darstellen, die klinischen Veränderungen der Trauersymptome über den Verlauf einer 10-wöchigen Nutzung der Intervention erfassen, und 3. wir bewerten an der Hand der Erfahrungen von eCoachs, die spezialisiert waren, die Nutzer der LEAVES Intervention zu unterstützen, inwiefern das Monitoring ihre Arbeit komplementiert.

Insgesamt leistet die vorliegende Dissertation einen Beitrag zur laufenden Forschung im Bereich Monitoringsysteme in eMental health für die Prävention von psychischen Problemen, gerade bei einschneidenden Ereignissen im Leben, wie den Verlust eines geliebten Menschen. Darüber hinaus wurde in dieser Dissertation erstmals FCM für die Konzipierung und Entwicklung eines Monitoringsystems für eine eMental health Anwendung verwendet und identifizierten wir Vorteile und Herausforderungen, die in zukünftigen Forschungsansätzen, die FCM im Bereich eMental health nutzen, berücksichtigt werden sollten.

Acknowledgements | Dankwoord

Hell yes, finally!! After almost an entire year of "being almost there", I can finally and with good conscience say that I have, in fact, reached the finish line! Little did I know when I embarked on this journey almost five years ago in March 2020. Back then, I was still finishing the last bits of my master's thesis at Nedap Healthcare, where I recently ended up as part of the Data Science team (again) - who would have thought that? ;-)

Although this thesis is published under my name, there is a long list of people whose contribution to this PhD I want to acknowledge and for whom I want to express my gratitude for supporting, advising, and accompanying me throughout my PhD journey.

First, thank you Dirk and Stephanie for taking up the task of (co-)promoting my PhD. Dirk; dankjewel dat je ten allen tijde het grote plaatje niet uit het oog verloor, daarvoor dat je mij aanmoedigde mijn PhD hier en daar met opportunisme aan te pakken in plaats van te proberen (en te falen) iedere stap te anticiperen en vooruit te plannen. Bedankt voor de filosofische discussies over de subtiele culturele verschillen tussen Nederlanders, Duitsers en Belgen, en vooral voor je oprechte waardering voor de verschillende manieren waarop verschillende mensen een PhD en in het algemeen het leven aanpakken.

Stephanie, toen je de leiding van het LEAVES-project in 2022 overnam kreeg je de begeleiding van mijn PhD er als bonus bij. Dankjewel dat je mij steeds hielp het verschil tussen hoofd- en bijzaken te zien, dat je oprecht geïnteresseerd was mij beter te leren kennen en te begrijpen, beginnend bij onze gezamenlijke trip naar de AAL week in Gdańsk, en dankjewel voor al onze open en filosofische discussies, vaak over de menselijke psyche. Mede dankzij jou heb ik de finish gehaald! En aan jullie beiden: bedankt voor jullie vertrouwen dat alles goed zal komen voor jullie aanmoedigingen toen ik besloot een pauze in te lassen van mijn PhD om terug te keren naar Zuid-Korea om nog eens de Zuid-Koreaanse cultuur in te duiken en te reizen.

Miriam Cabrita and Lex van Velsen, thank you for running the first stages of the co-promoter relay run and for guiding my initial steps on my PhD journey. I also would like to thank the members of the graduation committee for reading and critically evaluating my thesis and for being present at my defence.

A shout-out to my colleagues from the LEAVES consortium; thank you for supporting the design and execution of my study protocols. Jeannette and Sofia, thank you for your sincere interest in my research, for contributing your clinical expertise about grief, for supporting the setup of my studies, and for thoroughly co-authoring several of my research articles. Collaborating with you was a joy!

