2023

Psychiatric Diagnostic Decision-Making: Investigating the Theory of the Dual-Process Model

Christopher S. Kleva
Virginia Commonwealth University

Psychiatric Diagnostic Decision-Making:

Investigating the Theory of the Dual-Process Model

A thesis submitted in partial fulfillment of the requirements of the degree of

Master of Science at Virginia Commonwealth University

By: Christopher S. Kleva, M.A.

Master of Arts, Stony Brook University, May 2016

Bachelor of Arts, King's College, May 2014

Director: Jared Keeley, Ph.D.

Associate Professor of Psychology

Virginia Commonwealth University

Richmond, VA

April 2023

Acknowledgements

I am grateful for the support of many people while working on this project, including faculty members, colleagues, and friends. I am particularly thankful for the support of Dr. Jared Keeley, my research mentor and committee chair. I am grateful for his unwavering support, patience, and thoughtful guidance both on this project and throughout my graduate school training. I am also thankful to my other committee members, Bryce McLeod Ph.D., Kia Bentley PhD. and Gary Cuddeback, Ph.D., for their time, commitment, and valuable perspectives. Finally, I would not have been able to accomplish this milestone without the love and support of my amazing wife, Alyssa, and father, Ignacio, who have always supported my dreams. My father passed away before he could see me achieve this feat; however, I know he is looking down proud of me and I will continue to make him proud.

Table of Contents

List of Tables

List of Figures

Abstract

PSYCHIATRIC DIAGNOSTIC DECISION-MAKING: INVESTIGATING THE THEORY OF THE DUAL-PROCESS MODEL

By: Christopher S. Kleva, M.A.

A thesis submitted in partial fulfillment of the requirements of the degree of Master of Science at Virginia Commonwealth University.

Director: Jared Keeley, Ph.D.

Associate Professor of Psychology

Diagnostic decision-making is an important component of clinical practice; however, there is substantial diagnostic unreliability within mental health diagnoses. The lack of reliability emphasizes the importance of investigating diagnostic decision-making; however, the research to date is limited, primarily relying on a vague definition of decision-making based on the dual-process model. The present study is an exploratory attempt to apply the dual-process model to explain how mental health clinicians (n = 30, 73.3% cisgender female, 96.7% psychologists) arrive at making diagnostic decisions through the use of an interactive interview mechanism. For each participant, we are able to create a figure that displays their decision-making process. The number of times participants indicated System 2 processing based on diagnostic questions and certainty rating each ranged from 0 to 9 (M = 3.10, SD = 2.41 & M = 3.07, SD = 2.53, respectively). The number of questions asked (2 to 27, M = 9.30) were significantly correlated with switching diagnostic questions ($r(28) = .59$, $p < .001$) and switching their diagnostic certainty rating ($r(28) = .57$, $p < .001$). Findings provide evidence that the interactive interview is an effective mechanism to operationalize the dual-process model. The public health significance and implications of the interactive interview lies in the potential for individual clinicians and mental health teams to better understand their diagnostic decision-making process and improve overall accuracy. Future research may provide further support for the paradigm to offer

individualized training and consultation for clinicians and mental health teams to improve their decision-making process.

Keywords: clinician cognition, diagnosis, decision-making, dual process model

Introduction

Diagnostic decision-making is an important component of clinical practice, helping guide case conceptualization, selecting effective treatment interventions, and monitoring appropriate outcomes. Unfortunately, there is substantial diagnostic unreliability within mental health diagnoses (Narrow et al., 2013). A source of unreliability in diagnostic outcomes is that clinicians can arrive at different conclusions given the same information. This lack of reliability emphasizes the importance of investigating diagnostic decision-making; however, the research to date is limited. Past research has largely only focused on medical diagnoses while relying on a vague definition of diagnostic decision-making based on the dual-process model, a common model of thinking and decision-making.

The dual process model states that decision makers operate in two systems, System 1 and System 2. While the dual process model evolved out of cognitive psychology, it has influenced a number of other fields including behavioral economics (Kahneman & Tyversky, 1979), learning (Schnieder & Shiffrin, 1977), and social psychology (Devin, 1989). More recently, this model has been used to investigate clinical decision-making within medical settings. Croskerry (2009) and Norman et al. (2014) have used the dual process model as a way to explain diagnostic failures with the intention of increasing diagnostic reliability. Unfortunately, there are several limitations to the current literature on the dual-process model of decision-making.

By and large, current literature defines the dual-process model of decision making as conceptual frameworks of verbal theories. The central discriminant factor between System 1 and System 2 is reaction time; however, this operationalization is lacking. Studies have attempted multiple methodologies to operationalize the dual-process model of decision-making including: manipulating instructions (Bieleke et al., 2017), relying on participant self-report

(Walco & Risen, 2017), reaction time measurement (Markovits et al., 2013), and physiological measurements such as neuroimaging (Diederich & Trueblood, 2018) and skin conductance (Richards et al., 2018). Unfortunately, these different operationalizations are not without shortcomings and each will be addressed.

The proposed study will apply the dual-process model of decision-making to investigate how mental health clinicians arrive at making diagnostic decisions and, more specifically, improve the operationalization of the dual-process model. Alternative theories of decision-making were considered such as the "classical decision-making (CDM) paradigm"; however, the model posits that decision makers operate in a world of complete certainty in which they are presented with a specific problem, contemplate all choices and potential consequences prior to making a decision (Djulbegovic et al., 2018). The CDM model does not take into account unpredictable situations and fast-paced environments in which past experience, bias and heuristics can affect the decision-making process. The dual-process model takes into account bias and heuristics (e.g., availability bias).

Public Health Significance

Unfortunately, there is a vast degree of diagnostic unreliability (Regier et al., 2013), both the over and under diagnosis of certain mental disorders (Mitchell et al., 2009). The public health significance of psychiatric misdiagnosis lies in the impact on the overall health, quality of life and length of life of those individuals misdiagnosed. The initial misdiagnosis can result in the delay of appropriate treatments and harmful prescribing of pharmaceuticals, thereby leading to the worsening of symptoms of various mental disorders, including but not limited to bipolar disorder (Bowden, 2005; Ghaemi et al., 2000), depression (Mojtabai, 2013; Vermani et al., 2011), and obsessive-compulsive disorder (Glazier et al., 2015). For example, bipolar disorder

has been commonly misdiagnosed as unipolar depression and, on average, 48% of these individuals are not properly diagnosed until ten years after the first time they sought treatment (Berk et al., 2007; Hirschfeld et al., 2003). The improper prescribing of antidepressants when bipolar disorder is misdiagnosed as unipolar depression has been found to cause new or accelerated rapid cycling of hypomania or mania (Goldberg & Truman, 2003; Ghaemi et al., 2000). In addition, misdiagnosis leads to increased financial burden on society as well as on the individual as symptoms worsen impacting overall functioning and causing social and occupational impairment (Bessonova et al., 2020). For example, in 2013, the US economic burden of schizophrenia was reported to be US$155.7 billion (Cloutier et al., 2016), which had more than doubled since 2002, when the economic burden was US$62.7 billion (Wu et al., 2005). These numbers are alarming considering a recent study conducted at John Hopkins Bayview Medical Center found that over 50% of individuals with a primary diagnosis of schizophrenia were later re-diagnosed, most commonly with anxiety or mood disorders (Coulter et al., 2019).

It is also important to note that there are racial and sexual orientation-related disparities among those who are more likely to be misdiagnosed. For example, individuals of African descent are more often misdiagnosed with another disorder, such as schizophrenia, when bipolar disorder is the correct diagnosis, compared to those of European descent (Akinhanmi et al., 2018; Bailey et al., 2009). Gara et al. (2019) also found that African Americans are more often diagnosed with schizophrenia, even when full criteria for mood disorders are also met. They go on to emphasize the significant consequences of misdiagnosis, specifically treatment nonadherence or dropout, delayed treatment response, worsening morbidity and increased risk of suicide. In addition, inappropriate prescribing of antipsychotic medications can lead to increased

risk of adverse side effects including weight gain, diabetes, and certain movement disorders. Krishan Aggarwal et al. (2012) found that racial minorities diagnosed with schizophrenia are more likely to receive long-acting antipsychotic injections. They suggest that this is due to a biased belief that racial minorities are less adherent to antipsychotic medications. With respect to sexual orientation-related disparities, borderline personality disorder (BPD) is found to have been more commonly diagnosed in sexual minorities, more specifically bisexual individuals, compared to their heterosexual counterparts (Rodriguez-Seijas et al., 2021; Reuter et al., 2016). The potential overdiagnosis of BPD in sexual minorities further perpetuates stigma and marginalization while also further delaying appropriate treatment.

Diagnostic Decision-Making

Diagnostic decision-making is the process by which clinicians gather information by asking questions to evaluate a client's symptoms in order to develop a diagnostic hypothesis which is used to inform treatment (Higgs & Jones, 2008; Norman, 2005). Due to the multiple theory-grounded processes that make up diagnostic decision-making, this definition warrants dissection: (a) deciding how to assess the person, such as selecting questions in an unstructured interview; (b) recognizing the person's symptoms; (c) matching the endorsed symptoms to potential diagnoses; and (d) evaluating the fit of competing potential diagnoses, or competing hypotheses. It is important to note that the last step of diagnostic decision-making is choosing an appropriate treatment; however, this process is beyond the focus of this study, and we will only focus on the first four steps of the process.

The most common mechanism by which mental health professionals assess individuals is through diagnostic interviewing (Segal & Hersen, 2010). The three distinct types of diagnostic interviews are: (1) structured, (2) semi-structured, and (3) unstructured. The types of questions

that are asked by the interviewer is dictated by the format of the interview. Structured interviews provide the exact verbiage and series of questions that must be executed verbatim by the interviewer. Semi-structured interviews offer the same structure and standardization of questions and probes; however, interviewers have the flexibility to augment the interview by asking their own questions to further clarify the clinical profile (Miller, 2019). Unstructured interviews, as the name suggests, provides the interviewer "maximum latitude regarding what questions to ask, how to probe symptom patterns, and how much time to spend on different subject matters" (Miller, 2019, p. 31). While unstructured interviews are the most commonly used interviews by clinicians, there are major disadvantages including poor reliability and validity as well as a greater degree of susceptibility to biases (Rogers, 2001). The process of diagnostic interviewing, regardless of the type of interview, requires the interviewer to be able to identify diagnostic cues, signs and symptoms consistent with DSM-5 disorders (American Psychiatric Association, 2013a; Jones, 2010). This process is known as pattern recognition.

Primarily emerging from the field of cognitive psychology, pattern recognition refers to the process of classifying stimuli into distinct categories based upon previous knowledge and experience (Reed, 1972). For example, while conducting an unstructured clinical interview, the patient reports to the interviewer that they have been struggling with feelings of prolonged sadness and have lost interest in activities that use to bring them joy. The clinician, who has over twenty years of experience working with patients diagnosed with Major Depressive Disorder (MDD), recognizes that the two symptoms endorsed represent a pattern that is consistent with MDD. Pattern recognition is shown to be a key strategy used by clinicians to identify important clinical information and categorize said information into "an organized knowledge structure acquired and evolved through learning and extensive clinical experience" (Fernando et al., 2012,

p. 122). In medicine, the specific pattern for distinct disorders used by clinicians are “illness scripts”, which store information about prototypical patients or past experiences (Schmidt, Norman, & Boshuizen, 1990). A common theory explaining this process in psychiatry is prototype-matching (Genero & Cantor, 1987; Cantor et al., 1980).

A prototype refers to “a clinician’s conception of a hypothetical client who best exemplifies a particular disorder” (Garb, 2005, p. 71). According to the theory, clinicians make diagnostic decisions by comparing their patients to a prototype. The possibility that clinicians may be diagnosing using their own personally defined prototypes is important to recognize and may explain low levels of interrater reliability and diagnostic accuracy in psychiatric diagnoses. Additionally, the process of prototype-matching has been shown to categorize typical presentations more efficiently than atypical presentations, further explaining misdiagnosis (Garb, 2005). To better explain the concept of typical and atypical presentations, Cantor et al. (1980) discuss furniture. If we were to believe the prototype of furniture has four legs and a flat surface, we would think of a table and chair as typical presentations of the prototype. Unfortunately, according to the features we attribute to the prototype, a rug and desk lamp would be considered atypical because they do not have four legs and a flat surface. With respect to mental disorder, sadness may be considered a typical feature of a depression prototype; however, a poor memory may be considered an atypical feature.

Kim and Ahn (2002) posit the theory-based model of categorization which states that clinicians view the diagnostic process as causal theories, i.e., as a set of interlinked causes for their current problem. For example, participants were more likely to diagnose a hypothetical patient with a disorder if they endorsed causally central symptoms, those which directly map onto the participants’ theoretical representation (prototype) of the disorder, compared to causally

peripheral or isolated symptoms. Causally central symptoms and peripheral causal symptoms can be conceptualized similarly to typical and atypical presentations, as discussed above. It is important to note that during delayed recall, participants were more likely to attend to symptoms that were causally central to their own diagnostic hypothesis. The false memories suggest that participants may be biased to rely on their own personal theories especially when the symptomatology is ambiguous and conflicts with their own pattern recognition. Additionally, when symptomatology is ambiguous and prototype-matching encounters an atypical presentation, individuals may be forced to evaluate the fit of competing potential diagnoses, or competing hypotheses.

The hypothetico-deductive method requires individuals to assess symptom patterns and confirm their fit with competing diagnoses (Elstein, 1994). This has been shown to be the preferred method for less experienced clinicians (Kassirer & Kopelman, 1991). An alternative approach to deciding between competing hypotheses is using diagnostic decision trees. First (2014) offers a systematic approach to consider differential diagnoses, reflecting the DSM-5's (American Psychiatric Association, 2013a) organization and structure of mental disorders. The best approach to arriving at a diagnostic conclusion is a bottom-up approach, by which one begins the diagnostic process by considering the presenting problem. Decision points, whether or not a sign or symptom is present, continue throughout the tree, providing further support for specific diagnoses and ruling out others until a final diagnosis has been selected.

The definition of diagnostic decision-making provided above has been commonly used within the literature investigating the role of the dual process model in diagnostic decision-making among medical professionals (Pelaccia et al., 2011; Croskerry, 2009). In fact, the

prototype-matching theory involving typical presentations can be tied to System 1 processing while atypical presentations necessitate System 2 processing.

Dual-Process Model

Decades of research on reasoning and decision-making have come to agree on the presence and interaction of two distinct systems, typically referred to as System 1 and System 2 (Evans, 2008). System 1, considered the "intuitive system", operates automatically and instantly, with little to no conscious effort. Everyday activities that are considered automatic and are attributed to System 1 include selecting your favorite cereal to purchase at the grocery store, or which gas to fuel your vehicle. System 2, considered the "deliberate system", operates rationally and slowly, utilizing conscious mental processes prior to arriving at a decision. We often rely on System 2 when we are attempting to organize an introduction for a manuscript or selecting courses to take that will fit into your schedule for the upcoming semester. While there has been a general consensus on the existence of the dual-process model, there is a great deal of disagreement on the operationalization and interaction of the two systems.

Dual-Process Model 2.0

De Neys and Pennycook (2019) proposed an alternate conceptualization of the two systems. The traditional dual-process model marks System 1 and System 2 as separated by "intuitive" and "deliberate" thinking; however, theorists have put forth the idea that decisions in System 1 can be both intuitive and deliberate. Prior to making a decision, we often rely on traditional heuristics to arrive at an answer which can be described as an "intuitive" response. The intuitive response can be a logical and deliberate decision using little to no conscious effort. The instinctual and quick response is based on one's experience, relying on previous knowledge

and basic probability principles. The key difference in this alternative conceptualization is that, although System 1 and System 2 both result in deliberate decisions, System 2 requires more conscious effort and time to arrive at a decision. For example, an experienced psychologist working on the inpatient unit at a hospital meets a new patient and upon meeting them for the first time offers a preliminary diagnosis of schizophrenia. The psychologist relied on their experience as well as knowledge of the base rate of schizophrenia at the facility to make a logical and deliberate decision quickly. Unfortunately, by relying on heuristics and past experiences, System 1 thinking can result in biases and, ultimately, may result in errors in decision making (Croskerry, 2013; Croskerry, 2008). By the psychologist in the above example relying on past experience and base rates, they may be at risk of providing an inaccurate diagnosis. For example, it is possible that the psychologist neglected to consider medical causes for the symptoms (e.g., low blood sugar, encephalitis) or forgot to inquire about the patient's work history where they would have learned they were exposed to toxic industrial chemicals.

The alternative conceptualization of the dual-process model has led researchers to offer alternative models of how the two systems interact with one another. Do the two systems operate in a specific order or in parallel? The Default Interventionist theory proposes that there is a sequential ordering in the activation of the two models (Evans, 2008). The initial answer or decision that is generated is the result of System 1, and System 2 is only activated when further thought or deliberation is required. Recently, a hybrid model has been proposed that recognizes that a System 1 decision can result from both logical and heuristic processing. Additionally, it acknowledges that System 2 is optional, and activation is not always necessary. The Three Stage Dual-Process Model (Pennycook et al., 2015) provides reasoning for whether System 2 is activated. If an individual is presented a problem that elicits two System 1 responses (Stage 1),

conflict detection occurs because there are two competing responses (Stage 2). The conflict leads to the activation of System 2, in hopes of a resolution (Stage 3).

Theoretically, the Three Stage Dual-Process Model can explain the process by which many mental health clinicians go about making diagnostic decisions. When meeting a client for the first time, it is common for clinicians to have an initial set of diagnostic questions that they ask to all new clients. For example, as the clinician gathers further information, they begin to develop a preliminary diagnosis of major depressive disorder and alter questions to confirm said diagnosis. This would be considered Stage 1. As the clinician gathers more information that resembles patterns of symptomatology from previous patients, it helps to confirm their preliminary diagnosis and become more confident. However, during the process of asking questions, the clinician may receive a response that conflicts with their preliminary diagnosis, thereby, activating Stage 2. The clinician will then begin to recall differential diagnoses such as normative grief or sadness to determine the next line of questioning that will help arrive at the correct diagnosis. This conflict detection activates System 2 and the final stage of the model, Stage 3.

Unfortunately, the progression through stages are not always as systematic as presented. When a conflict is detected between competing ideas, the strength of the conflict determines whether System 2 will be activated. Further complication results from biased intuitions.

Bago and De Neys (2017) discovered that if an individual is confident enough in their System 1, initial intuitive response, if given the opportunity to engage in System 2, to rethink and potentially change their answer, they will fail to do so. The likelihood of staying in System 1 and not engaging System 2 is further exacerbated by heuristics and biases. Although heuristics provide the benefit for individuals to be able to make decisions quickly with less cognitive

energy, the payoff is that taking these mental shortcuts predispose individuals to ignore contradictory information. Furthermore, reliance on previous experience paired with unconscious biases directly affects clinical judgement, leading to diagnostic errors (Croskerry et al., 2013).

Limitations of Operationalization

As mentioned earlier, current literature defines the dual process model of decision-making as conceptual frameworks of verbal theories. Researchers have utilized a variety of methodologies to assess the existence of the dual-process model; however, each is not without its own limitations. These methods and their limitations are discussed below.

Reaction Time. The central discriminant factor between System 1 and System 2 is reaction time; however, this operationalization is inadequate. When attempting to capture each distinct system in action when making a decision, studies will often apply arbitrary thresholds for when System 1 processing ends and System 2 processing begins (De Neys, 2021). Markovits et al. (2013) applied such arbitrary thresholds in two experiments in which participants evaluated the validity of a conclusion that followed a series of supporting statements. In the first experiment, participants in the limited time condition were given 18 seconds to assess the validity of the conclusion while the other participants were given unlimited time. Researchers justified applying the 18 second time restriction based on pre-testing which concluded 18 seconds was sufficient to read the full series of statements with little time to deliberate before answering. In the second experiment, the limited time condition allowed participants 20 seconds to read the statements and provide an answer. Besides these time restrictions being arbitrary, applying such thresholds assumes that all participants move through System 1 and System 2 processing the same and ignores the individual differences in processing speed between participants.

Instructional Manipulation. Bieleke et al. (2017) were interested as to whether an individual's social value orientation, the way in which a person divides resources between themselves and others and weighs their welfare relative to the welfare of others, moderates the effects of intuitive versus reflective processing in an ultimatum game experiment. Participants were divided into either the control group or the experimental group, which consisted of the intuitive, System 1, or reflective, System 2, condition. The researchers directly manipulated the instructions for the participants to have a choice as to which strategy they would use for the ultimatum game. Participants were asked to either choose between the intuitive and neutral strategy or the reflective and neutral strategy. In anticipation that participants using the intuitive strategy may start deliberating for an extending period of time, researchers offered a plan to overcome this hurdle. Participants were encouraged to say to themselves, "If I start pondering at length, then I will tell myself: Listen to your guts!" Likewise, participants using the reflective strategy were advised to repeat the following statement, "If I start acting in a hasty way, then I will tell myself: Use your brain!" Presumably, these instructions would keep someone from using one system or the other.

Another method of instructional manipulation used by researchers to capture System 1 and System 2 processing is using the two-response paradigm. Thompson et al. (2011) provided explicit instructions to their participants to elicit which system they should use. Participants were told to complete the tasks as fast as possible, which is presumed to activate System 1, and were then able to return to the task later on, taking as much time as they need with the ability to change their response, which is presumed to be indicative of System 2. After providing each response, participants were asked on a scale of 1, "Guessing", to 7, "Certain I'm Right", on their

Feeling of Rightness. This component will be adopted and adapted for the present study which will be discussed further.

The limitation of operationalizing System 1 and System 2 by direct instructional manipulation is that researchers cannot conclude with a high degree of certainty that participants operated within their chosen condition. Additionally, for example, if a participant chose to operate intuitively but for each trial had to repeat the phrase, recognizing that they were no longer operating in System 1 would entail operating in System 2. This confounding variable limits the reliability of measuring and capturing System 1 and System 2 processing.