Cindy; ik hoefte er niet lang over na te denken wie ik zou vragen mijn paranimf te

zijn: na de vele RRD borrels waar wij met z'n tweeën weer in de stad eindigden, een aantal sportsessies bij Denisa en zelfs 1 keer samen bij knappe Koen en nadat we samen in de bieb de laatste stukken tekst voor onze boekjes getypt hadden wist ik het al; en natuurlijk omdat ik je nuchtere blik op dingen waar ik me veels te druk om kan maken enorm kan waarderen. Thanks dat je mij als paranimf bij de laatste loodjes ondersteunt en ik kijk ernaar uit dat we gaan proosten op het behalen van onze PhDs tijdens onze wijnreis!

Luc; dat wij een goed team zijn hebben we meer dan 1 keer bewezen bij de legendarische Luc & Lena pubquizzen. Als iemand die liever 2 of 3 keer nadenkt voordat ze ergens voor gaat heb ik enorm veel bewondering voor je enthousiasme voor alles wat je doet en hoe je alle ballen in de lucht weet te houden. Ik vind jou een ontzettend mooi mens! Dankjewel voor alle gezelligheid en sparsessies tijdens mijn PhD en dat je mij tijdens mijn eindsprint naar de finish als paranimf ondersteunt!

Tessa; je plek in het rijtje van paranimfen heb je meer dan verdiend! Naast alle leuke dingen die we samen beleefd hebben, had ik zonder alle sparsessies met jou en je mentale support bij het herschrijven van mijn discussie, het reviewen van dat ene manuscript dat 3 weken voor mijn boek-deadline een reject en major revisions kreeg en het doneren van je LaTeX template voor mijn boekje, afgelopen jaar meerdere keer de moed verloren. Dankjewel dat ik altijd op je luisterend oor en je wijze en to-the-point adviezen mocht vertrouwen!

The first 3 years of my PhD I spent mostly at Roessingh Research and Development (RRD) where I was welcomed warmly in the midst of the onset of a pandemic. To my RRD colleagues, especially my fellow junior researchers Bouke, Eline, Kira, Marian, and Vinish with whom I strode the longest in parallel on the PhD journey; thank you for all the coffee breaks, lunches, regularly ordered from Broodbode, the amazing wine tastings due to which I can now proudly call myself a "wine snob" (thanks Erik!), the boardgame nights and drinks. Dennis, thank you for going down the fuzzy cognitive maps rabbit hole together with me and for all the support getting the monitoring to work in the LEAVES application. Leendert, thank you for your unwavering interest in my not quite RRD-typical research and thereby, for making my research island feel a little more connected to the mainland.

But of course, the perk of being an external PhD is that there is a second research community that you get to explore and become a part of. Thank you to my lovely colleagues at the Human Media Interaction group (HMI) at the University of Twente for welcoming me warmly. In particular, thank you for all the foosball matches, (board)game nights, movie nights and outings (to Christmas markets and boardgame conventions amongst other destinations), for reviewing my manuscripts, for cozy Christmas lunches, and everything in between. My apologies to those whose ears may have suffered once or twice (or many more times) from the enthusiasm that tends to overwhelm me during a good foosball match (or any game for that matter). Hideki and Thomas, thank

&

you for a great first wild camping experience in Scotland and rest assured, I will always think of you when it comes down to it. Thank you for a great brainstorming session about the cover of my thesis, Ella! Charlotte, dankjewel voor je interesse in alles wat er in het leven van ons HMIers speelt en dat je de administratie van HMI draaiende houdt!

Verder wil ik mijn collega's bij Nedap bedanken en natuurlijk in het bijzonder team Product Analytics voor jullie begrip voor mijn PhD-chaos dit afgelopen jaar, jullie geduld en dat jullie elke mijlpaal op weg naar mijn promotie met mij gevierd hebben: Anne, Cas, Diederik, Kasper, Marlous, Linda, Robin en Steven. Ik had geen fijner team en geen fijnere werkgever dan Nedap kunnen wensen voor deze transitieperiode! En vanaf nu zetten we samen koers naar nieuwe uitdagingen (werkgerelateerd of sportief ;-))!

Kevin, Steven; dankjewel voor vele leuke uitgebreide kooksessies and serie-avonden in de afgelopen jaren! En Kevin, dankje dat je de samenstelling van mijn desktop uit je hoofd weet zodat ik altijd bij jou kan checken of een nieuw onderdeel wel/niet past als ik hem een upgrade wil geven.

Thijs; dankjewel dat je de eerste jaren van mijn PhD tijdens de pandemie aan mij zijde stond. Ik denk nog steeds met veel plezier aan de pubquizzen die we samen bezochten, onze roadtrip langs de westkust van de Verenigde Staten en de vele goede herinneringen samen.

Kelly, Mark, Bente; van dag 1 tot de finish zijn jullie er voor mij geweest, en natuurlijk waren jullie er al lang voordat ik aan het avontuur PhD begon. Vanuit de andere kant van de wereld bellen voor een peptalk voor een sollicitatie of gewoon om even bij te kletsen? De beste klussers bij een verhuizing? Gezellig een wijntje of biertje (of meer) samen doen of gewoon lekker chillen bij het Rutbeek? Allemaal vanzelfsprekend! Ik waardeer jullie vriendschap en het gevoel van geborgenheid dat jullie mij geven enorm! En uiteraard word ik nog steeds blij van het strak geschilderde plafonnetje in mijn woonkamer! Ik houd enorm van jullie en hoop dat we nog vele jaren samen door het leven gaan!

Mama, Papa; für euch war es immer am Wichtigsten, dass ich etwas mache, was mir Freude bringt, selbst wenn ich dafür nach einem Jahr Studium in Düsseldorf doch in die Niederlande umziehen wollte. Danke dafür, dass ihr meinen Werdegang in allen Phasen unterstützt und mit meinen Erfolgen mitgefiebert habt. Ich weiß, dass eure Tür für mich immer offen steht. Tim, Bruderherz, ein anderes Land und eine vollkommen unterschiedliche Herangehensweise an den Doktor und trotzdem haben wir Ähnliches erlebt während unserer Promotionszeit, vor allem der Tunnel am Ende bleibt derselbe, egal, wo man seinen Doktor macht. Danke für dein aufrichtiges Interesse in meine Forschung, in meine Reiseerlebnisse und das eine oder andere Korrekturlesen meiner deutschen Texte - hab dich lieb! Norbert, Christiane, Opa (und Oma) und Becky; auch für eure Unterstützung meines Werdeganges und fürs Mitfiebern bin ich sehr dankbar!

Ties; gelukkig wist je al een beetje waar je aan begon toen je op date ging met een PhD die aan het einde van haar PhD-traject staat. Tussen een verhuizing die net anders liep dan verwacht en wat tegenslagen bij het afronden van mijn these was jij de constante die mij meer dan 1 keer de nodige rust en energie gaf om er weer tegen aan te kunnen. Je bent geweldig geweest dit afgelopen jaar! Ik kijk er heel erg naar uit wat de komende jaren ons gaan brengen!

About the author

Lena Brandl was born October 4th 1993 in Essen, Germany. She went to Maria-Wächtler-Gymnasium where she completed her A-levels in 2012. Afterwards, she started her studies in German and Communication and Media Sciences at the Heinreich-Heine-Universität in Düsseldorf, Germany before relocating to Enschede, the Netherlands in 2013. At the University of Twente, she changed the field of her studies to Psychology. She completed her Bachelor of Science in Psychology (cum laude) in 2016. Her profound interest in interactions between humans and technology motivated her to continue her studies in the master's program Interaction Technology at the University of Twente, which she completed in 2020. At Roessingh Research and Development in Enschede, The Netherlands, Lena saw a chance to combine her multi-disciplinary background in Psychology and Computer Science in the LEAVES research project. Between 2020 and 2024, she researched and developed a module in an online self-help service for older mourners who lost their spouse to monitor their mental health and to recommend seeking offline (professional) support if necessary. For her research efforts, combined in this thesis, she obtained the degree Doctor of Philosophy (PhD).

Lena currently lives in Enschede, the Netherlands and continues to research and critically reflect on human interactions with technology.