Self-Report. Walco and Risen (2017) used a unique self-report design to study the role of acquiescence within the dual-process model of decision-making. Acquiescence is the phenomenon by which an individual explicitly recognizes that their intuitive decision is wrong; however, they refuse to alter their decision. The researchers tested this phenomenon using four different decision-making scenarios (e.g., lottery game, blackjack). All participants were educated on the difference between intuitive responses, presented as "consulting the gut", and rational responses based in reason, as well as how the two responses may differ at times. Participants were then introduced to a decision-making scenario and were asked to answer two questions prior to making their decision: (a) which option was more intuitively appealing, and (b) which option was the more rational choice. They were also asked how important it was for them to make decisions based on intuitive or rational analysis as well as how they would respond if there was a conflict between the two responses. Results from all four experiments support the existence of acquiescence. Nevertheless, the operationalization of the dual-process model using self-report is subject to the limitation of being reactive to experimental instructions. After participants were educated on the differences between intuitive and rational responses and

presented questions, it is possible that participants were influenced by the instructions and responded according to how they assumed researchers wanted them to respond.

Neuroimaging. As is common in the field of cognitive psychology, many theories have been tested under a neurobiological lens by applying neuroimaging and the dual-process model of decision-making is no different. De Martino et al. (2006) used functional magnetic resonance imaging (fMRI) to investigate which areas of the brain were activated when individuals are presented with framing effects in risky decision making. Framing effects refer to the bias individuals have to make decisions based on if the decision is presented as a positive (gain) or a negative (loss). The researchers concluded that framing effects were associated with increased activation in the amygdala, while the reduction of framing effects were associated with increased activation in the orbital and medial prefrontal cortex (OMPFC). Guo et al. (2017) showed that an increase in framing effects results from fast and intuitive System 1 processing. While it would be naïve to deny there are specific brain regions related to decision-making, it is harder to conclude which brain regions are activated during System 1 and System 2 processing. As was just demonstrated, conclusions are being drawn by pairing separate findings to justify the presence of specific brain regions during System 1 processing. Further, neuroimaging is expensive and impractical for widespread use, especially with hard-to-access populations like mental health professionals.

Skin Conductance. Richards et al. (2018) proposed that operating in System 1 increases the likelihood of the disposition effect, which states that individuals are more averse to experiencing significant losses rather than experiencing the enjoyment of significant gains (Shefrin & Statman, 1985). Goulart et al. (2013) investigated the relationship between the disposition effect and psychophysiological processes such as skin conductance. They found that

individuals with a high disposition effect also exhibited higher skin conductance responses. This finding suggests that emotional arousal is tied to System 1 processing. The limitation of using skin conductance as a proxy for this relationship is that there is an assumption that the relationship between emotional arousal and System 1 processing is universal and specific, ignoring individual differences and outside factors that may increase or decrease one's emotional arousal and skin conductance.

Dynamic Stochastic Dual-Process Model

Diederich and Trueblood (2018) proposed a specific framework for investigating the dual-process model: the dynamic stochastic model. This model was designed to measure how decision makers change preferences over time (dynamic) between two potential options (stochastic). Through computational modeling, the researchers are able to explain the timing and interaction between the two systems. Decision makers gather evidence over time regarding two distinct options until they have reached a threshold where enough information has been gathered that a choice can be made between the two options. Unique to this model is that it can account for choice and response-time data, addressing the various limitations seen in traditional, verbal theories of the dual-process model. Additionally, the stochastic model can explain whether the two systems interact simultaneously or sequentially, as seen in the Three Stage Model.

The foundation of the stochastic model is attributing System 1 choices to Prospect Theory (PT; Kahneman & Tversky, 1979) and System 2 choices to Expected Utility Theory (EU; vonNeumann & Morgenstern, 1953). PT suggests that decision makers base their choices on perceived gains opposed to perceived losses. This is known as "loss aversion." Tversky and Kahneman (1992) proposed that losses are more emotionally involved than gains. Additionally, there is also a framing effect that is at play. For example, if an individual is given two potential

options, both equal in value, but one is presented in terms of potential gains and the other in terms of potential losses, the individual is more likely to choose the option with the potential gains (Weller, Levin, & Denburg, 2011). EU is typically conceptualized as decision making under uncertainty. When a decision maker (DM) is unsure of the outcome of any one decision, they will weigh the average of all potential outcomes and make a decision based on the highest expected utility. A common example is insurance. DMs accept that the expected value of paying monthly insurance is superseded by the possibility of large-scale losses if something were to happen and they did not have insurance. This is known as "risk aversion."

Diederich and Trueblood (2018) tested their model using a gain-loss paradigm presented in De Martino et al. (2006). DMs were presented with different starting monetary amounts (e.g., $25, $50, $75, $100) and were then asked to choose between a sure option to keep a portion of the initial amount or a gamble option to possibly keep all of the initial amount. The gamble options were offered in four different probabilities: 20%, 40%, 60%, and 80%. All DMs were presented with six trials. Whether or not DMs chose the sure or gamble option depended on the frame. When the task was framed as a gain, DMs were more likely to prefer the sure option, opposed to when the task was framed as a loss, DMs were more likely to prefer the gamble option. For example, the sure option consisted of keeping $20 of the original $50 and the gamble option consisted of potentially losing $30 of the original $50. Both options are mathematically the same, but the frame resulted in DMs treating them differently. Researchers found that the longer a DM operated in System 1, the stronger the framing effect. DMs operated in System 2 when the gamble and sure options are the same and the risk is neutral. For example, Figure 1 represents the model from six individual trials when the task is framed as a gain. The upper and lower limits represent the gain and sure option, respectively, with the vertical dotted line

representing the arbitrary threshold defining System 1 and System 2. In this experiment, two trials chose the sure option while operating in System 1 while the other four trials made their choice in System 2. This model will be the basis for the proposed study.

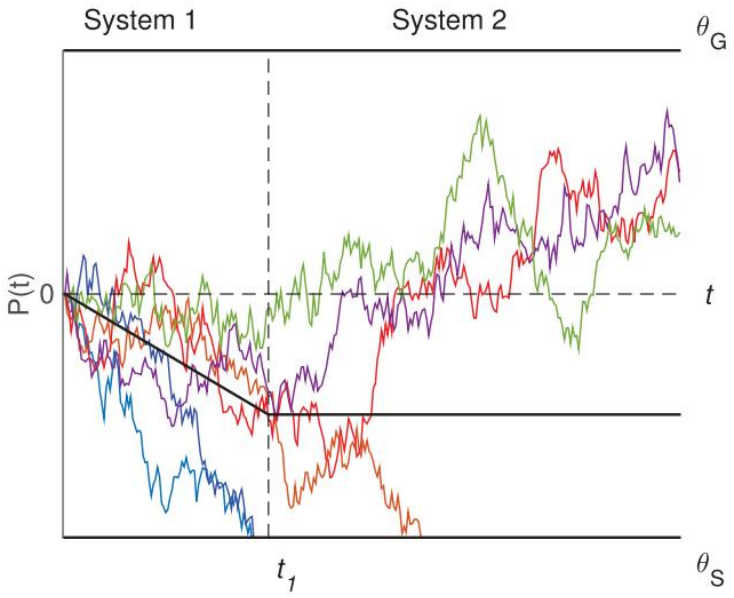


Figure 1. Example of two-stage dynamic-stochastic dual-process model

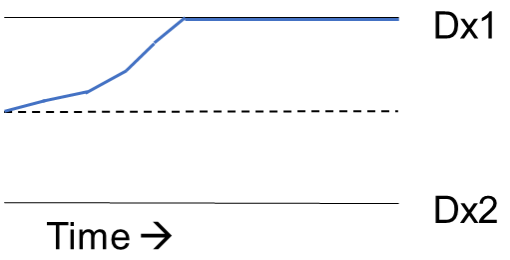
Note: Diederich and Trueblood (2018)

Proposed Study

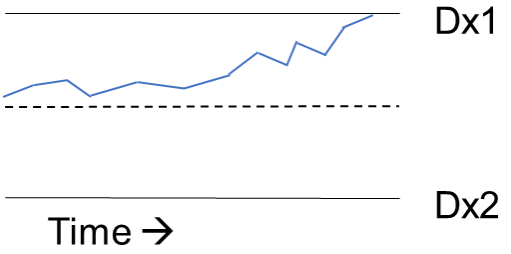
Previous research has struggled to move beyond defining the dual-process model of decision-making as simply a conceptual framework of verbal theories. A primary reason is the weak operationalization of how the two systems differ. To date, researchers are unconvinced about how to define and measure the two systems and, therefore, have attempted various methods. As was addressed earlier, each of these methods have limitations. Additionally, there is limited research on the application of the dual-process model of decision-making to explain psychiatric diagnostic decisions, specifically.

The proposed study seeks to apply the dual-process model to explain how mental health clinicians arrive at making diagnostic decisions through the use of an interactive interview mechanism. More specifically, the general aim of the proposed study is an exploratory attempt to apply Diederich and Trueblood's (2018) dynamic stochastic model of the dual-process model to explain psychiatric diagnostic decisions. Through the use of the interactive interview mechanism, we intend to attempt to measure which system (System 1 or System 2) clinicians are using to make diagnostic decisions both overall and at any given moment. After receiving each new prompt or new piece of information within the interview, participants will be asked to complete two scales that will serve as indicators as to which system participants are operating. For the first scale, participants will be asked to rate whether they are leaning towards one diagnosis or the other, with the option of staying neutral in the middle of the scale. In the second scale, participants will rate their confidence of their diagnostic rating on the previous scale with a sliding scale of 0 to 100 (See Appendix A). Additionally, the interactive component of the interview will provide an additional point of operationalization. Participants will be given a choice of questions to ask after each new piece of information in the interview. Their choice of question will be an additional indication of which system they are using, as the questions can be either consistent or inconsistent with the diagnosis they have selected. For each participant, we will be able to create a figure that displays the decision-making process, similar to those in Diederich and Trueblood (2018). For example, Figure 2A represents a participant who was confident throughout the interactive interview. They asked confirmatory questions to further their confidence towards diagnosis one and likely operated primarily in System 1. Possible outcomes for participants operating in System 2 would look similar to Figure 2B and 2C. In the former, the participant's diagnostic certainty fluctuated consistently within a single diagnosis;

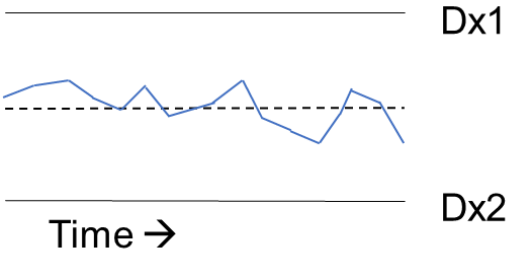
however, in the latter, the participant's certainty fluctuated between both potential diagnoses. In addition to applying the dual-process model of decision-making to psychiatric diagnostic decisions, using the interactive mechanism will also enable us to recognize patterns of diagnostic questioning and their relation to diagnostic certainty and confidence ratings.