Publications

International journal papers

Brandl, L., Jansen-Kosterink, S., Brodbeck, J., Jacinto, S., Mooser, B., & Heylen, D. (2024). Moving Towards Meaningful Evaluations of Monitoring in eMental Health based on the Case of a Web-Based Grief Service for Older Mourners: Mixed Methods Study. *JMIR Formative Research*;8:e63262. doi:10.2196/63262. (**Included as chapter 6**).

van Velsen, L., Schokking, L., Siderakis, E., Knospe, G. M., **Brandl, L.**, Mooser, B., Madörin, S., Jacinto, S., Gouveia, A., & Brodbeck, J. (2023). Development of an online service for coping with spousal loss by means of human-centered and stakeholder-inclusive design: the case of LEAVES. *Death Studies*;48(2), 187–196. doi:10.1080/07481187.2023.2203680.

Brandl, L., van Velsen, L., Brodbeck, J., Jacinto, S., Hofs, D., & Heylen, D. (2023). Developing an eMental health monitoring module for older mourners using Fuzzy Cognitive Maps. *Digital Health*;9. doi:10.1177/20552076231183549. (**Included as chapter 4**)

Brodbeck, J., Jacinto, S., Gouveia, A., Mendonça, Madörin, S., **Brandl, L.**, Schokking, L., Rodrigues, A. M., Gonçalves, J., Mooser, B., Marques, M. M., Isaac, J., Nogueira, V., Pires, A. M., van Velsen, L. (2022). A Web-Based Self-help Intervention for Coping With the Loss of a Partner: Protocol for Randomized Controlled Trials in 3 Countries. *JMIR Research Protocols*;11(11):e37827. doi: 10.2196/37827.

Brandl, L., Cabrita, M., Brodbeck, J., Heylen, D., & van Velsen, L. (2022). Consulting the Oracle: A Delphi study for determining parameters for a mental health user profile and a personalization strategy for an online service to aid grieving older adults. *Internet Interventions*;28:100534. doi:10.1016/j.invent.2022.100534. (**Included as chapter 2**)

van Velsen, L., Cabrita, M., op den Akker, H., **Brandl, L.**, Isaac, J., Suárez, M., Gouveia, A., Evans, N., de Sousa, R. D., Rodrigues, A. M., Canhāo, H., DELA Natura- En Levensverzekeringen NV, Blok, M., Alcobia, C., Brodbeck, J. (2020). LEAVES (optimizing the mentaL health and resiliencE of older Adults that haVe lost thEir spouSe via blended, online therapy): Proposal for an Online Service Development and Evaluation. *JMIR Research Protocols*;9(9):e19344. doi: 10.2196/19344.

Submitted for publication

As of December 2024

Brandl, L., Jansen-Kosterink, S., Hofs, D., & Heylen, D. Fuzzy cognitive maps for

algorithmic decision making in eMental health. (**Included as chapter 3**).

Brandl, L., Jansen-Kosterink, S., Schokking, L., Siderakis, E., & Heylen, D. Design implications for transitioning from self-help to offline support in eMental health services for older adults. (**Included as chapter 5**).

Conference contributions

Brandl, L., Jansen-Kosterink, S., Brodbeck, J., Jacinto, S., Hofs, D., & Heylen, D. (2023). Decision-making algorithms using Fuzzy Cognitive Maps for eMental Health. Poster presentation at the *TechMed Research Day 2023*. Enschede, the Netherlands.

Brandl, L., Jansen-Kosterink, S., Schokking, L., Siderakis, E., & Heylen, D. (2022). Older adults' usage and appreciation of a monitoring system to deliver support on-demand in an online grief intervention. Poster presentation at the *European Week of Active and Healthy Ageing 2022*. Gdańsk, Poland.