A.



B.



C.

Figure 2. Expected dynamic-stochastic dual-process model for proposed study. Panel A: Participant operated primarily in System 1. Panel B: Participant fluctuated between System 1 and System 2; however, consistently within a single diagnosis. Panel C: Participant operated in both systems and diagnostic certainty fluctuated between both potential diagnoses.

Throughout the interactive interview participants are attempting to decide between two potential diagnoses: major depressive disorder and a normal grief response. These two differential diagnoses were chosen because of the high rate of comorbidity (Sung et al., 2013) and the longstanding debate between categorizing grief or bereavement as a diagnosable psychological disorder or as a normative response to loss (Maj, 2012; Wakefield & First, 2012; Wakefield, 2013). The "bereavement exclusion" (BE) for major depressive disorder (MDD), which was first introduced in the DSM-III (American Psychiatric Association, 1987), recognized depressive symptoms as a culturally, normative response for bereaved individuals. The BE was carried over to the DSM-IV (American Psychiatric Association, 2000); however, it was removed from the DSM-5 (American Psychiatric Association, 2013a). Ultimately, while the developers of the DSM-5 provided some guidance and insight as how to best differentiate MDD and normative grief, they noted, "this decision inevitably requires the exercise of clinical judgement…" (American Psychiatric Association, 2013b, p. 161). This study will capitalize on the difficult distinction between these two clinical presentations to investigate clinicians' decision-making process.

The public health significance and implications of the proposed paradigm lies in the potential for individual clinicians and mental health teams to better understand their diagnostic decision-making process and improve overall accuracy. The paradigm could be used to offer individualized training and consultation for clinicians and mental health teams, as a whole, to improve their decision-making process. The probabilistic evaluations of an individual's use of biases and heuristics would inform the individualized training. Additionally, feedback for clinicians can be provided for them to learn about their own biases/heuristics, and develop techniques to increase the frequency of considering and ruling out differential diagnoses.

Croskerry (2002) describes similar techniques, including pattern recognition, as a cognitive debiasing approach to improve clinical decision-making among Emergency Department clinicians. Furthermore, the individualized training and consultation would provide an added benefit if adopted by doctoral-level graduate programs. Research has shown that diagnostic inaccuracy is not a result of deficits in clinical knowledge but rather differences in the way clinicians think (Croskerry, 2017). Stark and Fins (2014) would argue that it is an ethical obligation of every graduate program to provide specific training in critical thinking and decision-making, in order to recognize and limit the effects of cognitive biases. If the methodology proposed in this project is successful at detecting the manner in which clinicians make their decisions, it could be used as an assessment and training tool in addition to its research applications.

Method

Participants

The proposed study is exploratory in nature and so an *a priori* hypothesis-based power analysis was not possible. Nonetheless, a reasonable goal would be to seek a sample that is large enough to create sufficient variability on clinicians' display of System 1 and System 2 processing. According to the Central Limit Theorem, an *n* of 30 would create a sampling distribution that approaches normality on any single measurement (Anderson, 2010). This is a practical sample size that would offer response variability.

Licensed psychologists were recruited from state licensing boards. The decision to only recruit psychologists rather than a diverse sample of mental health professionals is supported by the research showing there is no difference in diagnostic accuracy between professions (i.e.,

psychologists, counselors, social workers) or years of experience (Gaebel et al., 2020). Many of these licensee lists are publicly available or can be obtained by paying a small fee. We recruited a broad sample of licensed psychologists by obtaining licensee lists from six states: Arkansas, Maine, Minnesota, Nevada, Rhode Island, and Wisconsin. Our decision to recruit from these specific states was in part because each list contains email addresses. Our recruitment strategy involved emailing all licensees with a valid email address on the obtained lists. Our decision to solely contact those with an email address is because mailing recruitment letters would be cost prohibitive.

The total participant pool including all six state lists is 6957 participants. A breakdown of the number of psychologists within each state licensee list is as follows: Arkansas (n = 580), Maine (n = 168), Minnesota (n = 3690), Nevada (n = 458), Rhode Island (n = 654), and Wisconsin (n = 1407). It is important to note that these sample sizes between states are not equal; however, our analytic plan does not entail investigating differences between states and, therefore, the recruitment strategy does not intend on obtaining a representative sample. Demographic variables were collected for descriptive purposes only. To ensure that the interview paradigm captures the most realistic situation of diagnostic decision-making, all participants were required to complete the study in one sitting.

Measures

Demographic Information. Participants were asked their age, gender, race, profession, years of experience, and whether they supervise other mental health professionals (See Appendix B for the comprehensive list of questions).

Diagnostic Interview. The interactive interview acted as a "create your own adventure" mechanism in which participants asked questions and received responses from a mock client. Participants were presented with a brief vignette explaining that they are meeting a client for an intake interview. They present with a downcast look, walking slowly and, when asked the reason for the session, they state, "I've been feeling really down recently." (See Appendix C). Participants then chose from two question options. The first question is focused on how long the client has been feeling down while the second question is concentrated on whether there have been any big changes in the client's life. Each of the two questions are consistent with separate diagnostic conclusions and, therefore, provided separate responses. For all questions after the first, participants were able to choose from a drop-down menu of multiple questions, each of which are more consistent with one diagnostic option or another. As the interview evolved, more question options became available, with some questions being dependent upon certain questions having already been asked. All questions were developed to be consistent with the diagnostic guidelines of Major Depressive Disorder or a normal grief response. Questions were designed to reflect specific guidelines as well as each differential diagnosis for both conditions. Questions and answers were peer-reviewed and revised to ensure that all required features were clearly presented and recognizable.

Rational-Experiential Inventory (REI). This measure is based on Cognitive-Experiential Self-Theory which asserts that individuals process information through two distinct, parallel systems of thinking that interact: rational (deliberate) and experiential (intuitive). The REI has been validated to identify thinking styles (rational vs. experiential) within various populations: adolescents (Shirzadifard et al., 2018), paramedics (Jensen et al., 2016), and nurses (Alba, 2018). The original version consists of 31 items which are evenly divided to make up two

distinct scales: Faith in Intuition (experiential; System 1 thinking) and Need for Cognition (rational; System 2 thinking; Epstein et al., 1996). The REI was modified by Pacini and Epstein (1999; See Appendix D) to expand the measurement of the two distinct thinking styles across two dimensions. Rational engagement and Rational ability formed the Rational thinking scale while Experiential engagement and Experiential ability formed the Experiential thinking scale. Each subscale is comprised of 10 items and are scored on a five-point Likert scale from "Definitely False" to "Definitely True." Higher subscale scores indicate a higher affinity for the specific thinking style. Previous studies have reported a reliability coefficient (Cronbach's α) of at least .86 (Pacini & Epstein, 1999; Witteman et al., 2009). The REI was used as a means of informing the likelihood that indications of System 1 and System 2 thinking in the interview corresponded to participants' self-reported thinking preferences.

Rank Order. Participants were presented a separate page where they were asked to rank order the usefulness of the diagnostic questions. The rank order was based on the perceived importance and usefulness in helping the participant to confidently arrive at a diagnosis. The page listed all of the diagnostic questions available to the participant during the interview, and they rank ordered them, starting with most important.

Clinical Relevance Questions. Participants were asked to provide qualitative responses to better understand their thought process throughout the study (See Appendix E). These questions included: a) What diagnostic features were most indicative for you to arrive at a diagnostic decision, b) What diagnostic features were absent which would have better supported the differential diagnosis, and c) What diagnostic materials (e.g., DSM, ICD) or other tools (e.g., SCID) guided your decision-making process when selecting diagnostic questions to ask?

Procedure

Participants received an email invitation to participate in the study through Qualtrics, a web-based survey program. Written, informed consent was provided to those interested participants. Upon entry into the study, participants were provided a prompt about a mock client arriving for their initial appointment. Participants were then presented with two sliding scales. For the first scale, major depressive disorder and a normal grief response were at opposite ends, and participants were asked to rate whether they are leaning towards one diagnosis or the other. They also had the option of staying neutral in the middle of the scale. In the second scale, participants rated their certainty of their diagnostic rating on the previous scale with a sliding scale of 0 to 100. Participants were then provided two potential question options, "Have any big changes happened in your life recently?" and "How long have you been feeling this way?" After selecting a question, they received a preprogramed response to the question. After each new prompt or piece of information, participants provided new ratings on the two sliding scales described previously, and then selected their next question. For the second question, participants were able to choose from a drop-down menu of multiple, different question options. The question-and-answer format followed by the two rating scales continued until one of three possibilities occurred: (a) the participant asked all possible questions and needed to provide a final diagnosis, (b) the participant provided a diagnostic rating of 0 or 100, in which case they were asked if they no longer wanted to ask any further questions and have arrived at a diagnostic conclusion, or (c) at any point during the sequence, the participant chose to no longer ask any questions and provided their final diagnosis. Following the conclusion of the interactive interview, participants completed the demographic questionnaire and the Rational-Experiential Inventory. Lastly, participants were presented with the clinical relevance questions and, separately, rank ordered their perceived importance of diagnostic questions.

Statistical Analysis Plan

SPSS 27.0 was used for all statistical analyses. Descriptive statistics (means, standard deviations, and frequencies) were calculated to verify that data meet the assumptions of the planned analyses. Being that this study was essentially exploratory, all statistical analyses were descriptive and idiographic in nature.

Aim 1. The first and primary aim of the proposed study was to investigate whether the interactive interview is an effective mechanism to operationalize the dual-process model. We will examine a variety of indicators both independently and in tandem to classify the decision-making process of participants. First, changes in a participant's diagnostic rating over time could indicate which system they are using. Specifically, no change in the rating or an increase in confidence in a diagnostic conclusion could indicate either System 1 or System 2; however, a decrease in confidence or a switch to the opposite diagnosis as being more likely would necessitate System 2 processing. Second, the consistency of individuals' diagnostic rating and question choice can indicate System 1 vs. System 2 processing. Specifically, if the person's choice of diagnostic question is consistent with their diagnostic rating, they could be operating in either System 1 or System 2. If they pick a diagnostic question that is inconsistent with their diagnostic rating (e.g., rating depression as more likely but picking a question that investigates the possibility of grief), it would indicate System 2. Third, the pattern of an individuals' response across trials could also indicate the predominant type of processing. For example, panel A of Figure 2 would indicate likely System 1 processing where an individual quickly reached a diagnostic conclusion with little variability. In contrast, panels B and C of Figure 2 would both represent patterns consistent with System 2 processing. The REI will be used as a validator to check whether the interactive interview can distinguish between System 1 and System 2

processing. Specifically, we expect that participants who self-report a preference for rational thinking will be more likely to engage in System 2 processing and participants who self-report a preference for experiential thinking will be less likely to engage in System 2 processing.

Aim 2. The second aim of the proposed study is to recognize and better understand patterns of diagnostic questioning and their relation to diagnostic certainty and confidence ratings. This will be done by reviewing the rank order of diagnostic questions, client responses, and evaluating the qualitative data provided by participants.