Brandl, L., van Velsen, L., Brodbeck, J., Jacinto, S., & Heylen, D. (2022). Developing an eMental Health monitoring module for older mourners using Fuzzy Cognitive Maps. Abstract and oral presentation at the *Supporting Health by Technology Conference (HealthBy Tech 2022)*. Groningen, the Netherlands.

Brandl, L., Cabrita, M., Brodbeck, J., Heylen, D., & van Velsen, L. (2021). A Delphi study for delivering blended care and personalizing eHealth to support grieving older adults. Abstract and oral presentation at the *Supporting Health by Technology Conference (HealthBy Tech 2021)*. Online, the Netherlands.

Brandl, L., Cabrita, M., Brodbeck, J., Heylen, D., & van Velsen, L. (2021). A Delphi study for escalating self-help to more intensive support and for personalizing an online intervention to support grieving older adults. Oral panel presentation (*Panel: How to Personalise an Internet Intervention?*) at the *52nd International Meeting of the Society for Psychotherapy Research*. Online, Germany.

Brandl, L. & van Velsen, L. (2021). LEAVES: A virtual agent for processing grief in later life. Oral presentation at the *Workshop on Personalized Intelligent Conversational Agents*. Online, the Netherlands.

Progress range

The following publications have been published in the Progress range by Roessingh Research and Development, Enschede, the Netherlands. Copies can be ordered, when available, via info@rrd.nl.

1. Pot, J.W.G.A, Boer H., van Harten, W.H., Hermens, H.J., Seydel, E.R. Comprehensive Need-Assessment. Ontwikkeling van een meetinstrument voor zorgbehoeften en kwaliteitsbeoordeling door patiënten. September 1994, ISBN 90-25452-01-22
2. Van Leerdam, N.G.A, Hermens, H.J. Revalidatietechnologie in Euregio. July 1995, ISBN 90-75452-02-0
3. Duda, L., van Noort, L.O., Röseler, S., Greitemann, B.O.L, van Harten, W.H., Klazinga, N.S. Rehabilitation in Germany and the Netherlands, A comparison of two rehabilitation systems. August 1995, ISBN 90-75452-03-9
4. Hermens, H.J., Nene, A.V., Zilvold, G. Electrophysiological Kinesiology, Proceedings of the 11th congress of the International Society of Electrophysiology and Kinesiology in Enschede, the Netherlands 1996. October 1996, ISBN 90-75452-04-7
5. Van Harten, W.H. Bouwen aan een kwaliteitssysteem in de revalidatiezorg. Een poging tot constructieve technology assessment van een kwaliteitssysteem in een gezondheidszorginstelling. December 1997, ISBN 90-75452-07-1
6. Baardman, G., IJzerman, M.J. Design and evaluation of a hybrid orthosis for people with paraplegia. November 1997, ISBN 90-75452-8-X
7. Hutten, M.M.R. Lumbar Dynamometry: A useful method for assessment of patients with chronic low back pain? November 1999, ISBN 90-75452-13-6
8. Van der Salm, A., van Harten W.H., Maathuis, C.G.B. Ketenkwaliteit Cerebrale Parese Zorg. Een beschrijving van de cerebrale parese zorg en mogelijke verbeteringen hierin. April 2001, ISBN 90-75453-19-5
9. Nederhand, M.J. Muscle activation patterns in post traumatic neck pain. March 2003, ISBN 90-75452-27-6
10. Jannink, M.J.A. Usability of custom-made orthopaedic shoes in patients with degenererative disorders of the foot. September 2004, ISBN 90-75452-28-4
11. Blokhorst, M.G.B.G. State-dependent factors and attention in whiplash associated disorder. January 2005, ISBN 90-365-2111-4
12. Buurke, J.H. Walking after stroke. Co-ordination patterns & functional recovery. February 2005, ISBN 90-365-140-8
13. Van der Salm, A. Spasticity reduction using electrical stimulation in the lower limb of spinal cord injury patients. October 2005, ISBN 90-365-2253-6
14. Snoek, G.J. Patient preferences for reconstructive interventions of the upper limb in tetraplegia. December 2005, ISBN 90-365-2255-2
15. De Kroon, J.R. Therepeutic electrical stimulation of the upper extremity in stroke. December 2005, ISBN 90-365-2269-2
16. Van Dijk, H. Motor skill learning, age and augmented feedback. March 2006, ISBN 90-365-202-9
17. Mes, C.A.J. Improving non-optimal results in chronic pain treatment. January 2007, ISBN 90-365-2435-0
18. Voerman, G.E. Musculoskeletal neck-shoulder pain: a new ambulant myofeedback intervention approach. March 2007, ISBN 90-365-2460-1

19. Kallenberg, L.A.C. Multi-channel array EMG in chronic neck-shoulder pain. March 2007, ISBN 90-365-2459-8
20. Huis in 't Veld, M.H.A. Work-releated neck-shoulder pain. The role of cognitive-behavioural factors and remotely supervised treatment. December 2007, ISBN 978-90-365-2584-8
21. Fleuren, J.F.M. Assessment of spasiticity. From EMG to patients' perception. October 2009, ISBN 978-90-365-2869-6
22. Reenalda, J. Dynamic sitting to prevent pressure ulcers in spinal cord injured. October 2009, ISBN 978-90-365-2884-9
23. Prange, G.B. Rehabilitation robotics. Stimulating restoration of arm function after stroke. October 2009, ISBN 978-90-365-2901-3
24. Vos-van der Hulst, M. Prognostic factors and underlying mechanisms in chronic low back pain. February 2010, ISBN 978-90-365-2881-8
25. Kottink-Hutten, A.I.R. Assessment of a two-channel implantable peroneal nerve stimulator post-stroke. February 2010, ISBN 978-90-365-2959-4
26. Van Weering, M.G.H. Towards a new treatment for chronic low back pain patients. Using activity monitoring and personalized feedback. May 2011, ISBN 978-90-365-3305-8
27. Gulmans, J. Crossing boundaries. Improving communication in cerebral palsy care. February 2012, ISBN 978-90-365-3305-8
28. Molier, B.I. Influence of augmented feedback on learning upper extremity tasks after stroke. March 2012, ISBN 978-90-365-3293-9
29. Dubbeldam, R. Towards a better understanding of foot and ankle kinematics in rheumatoid arthritis. The effects of walking speed and structural impairments. October 2012, ISBN 978-90-365-3407-9
30. Evering, R.M.H. Ambulatory feedback at daily physicial activity patterns. A treatment for the chronic fatigue syndrome in the home environment? April 2013, ISBN 978-90-365-3512-0
31. Malhorta, S. Does spasticity interfere with functional recovery after stroke? A novel approach to understand, measure and treat spasticity after acute stroke. November 2013, ISBN 978-90-365-2567-0
32. Tabak, M. Telemedicine for patients with COPD. New treatment approaches to improve daily activity behaviour. February 2014, ISBN 978-94-6108-590-0
33. Trompetter, H.R. ACT with pain. Measurement, efficacy and mechanisms of Acceptance and Commitment Therapy. September 2014, ISBN 978-90-365-3708-7
34. Op den Akker, H. Smart tailoring of real-time physical activity coaching systems. October 2014, ISBN 978-90-365-3762-9
35. Jansen-Kosterink, S.M. The added value of telemedicine for physical rehabilitation. December 2014, ISBN 978-90-823196-0-6
36. Velstra, I.M. Advanced insights in upper limb function of individuals with cervical spinal cord injury. December 2015, ISBN 978-90-365-3929-6
37. Kloosterman, M.G.M. Keep on rolling. Functional evaluation of power-assisted wheelchair use. June 2016, ISBN 978-90-365-4299-9
38. Prinsen, E.C. Adapting to change. Influence of a microprocessor-controlled prosthetic knee on gait adaptations. December 2016, ISBN 978-90-365-4206-7
39. Wolvers, M.D.J. A coach in your pocket. On chronic cancer-related fatigue and physical behavior. March 2017, ISBN 978-90-365-4299-9
40. Cabrita, M. Active and pleasant ageing supported by technology. November 2017, ISBN 978-90-365-4407-8
41. Haarman, J.A.M. TIBAR. Therapist Inspired Balance Assisting Robot. November 2017, ISBN 978-90-365-4407-8