Results

Of the 6957 clinicians who were invited from various state licensee lists, 246 clicked the link to view the consent form and 57 consented. Of those who consented and went on to take part in the interactive interview, 30 (12.19%) provided complete data for inclusion in the present analysis. Participants that completed the study were younger (completed $M = 50.57$, $SD = 12.71$; not completed $M = 57.93$, $SD = 17.51$; $t(55) = -1.83$, $p < .05$) and with approximately seven years less experience, on average (completed $M = 22.03$, $SD = 10.45$; not completed $M = 29.11$, $SD = 15.58$; $t(55) = -2.03$, $p < .005$). The majority (73.3%) of participants were cisgender females. Almost all participants described themselves as Caucasian (93.3%). See ***Table 1*** for additional details regarding demographic and other participant features.

Table 1. Participant Demographics

	***f* (%)**
Gender	
Cisgender Female	22 (73.3)
Cisgender Male	7 (23.3)
Prefer not to answer	1 (3.3)
Race/Ethnicity	
Caucasian	28 (93.3)

East Asian/ Pacific Islander	1 (3.3)
Latinx	1 (3.3)
Highest Degree Earned	
PhD	20 (66.7)
PsyD	9 (30.0)
Masters	1 (3.3)
Provides Direct Supervision	
Yes	18 (60.0)
No	11 (36.7)
Prefer not to answer	1 (3.3)
Professional Setting*	
Outpatient	23 (76.7)
Inpatient (non-psychiatric)	2 (6.7)
Psychiatric Hospital	3 (10.0)
University Setting	3 (10.0)
Private Practice	10 (33.3)
Rehabilitation Facility	2 (6.7)
Nursing Home	1 (3.3)
Telehealth	11 (36.7)
Forensic	2 (6.7)
Foster Care	1 (3.3)
Typical Caseload per Week	
1-5	4 (13.3)
6-10	5 (16.7)
11-15	3 (10.0)
16-20	3 (10.0)
21-35	6 (20.0)
26-30	4 (13.3)
More than 30	3 (10.0)
Prefer not to answer	2 (6.7)
Age Range of Patients*	
11 years or under	9 (30.0)
12-17	11 (36.7)
18-24	15 (50.0)
25-39	18 (60.0)
40-60	15 (50.0)
Over 60	10 (33.3)

	M (SD)
Age	50.57 (12.71)

Years of Experience	22.03 (10.45)

*categories are not mutually exclusive

Over 70% of participants accurately diagnosed the mock patient with having a normal grief response. The number of questions participants asked before providing a final diagnosis ranged from 2 to 27 (M = 9.30, SD = 5.78). Of the thirty-six potential questions available, four questions were never chosen by participants. Surprisingly, that included question A15, which assessed for suicidality. See ***Table 2*** for additional details regarding the frequency of each diagnostic question asked by participants.

Table 2. Diagnostic Questions

	f
A1**:** How long have you been feeling this way?*	31
A2: Tell me a little more about what you mean when you say you feel down.*	24
A3: Have you been feeling down consistently during this time or has the feeling come and gone?	11
A4: How has your sleep been?	6
A5: How is your energy level during the day?	5
A6: Have you noticed or has anyone else said that you seem to be moving slower than usual?	1
A7: Any difficulties concentrating? Reading a book or following a movie?	4
A8: How have your symptoms impacted you at work?	2
A9: Does how you are feeling stop you from enjoying things you normally like doing?	6
A10: What do [did] you do with your husband?	0
A11: Have you stopped doing these activities?	1
A12: What has led you to stop doing these things?	0
A13: Have there been any changes in your appetite?	3
A14: Have you noticed any significant weight gain or weight loss?	3
A15: Have you had any thoughts about wanting to kill yourself [or wanting to die too]?	0
A16: How do you feel about the future?	8
A17: Do you have any feelings of guilt?	4
A18: Have you been drinking or using any other drugs during this time?	5
A19: Have you ever had a period of time where you felt much better than usual, like you were really excited or energetic?	3
A20: Have you been diagnosed with any medical conditions or are you taking medication for anything?	6

A21: Have you started developing beliefs that others may view as odd or seen/heard things that others were not able to see or hear?	3
B1: Have any big changes happened in your life recently?	28
B2: I'm so sorry to hear that. How did he pass away?	16
B3: May I ask what happened? Were you with him?	16
B4: Do you feel guilty about what happened?	7
B5: How have things been with your family and friends?	12
B6: How have you been feeling physically?	8
B7: Is it painful to think of memories of your husband?	4
B8: Have you done anything to avoid thinking of the accident or your husband because it's too difficult?	8
B9: How have you been coping with the loss? *	26
B10: Have you lost other important people in your life before?	6
B11: How often do you find yourself thinking about your husband?	7
B12: What are your beliefs about death and what happens when we die?	5
B13: Has this experience changed your outlook on life?	7
B14: Have you been able to do anything to honor your husband's life?	3
B15: Do you still speak to your husband? Or even think he is still here?	0

*Participant asked this question twice

Rank Order

Table 3 presents the results of the rank ordering of diagnostic questions with eighteen (60%) participants successfully engaging in the task by ranking at least one question. The question that was ranked first (i.e., most important) most often was A2 (n = 10; 33.33%) followed by A1 and B1 (each n = 3; 10.00%). These were also the questions with the highest average ranking. The most common patterns within the rank ordering were A1-B1 and A2-A1, each occurring across six participants (20.00%).

The consistency between the order in which participants asked questions during the interactive interview and how they ranked questions based on importance was examined to rule out if the rank ordering was a simple reflection of the order in which the questions were asked. By doing so, we can determine if participants ask their series of diagnostic questions based on importance or another metric. In addition, several questions were not available to be chosen to ask the mock patient because they were contingent on other questions being answered first. As a

result, it would be wrong to rely on the order of questions asked as an indicator of importance. A correlation near 1.00 would indicate that the rank order responses were identical to the order in which the questions were asked. Although eighteen participants completed the ranked order task, sixteen (53.33%) participants rank ordered a similar number of questions compared to how many they asked during the interview. Correlations between participants' rank order and the order they asked the questions during the interview ranged from -0.48 to 0.98 with an average correlation of 0.21.

Table 3. Rank Order of Diagnostic Questions

	Number of Times Ranked f (%)	***Average Rank***
A1**:** How long have you been feeling this way?	18 (100)	2.50
A2: Tell me a little more about what you mean when you say you feel down.	17 (94.44)	2.24
A3: Have you been feeling down consistently during this time or has the feeling come and gone?	8 (44.44)	6.00
A4: How has your sleep been?	8 (44.44)	7.25
A5: How is your energy level during the day?	4 (22.22)	10.50
A6: Have you noticed or has anyone else said that you seem to be moving slower than usual?	1 (5.56)	7.00
A7: Any difficulties concentrating? Reading a book or following a movie?	3 (16.67)	12.33
A8: How have your symptoms impacted you at work?	3 (16.67)	10.33
A9: Does how you are feeling stop you from enjoying things you normally like doing?	4 (22.22)	9.00
A10: What do [did] you do with your husband?	1 (5.56)	12.00
A11: Have you stopped doing these activities?	3 (16.67)	7.33
A12: What has led you to stop doing these things?	1 (5.56)	13.00
A13: Have there been any changes in your appetite?	2 (11.11)	12.00
A14: Have you noticed any significant weight gain or weight loss?	2 (11.11)	14.50
A15: Have you had any thoughts about wanting to kill yourself [or wanting to die too]?	3 (16.67)	18.67
A16: How do you feel about the future?	8 (44.44)	7.50
A17: Do you have any feelings of guilt?	3 (16.67)	11.00
A18: Have you been drinking or using any other drugs during this time?	4 (22.22)	14.25
A19: Have you ever had a period of time where you felt much better than usual, like you were really excited or energetic?	0 (0.0)	-

A20: Have you been diagnosed with any medical conditions or are you taking medication for anything?	4 (22.22)	16.50
A21: Have you started developing beliefs that others may view as odd or seen/heard things that others were not able to see or hear?	3 (16.67)	19.67
B1: Have any big changes happened in your life recently?	14 (77.78)	2.79
B2: I'm so sorry to hear that. How did he pass away?	9 (50.0)	7.22
B3: May I ask what happened? Were you with him?	7 (38.89)	8.14
B4: Do you feel guilty about what happened?	0 (0.0)	-
B5: How have things been with your family and friends?	5 (27.78))	6.20
B6: How have you been feeling physically?	5 (27.78)	6.67
B7: Is it painful to think of memories of your husband?	2 (11.11)	7.00
B8: Have you done anything to avoid thinking of the accident or your husband because it's too difficult?	3 (16.67)	11.00
B9: How have you been coping with the loss?	12 (66.67)	5.33
B10: Have you lost other important people in your life before?	4 (22.22)	11.25
B11: How often do you find yourself thinking about your husband?	5 (27.78)	9.80
B12: What are your beliefs about death and what happens when we die?	0 (0.0)	-
B13: Has this experience changed your outlook on life?	2 (11.11)	5.00
B14: Have you been able to do anything to honor your husband's life?	3 (16.67)	8.00
B15: Do you still speak to your husband? Or even think he is still here?	0 (0.0)	-

Ranked: Most Important Question	
A1: How long have you been feeling this way?	3 (16.67)
A2: Tell me a little more about what you mean when you say you feel down.	10 (55.56)
B1: Have any big changes happened in your life recently?	3 (16.67)
B2: I'm so sorry to hear that. How did he pass away?	1 (5.56)
A16: How do you feel about the future?	1 (5.56)

Common Trends	
A1- B1 How long have you been feeling this way? Have any big changes happened in your life recently?	6 (33.33)
A2- A1 Tell me a little more about what you mean when you say you feel down. How long have you been feeling this way?	6 (33.33)

Common Question Patterns

At the start of the interactive interview, participants are provided the option of asking one of two questions: A1, which is designed to serve as an MDD-related question and B1, which is designed to serve as a normal bereavement question. The overwhelming majority (80.00%; n = 24; $\chi^2(1) = 5.40$, $p = .02$) of participants chose to ask A1 ("How long have you been feeling this way?") as their initial question of the diagnostic interview, while the remaining six (20.00%) participants selected B1 ("Have any big changes happened in your life recently?") as their initial question. Of those participants who chose A1, fifteen followed up by asking another MDD-related question, A2 ("Tell me a little more about what you mean when you say you feel down"), and only eight switched to asking a normal bereavement question, B1.

A common pattern observed within each participant's selection of questions was a series of diagnosis specific questions before switching to the other diagnosis-specific questions. More specifically, there appeared to be several common clusters of questions asked in a distinct order. For example, the sequence of A1-A2-B1 occurred thirteen times (43.33%) among all participants. Additionally, B1-B2-B3 appeared ten times (33.33%). Notably, these questions would align with the diagnostic criteria for MDD. Specifically, they identify when symptoms began as well as if the mock patient is experiencing (a) depressed mood and/or (b) loss of interest or pleasure.

Prediction 1: The interactive interview will be an effective mechanism to operationalize the dual-process model.