42. Nijenhuis, S.M. Roll up your sleeves! Technology-supported arm and hand training at home after stroke. April 2018, ISBN 978-90-365-4510-5
43. Cranen, K. Acceptance of telerehabilitation in chronic pain: the patients' perspective. June 2018, ISBN 978-90-365-4555-6
44. Boerema, S.T. Sensing human activity to improve sedentary lifestyle. September 2018, ISBN 978-90-365-4604-1
45. Radder, B. The wearable hand robot. Supporting impaired hand function in activities of daily living and rehabilitation. November 2018, ISBN 978-90-365-4658-4
46. Krabben T. A reaching hand. Towards an active therapeutic device for the upper extremity following stroke. December 2018, ISBN 978-90-365-4660-7
47. Timmerman, J.G. Cancer rehabilitation at home. The potential of telehealthcare to support functional recovery of lung cancer survivors. January 2019, ISBN 978-90-4701-7
48. Nikamp-Simons, C.D.M. The sooner the better?! Providing ankle-foot orthoses in the rehabilitation after stroke. May 2019, ISBN 978-90-365-4747-5
49. Achterkamp, R. Towards a balanced and active lifestyle. June 2019, ISBN 978-94-6323-656-0
50. Engbers, C. Keep Cycling. How technology can support safe and comfortable cycling for older adults. September 2019, ISBN 978-90-365-4848-9
51. Ommeren, A.L. Offering a helping hand. Getting a grip on needs and preferences of stroke patients regarding soft-robotic technology supporting hand function. October 2019, ISBN 978-90-365-4835-9
52. Ter Stal, S. Look who's talking. Appearance of embodied conversational agents in eHealth. March 2021, ISBN 978-90-365-5126-7
53. Beinema, T.C. Tailoring coaching conversations with virtual health coaches. December 2021, ISBN 978-90-365-5260-8
54. Hurmuz, M. Z. M. eHealth-In or out of our daily lives?: Measuring the (non-) use of eHealth in summative evaluations. June 2022, ISBN 978-90-365-5360-5
55. Broekhuis, M. Meet my HUBBI: he's an expert on eHealth usability. September 2022, ISBN 978-90-365-5443-5
56. Schulte, R. V. Up to one's knees in data: Data-driven intent recognition using electromyography for the lower limb. December 2022, ISBN 978-90-365-5486-2
57. Bessler-Etten, J. Safety first in rehabilitation robots! Investigating how safety-related physical human-robot interaction can be assessed. January 2023, ISBN 978-90-365-5503-6
58. Zandbergen, M. A. Moving forwards by going outside: Inertial measurement unit-based monitoring of running biomechanics. February 2023, ISBN 978-920-365-5513-5
59. Marotta, L. Development of inertial sensor-based methods to assess physical fatigue in running applications. February 2023, ISBN 978-90-365-5507-4
60. Tenniglo, M. J. Stiff knee gait in stroke: Walking down the road of different treatment options. December 2023, ISBN 978-94-6419-964-2
61. Oberschmidt, K. Who, When, How - Guiding the active involvement of stakeholders in eHealth Action Research. January 2024, ISBN 978-90-365-5866-2
62. Rikhof, C. Stimulating rehabilitation - A new hybrid rehabilitation device combining robot and functional electrical stimulation for early stroke rehabilitation. December 2024, ISBN 978-90-365-6353-6
63. Brandl, L. Guiding older mourners from online self-help to offline support using algorithmic mental health monitoring. January 2025, ISBN 978-90-365-6436-6

&

Roessingh
Research and Development
Progress in rehabilitation science

63

UNIVERSITY OF TWENTE.