Results indicated that by examining a variety of indicators both independently and in tandem, the interactive interview was able to describe the decision-making process of participants. Changes in a participant's diagnostic rating or certainty over time could indicate which system they are using. Specifically, no change in the rating or an increase in confidence in

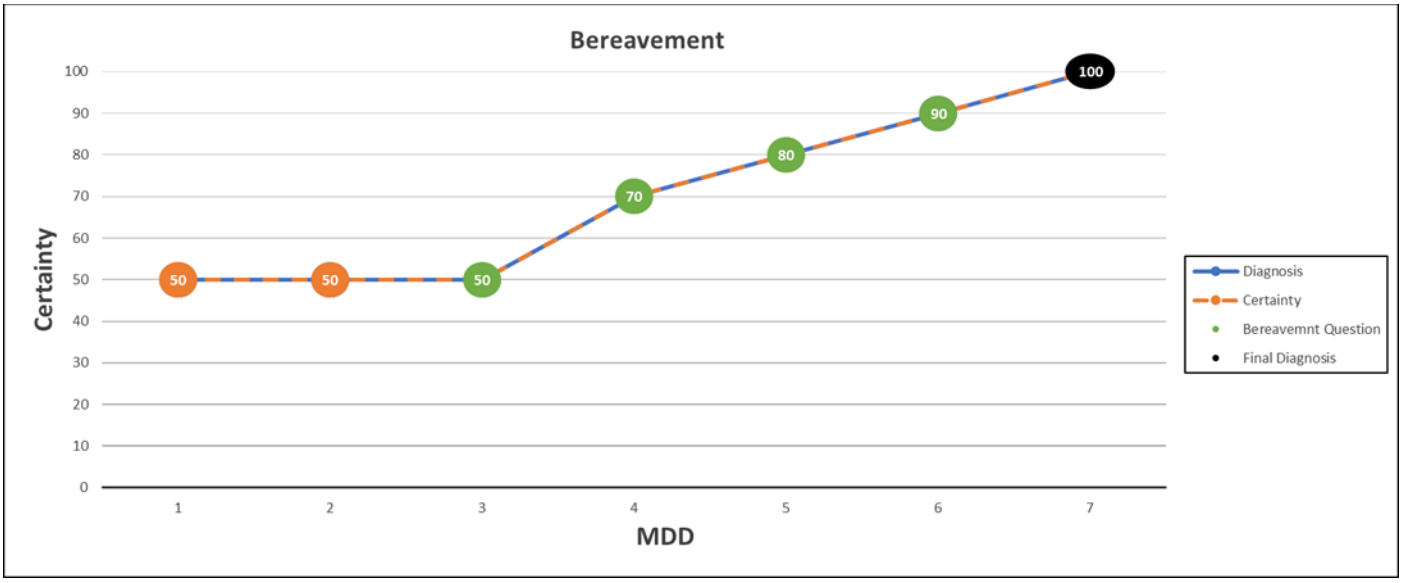
a diagnostic conclusion could indicate either System 1 or System 2; however, a decrease in confidence or a switch to the opposite diagnosis as being more likely would necessitate System 2 processing. The number of times participants indicated System 2 processing based on diagnostic questions and certainty rating each ranged from 0 to 9 (M = 3.10, SD = 2.41 & M = 3.07, SD = 2.53, respectively). The number of questions participants asked and switching diagnostic questions ($r(28) = .59$, $p < .001$) and switching their diagnostic certainty rating ($r(28) = .57$, $p < .001$) were significantly correlated. This suggests participants who ask more questions have more opportunities to engage in System 2, thereby, considering and asking questions related to differential diagnoses. See ***Table 4*** for additional information about metrics indicating the activation of System 2.

Table 4. Means and Standard Deviations of Metrics of Possible Processing Types

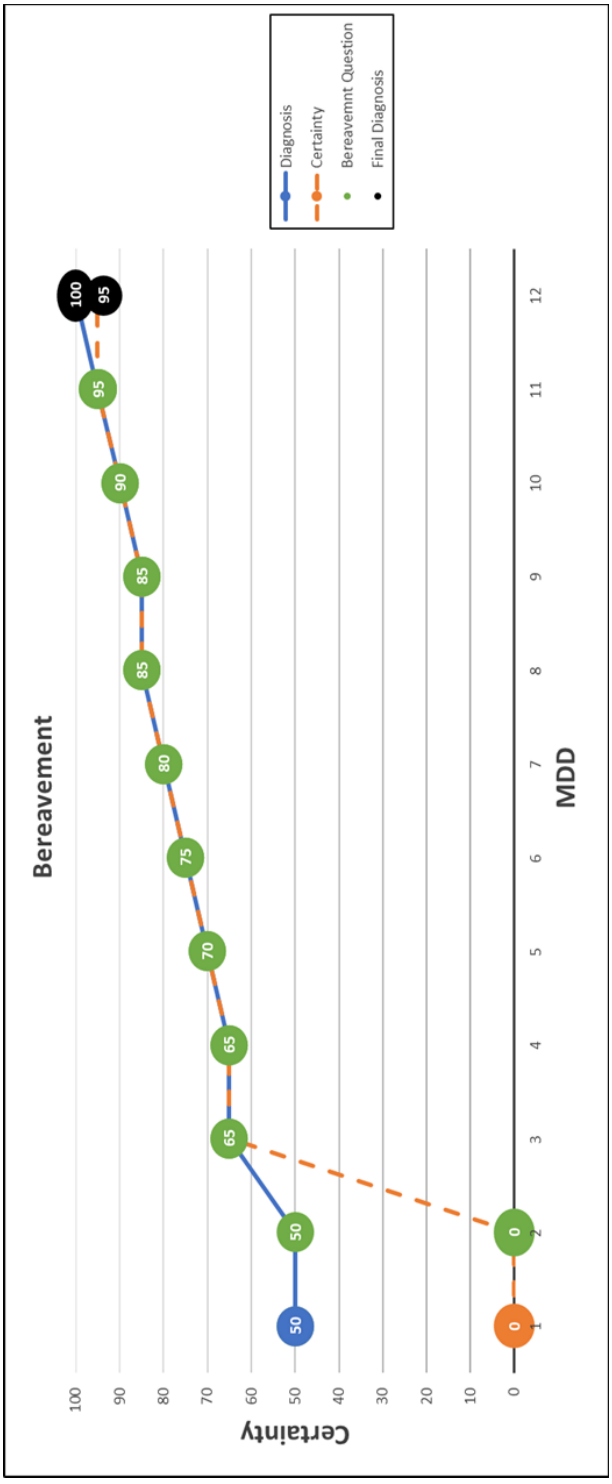
	Min	***Max***	***M***	***SD***
Questions Asked (total)	2	27	9.30	5.78
Switching By Dx Question	0	9	3.10	2.41
Switching By Certainty Rating	0	9	3.07	2.53
Longest Series Without Switching	1	13	4.93	3.00

Figures 3A and 3B are prime examples of likely System 1 processing. At the beginning of the interview, each participant was unsure of their diagnostic rating and remained neutral, until they switched to asking questions distinct to the opposite diagnosis. After switching from MDD questions to normal bereavement questions, their diagnostic rating and certainty of such rating progressively increased until they arrived at a final diagnosis. This pattern can also be observed in panel A of Figure 2 in which a diagnostic conclusion was quickly achieved with little to no variability. A total of thirteen (43.33%) participants exhibited these patterns (See Appendix F).

Figure 3. Examples of System 1 Processing



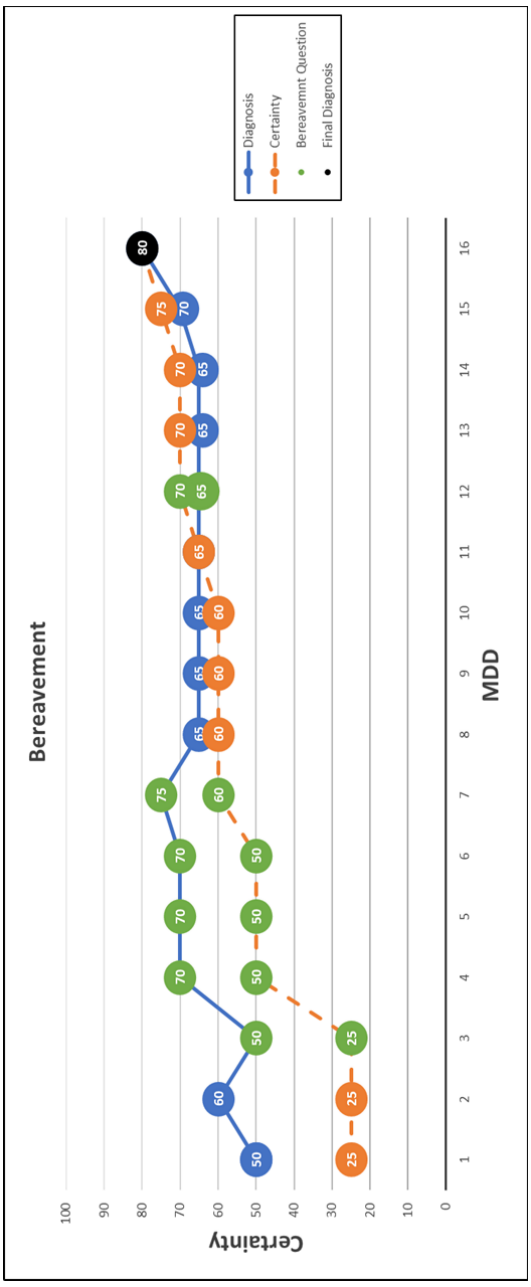
A.


Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
Bereavement
MDD
Certainty
100
95
95
90
85
85
80
75
70
65
65
50
50
0
0

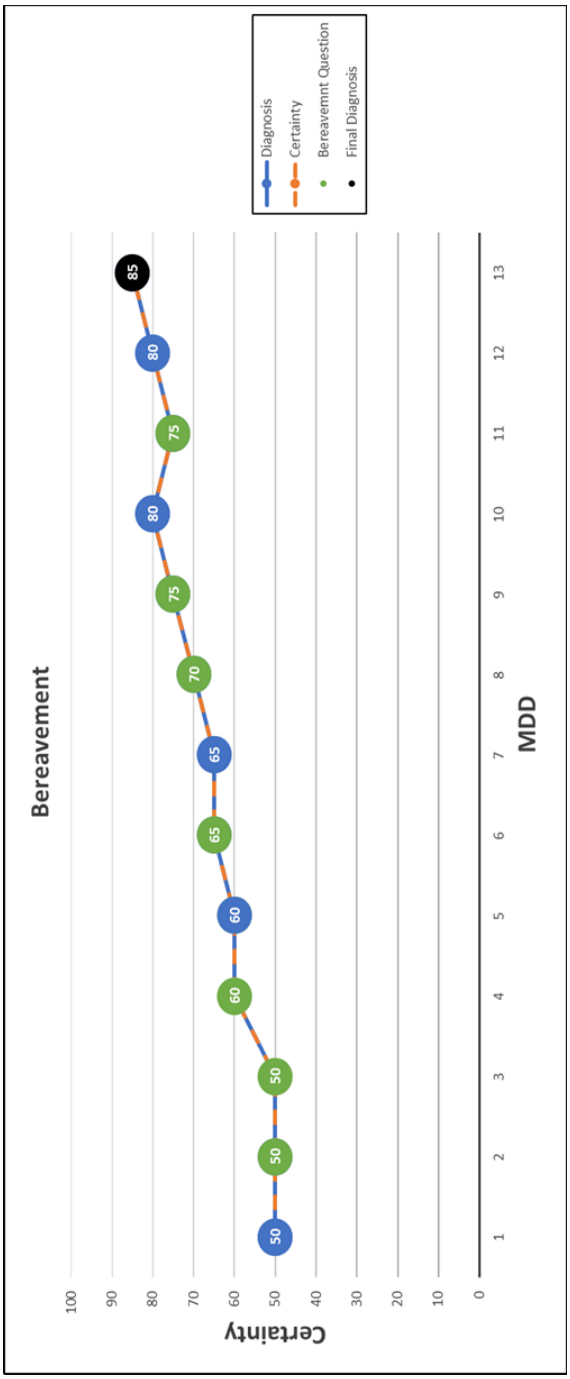
B.

Figures 4A and 4B are examples of the likely presence of both System 1 and System 2 processing. The consistency of individuals' diagnostic rating and question choice can indicate System 1 vs. System 2 processing. Specifically, if the person's choice of diagnostic question is consistent with their diagnostic rating, they could be operating in either System 1 or System 2. If they pick a diagnostic question that is inconsistent with their diagnostic rating (e.g., rating depression as more likely but picking a question that investigates the possibility of grief), it would indicate System 2. Figure 4A displays several instances of assumed System 1 processing, where the series of questions are diagnosis-consistent and their certainty rating remains stable or increases. This is observed in questions 1-2, 3-7, 8-11, and 13-15. System 2 processing is activated when there is a change in diagnostic questioning and, thereby, affecting the diagnostic rating. For example, between question 2-3 and 7-8 there is a ten point reduction in diagnostic rating when the participant switches their diagnostic questioning strategy. Conversely, in Figure 4B, there are less instances of likely System 1 processing (e.g., question 2-4) and a longer sequence of System 2 processing (e.g., question 4-8 and 9-13). This participant is consistently switching between MDD and normal bereavement specific questions. Although, the diagnostic rating and certainty rating are increasing throughout the interview, the participant asks bereavement specific questions to confirm their preliminary diagnosis; however, this person also uses the strategy of asking differential diagnosis questions (e.g., MDD) to further confirm their preliminary diagnosis. The act of considering differential diagnosis is an example of System 2 processing. Similar patterns can be observed in panels B and C of Figure 2, which both represent patterns consistent with System 2 processing. A total of seventeen (56.67%) participants displayed a similar pattern (See Appendix G for participants depicting these patterns).

Figure 4. Examples of System 2 Processing


Bereavement
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
Certainty
MDD

A.



B.

Rational-Experiential Inventory (REI)

Means and standard deviations from the REI scales and subscales are presented in ***Table 5***. Rational-total scores ranged from 21 to 80 (*M* = 39.40; *SD* = 11.25) while the two subscales,

Ability and Engagement ranged from 10 to 35 (M = 17.70; SD = 5.07) and from 10 to 40 (M = 19.77; SD = 6.15), respectively. Overall, the Experiential scale and subscale scores were higher. Experiential-total scores ranged from 34 to 68 (M = 53.80; SD = 9.44) and the two subscales, Ability and Engagement ranged from 12 to 35 (M = 26.13; SD = 5.62) and from 16 to 35 (M = 27.67; SD = 4.66), respectively. Overall, participants' thinking styles were significantly more experiential than rational ($t(29) = -5.77$, $p < .001$). Furthermore, participants scored significantly higher on the experiential subscales, Ability ($t(29) = -6.72$, $p < .001$) and Engagement ($t(29) = -5.74$, $p < .001$), compared to the Rational counterparts.

Table 5. Means and standard deviations of the REI scales and subscales

	Min	***Max***	***M***	***SD***
Rationality (total)	21	80	39.4	11.25
Rational ability	10	35	17.7	5.07
Rational engagement	10	40	19.77	6.15
Experientiality (total)	34	68	53.8	9.44
Experientiality ability	12	35	26.13	5.62
Experientiality engagement	16	36	27.67	4.66

Clinical Relevance Qualitative Data

The proposed study also aimed to recognize and better understand patterns of diagnostic questioning and their relation to diagnostic certainty and confidence ratings by reviewing the qualitative data provided by participants. The first open-ended question asked was, "What diagnostic features were most indicative for you to arrive at a diagnostic decision?" The two most popular features were, "problems did not begin prior to husband's death", and "brief period of symptoms." Each were noted by 13 (43.33%) participants. The second question focused on, "What diagnostic features were absent which would have better supported the differential diagnosis?" Several participants expressed additional information about the patient's history of

depression (*n* = 9, 30.00%) would have been helpful. Additionally, information about sleep, appetite, participation in enjoyable activities and suicidal ideation were noted as missing (although information about most of these symptoms was available in questions that participants may not have selected).

Notable responses related to participant diagnostic decision-making include: “I don't focus heavily on the ’right’ diagnosis. Rather, I focus on getting in the ball park and on severity” and “bereavement isn't a DSM V TR or ICD 10 diagnosis. May have considered an adjustment disorder with depressed mood (DSM-V) and ICD 10 grief code for dx.” Relatedly, when asked which sources participants rely on when making diagnostic decisions, twenty-five (83.33%) stated they use the DSM and only four (13.33%) acknowledging relying on their own clinical experience. See ***Table 6*** for additional details regarding open-ended responses provided by participants.

Table 6. Open-Ended Clinical Relevance Questions

	f (%)
What diagnostic features were most indicative for you to arrive at a diagnostic decision?	
Brief period of symptoms	13 (43.33)
Recent death of husband	13 (43.33)
Positive feeling/preserved self-esteem	5 (16.66)
What diagnostic features were absent which would have better supported the differential diagnosis?	
History of depression/family mental health	9 (30)
What diagnostic materials (e.g., DSM, ICD) or other tools (e.g., SCID) guided your decision-making process when selecting diagnostic questions to ask?	
DSM	25 (83.33)
Years of Experience	4 (13.33)
ICD	1 (3.33)
Functional Analysis	1 (3.33)
Do you agree that Major Depressive Disorder and normative grief exist on a spectrum?	

Yes	22 (73.33)
Unsure	2 (6.67)
No	1 (3.33)
Specify	
Both can exist at the same time but one does not include or exclude the other from existence	1 (3.33)
A time spectrum maybe. If grief is not overcome, depression steps in.	1 (3.33)

Correlations Between Clinician Demographics, Indicators & Interview Variables

Overall, the total number of questions asked was significantly correlated with indicators of System 2 based on diagnostic question ($r(28) = .59$, $p < .001$), switching certainty rating ($r(28) = .57$, $p < .001$), and longest series of questions asked without indicating System 2 ($r(28) = .81$, $p < .001$). Contrary to what was expected, results indicate that the Rational scale of the REI was not significantly correlated with the total number of questions asked ($r(28) = -.02$, $p = .93$), indicating System 2 based on diagnostic question ($r(28) = .001$, $p = .99$), or switching certainty rating ($r(28) = .09$, $p = .64$). In addition, the Experiential scale of the REI was not significantly correlated with the total number of questions asked ($r(28) = .20$, $p = .29$), indicating System 2 based on diagnostic question ($r(28) = .27$, $p = .15$), and switching certainty rating ($r(28) = .29$, $p = .12$).

The sample was split into two groups: those with limited System 2 activation and those with multiple System 2 activations. As expected, participants with multiple System 2 activations ($M = 3.94$, $SD = 2.66$) displayed more indicators of System 2 processing based on diagnostic question compared to those participants with limited System activation ($M = 2.00$, $SD = 1.53$), $t(28) = -2.35$, $p < .05$. Participants with multiple System 2 activations ($M = 4.12$, $SD = 2.71$) displayed more indicators of System 2 processing based on switching certainty rating compared to those participants with limited System activation ($M = 1.69$, $SD = 1.44$), $t(28) = -2.92$, $p < .05$. There were no other statistically significant results between these two groups with regard to

scores on the Rational scale of the REI (limited System 2 $M = 37.77$, $SD = 7.10$; multiple System 2 $M = 40.65$, $SD = 13.70$; $t(28) = -.69$, $p = .25$), and Experiential scale of the REI (limited System 2 $M = 53.23$, $SD = 7.87$; multiple System 2 $M = 54.24$, $SD = 10.72$; $t(28) = -.28$, $p = .39$). Additionally, there were no other statistically significant results between these two groups with regard to years of experience (limited System 2 $M = 20.77$, $SD = 8.77$; multiple System 2 $M = 23.00$, $SD = 11.75$; $t(28) = -.57$, $p = .29$), or providing supervision ($\chi^2(1) = .003$, $p = .96$).

There was a significant difference in the total number of questions asked between those participants who diagnosed Bereavement ($M = 8.14$, $SD = 3.81$) and those who diagnosed MDD ($M = 15.00$, $SD = 6.16$) as the final diagnosis, $t(24) = -3.02$, $p = .006$. There was also a significant difference in the length of questions asked without evidence of System 2 processing between those participants who diagnosed Bereavement ($M = 4.36$, $SD = 2.34$) and those who diagnosed MDD ($M = 7.75$, $SD = 3.86$) as the final diagnosis, $t(24) = -2.41$, $p = .012$.

Participants who were older and who had more years of experience were significantly more familiar with the diagnosis of prolonged grief disorder (PGD), $r(28) = .48$, $p = .008$ and $r(28) = .37$, $p = .045$, respectively. Participants who provide supervision had more years of experience ($M = 24.11$; $SD = 10.40$) than those who do not provide supervision ($M = 17.45$; $SD = 9.25$), $t(27) = -1.74$, $p = .05$.

Participants who reported providing supervision scored lower on the Rational scale of the REI (supervisors $M = 35.11$; $SD = 8.69$; non-supervisors $M = 46.27$, $SD = 12.37$; $t(27) = 2.86$, $p = .004$). There were no other statistically significant results between these two groups with regard to the interactive interview including the number of questions asked (supervisors $M = 9.00$, $SD = 4.06$; non-supervisors $M = 10.36$, $SD = 7.98$; $t(27) = .61$, $p = .55$), switching to System 2 based on diagnostic question (supervisors $M = 3.39$, $SD = 2.30$; non-

supervisors $M = 2.82$, $SD = 2.68$; $t(27) = -.61$, $p = .55$) or switching to System 2 based on certainty rating (supervisors $M = 3.22$, $SD = 2.29$; non-supervisors $M = 2.91$, $SD = 3.08$; $t(27) = -.31$, $p = .76$).

Discussion

This study aimed to investigate the effectiveness of an interactive interview in operationalizing the dual process model of decision-making using a variety of indicators. The interactive interview provided a variety of indicators regarding the decision-making process of participants by examining changes in a participant's diagnostic rating or certainty over time suggesting whether they were using System 1 (intuitive) or System 2 (analytic) processing. Participants did provide a variety of patterns of response to the interactive interview, with some providing no indication of overt System 2 thinking and others providing multiple indicators and an overall pattern of deliberative thinking.

Overall, the interactive interview mechanism appeared to be a successful methodology to operationalize the dual-process model of decision-making. This is a significant improvement from previous flawed methodologies including assigning arbitrary reaction time thresholds and instructional manipulations, which have been used to assess the presence of the two systems. The interactive interview identified individual decision patterns, moment to moment, as well as gathering a better comprehensive understanding of an individual's diagnostic decision-making process from start to finish. As a result, we were able to divide the sample into two groups; those with limited System 2 activation and those with multiple System 2 activations. While it is encouraging that the moment-to-moment decision patterns were significantly correlated with one's overall activation of System 2, there was no correlation with responses on the Rational and Experiential scales of the REI.

The study also identified common patterns and clusters of questions asked by participants during the interview. Thirty participants provided complete data for the analysis with over 70% accurately diagnosing the mock patient with a normal grief response. This is an encouraging finding, in particular because of the diagnostic challenge when working with bereaved patients. Mental health professionals are faced with the risk of pathologizing a natural grief response or neglecting to diagnose and treat a debilitating mental disorder. While the DSM-5 relied on mental health professionals to use their clinical judgment when making this decision, Prolonged Grief Disorder (PGD) in the ICD-11 (WHO, 2020) and new addition to the DSM-5 TR (American Psychiatric Association, 2022) provide clear guidelines to differentiate a normal grief response from PGD. The primary difference between the two diagnostic systems is the duration since the death of someone close to the bereaved. DSM-5 TR requires a period of at least one year since the death, while ICD requires at least six months since the death. Although participants in the current study were not provided the option to diagnose PGD, all participants selected either A1 or B1, which indicated it had been about two months since the death of the mock patient's husband. As a result, participants may have thought it was premature to pathologize the patient's grief by selecting MDD as the final diagnosis. In fact, about half of all participants stated that the "brief period of symptoms" and "recent death of husband" contributed to their final diagnosis. Additionally, one participant shared, "as a woman who lost a spouse to cancer, I can tell you 2 months is insufficient time for bereavement to pass." The majority of participants (73.33%) endorsed believing MDD and normative grief exist on a spectrum, which further suggests participants may have thought not enough time had passed since the death for the mock patient to be diagnosed with MDD.

Participants asked a range of two to twenty-seven questions before providing a final diagnosis. This indicates that some participants were thorough with their diagnostic questioning while others were quick to provide a diagnostic conclusion. Interestingly, there was a significant difference in the amount of questions asked depending on the final diagnosis. On average, participants who provided a final diagnosis of MDD asked about seven more questions compared to those who diagnosed a normal grief response. While it is concerning that over 20% of participants asked five or less questions before providing a diagnosis, it is encouraging that those who diagnosed MDD averaged asking fifteen questions. This indicates that participants were thorough with their diagnostic questioning prior to diagnosing the mock patient with a mental disorder. Although these participants asked more questions, if this were a patient in a real clinical setting, they would have been misdiagnosed. Unfortunately, this is a consistent problem within the mental healthcare system with rates of misdiagnosis of MDD as high as 65.9% (Vermani et al., 2011).

Of the thirty-six potential questions, there were only four which were never asked across all participants. Most surprisingly, the question assessing suicidality (A15) was never asked by any participant. However, three participants chose to rank the question in importance, averaging a rank of 16.67. A potential explanation for this diagnostic oversight may be participants view suicidality as a transdiagnostic symptom rather than one that is specific to MDD or its differential diagnosis. The task of all participants was to go through the interactive interview with the goal of providing a diagnostic conclusion. Assessing suicidality is an essential skill which should occur with each and every patient; however, it is primarily used to inform treatment rather than diagnosis. If participants approached the interactive interview paradigm focusing only on the goal of arriving at a diagnostic conclusion, without concern for developing

a treatment plan, thereby eliminating any risk, perhaps suicidality was no longer a clinical concern. Nevertheless, it is worth noting that suicidality was never assessed across any of the participants.

Overall, participants scored higher on the Experiential scale of the REI, indicating a self-reported preference for intuition during decision-making. This may be expected being that the sample is older (M = 50.57, SD = 12.71) with more years of experience (M = 22.03, SD = 10.45); however, this is contrary to the present available literature. Aarts et al. (2012) also used the REI and found that there was not a significant difference in thinking style (Rational or Experiential) between experienced psychologists (age range 39–63, M = 52.08, SD = 7.66; years of experience range 10-38, M = 20.64, SD = 8.23) and novice psychologists (age range 21–30, M = 23.76, SD = 2.53; less than 1 year of experience).

Similarly, it would be expected that mental health professionals who provide supervision would score higher on the Rational scale of the REI. Being that these participants work alongside trainees, it is likely they frequently discuss the presence of bias in clinical work and, therefore, expected that they may be more deliberate when making diagnostic decisions to ensure they are considering all differential diagnoses. Additionally, although the steps of diagnostic decision-making may be implicit for supervisors, they focus on making the steps of diagnostic decision-making explicit for their supervisees to ensure they are able to describe the steps they went through to arrive at a diagnostic conclusion (Zalzala & Gagen, 2023). Interestingly, in the current study, participants who provide supervision scored lower on the Rational scale of the REI and did not ask more questions; however, they did show slightly more indicators of System 2 thinking. This finding suggests that, despite scoring lower on the Rational scale, these participants were more intentional choosing the questions to ask. A potential explanation for this

discrepancy, scoring lower on the Rational scale of the REI but displaying more indicators of System 2 thinking, is that the REI taps into a more, general pattern of thinking that does not correlate strongly with individual moments in decision-making. Research suggests that attitude-behavior relationships are not entirely linear, as was previously believed (Bechler et al., 2021). Furthermore, asking less questions but being more intentional with choosing which questions to ask is likely an advantageous skill to develop in fast paced mental health settings such as inpatient hospitals, where, due to increased rates of admissions, mental health professionals are limited to how much time they are able to spend with each patient.

Limitations

The justification for this present study also serves as the primary limitation in that there is not a "gold standard" operationalization of the dual-process model. As outlined earlier, the ways in which the dual-process model has been operationalized in previous studies are inherently flawed. If the interactive interview mechanism was to be validated using these flawed indicators, it would lend to decreased confidence being that the foundation is weak. As a result, it is difficult to validate the interactive interview mechanism as an improved operationalization without better indicators.

Although the interactive interview paradigm displayed the ability to operationalize the dual-process model and identify common patterns in participants' selection of questions, the design may have limited the fullest ability of the paradigm. As stated earlier, one source of diagnostic unreliability is mental health professionals arriving at different diagnostic conclusions despite having the same information. Although the same information was available to all participants, they did not all receive the same information because they chose which diagnostic questions to ask and, therefore, the answers. In the present study, participants were asked to

arrive at a diagnostic conclusion between only two predetermined possibilities: MDD and normal grief response. It is possible that the similarities between these two diagnostic conclusions as well as the majority of participants endorsing their existence on a spectrum limited the range of uncertainty throughout the interactive interview.

Unfortunately, there are several limits to the generalizability of the findings. First, the study's sample size (n = 30) is small. The small sample may have limited the ability to identify common patterns in decision making and led to some variables not showing significant results. A potential explanation for this discrepancy is the study was conducted during the COVID-19 pandemic. During this time, mental health professionals have experienced a dramatic increase in stressors caused by the growing need for mental health services and changing work environments (Billings et al., 2021). As a result, these factors likely lowered their willingness to participate in this study. Additionally, it is important to consider mental health professionals' hesitancy to take part in the study was related to the "Sentinel effect", in which they may change their behavior because they are being evaluating (Veloski et al., 2005). In this case, rather than changing how they would have responded in the study, they may have chosen to not engage at all.

The study's reliance on a mock patient scenario may also limit the generalizability of the results to real-world clinical settings. This is in part because the interactive interview is focused only on a portion of a complex diagnostic decision-making process with multiple components. The study is aimed to investigate how people go about collecting the available information (choosing the questions to ask) as an indication of their thought process and putting the information together to form a diagnostic conclusion. However, information gathering also involves behavioral indicators like the patient's affect, psychomotor responses, speech patterns,

etc. Both the cognitive decision-making component and the information gathering process blend into the overall process of pattern recognition. In this study, we were forced to separate the two and control all inputs so we may focus exclusively on the cognitive component. Nevertheless, vignette methodology has been shown to be a valid analog for decision-making in clinical settings (Evans et al., 2015). In fact, assessing the diagnostic decision-making process in this way is more ecologically valid compared to written case vignettes. In addition, using the interactive interview methodology may be increasingly more relevant considering the exponential rise in telehealth (Shaver, 2022) and development of virtual reality therapy interventions (Sampaio et al., 2021). Delivering therapy virtually limits the number of behavioral indicators being provided and, therefore, that can be used to assist in diagnostic decision-making. As a result, the cognitive component of how mental health professionals go about collecting the available information becomes increasingly more important.

Future Studies

The interactive interview mechanism should continue to be validated to maximize confidence that it is an effective operationalization of the dual-process model of decision-making. The methodology should also be improved to become more sophisticated. As stated earlier, a limitation of the present study is that behavioral indicators like the patient's affect, psychomotor responses, and speech patterns were excluded. Rather than only providing participants with pre-programmed text responses, future studies can provide pre-programmed video responses which would allow participants to also consider behavioral indicators. This would also more closely mimic telehealth services. Future versions of the interactive interview should test mouse tracking as another validator. Mathur and Reichling (2019) designed a user-friendly, open-source software that is designed to easily embed into surveys using the Qualtrics

platform to allow researchers the ability to monitor and measure the movement of participants' mouse cursors. Gathering this data will provide an additional validation tool as to whether participants are operating in System 1 vs. System 2. Specifically, this will be a useful tool when participants are deciding on the drop-down list of questions to ask, both rating scales and rank ordering of questions.

Future iterations of the interactive interview could also allow the participant to choose which diagnoses they are considering. This would better mimic the diagnostic process and provide better insight of differences between mental health professionals. For example, what factors would lead to half of the sample contemplating between Major Depressive Disorder and Posttraumatic Stress Disorder, and the other half between Major Depressive Disorder and Generalized Anxiety Disorder? An additional indicator of System 2 processing with this improved methodology is if a participant changes the diagnoses they are considering at any point throughout the interview. This would indicate the participant is actively considering differential diagnoses, which requires System 2 thinking.

Future studies can also attempt to better understand the intersection between psychiatric diagnostic decision-making and neuroscience by having participants complete the interactive interview while undergoing an fMRI (van den Berg et al., 2020). For example, episodic memory is believed to be activated during System 2 processing and so, fMRI would allow researchers to determine activation of the hippocampus, the region of the brain involved in encoding and retrieval of episodic memory (Dickerson & Eichenbaum, 2009).

One of the major implications of the present methodology is the ability for the interactive interview to answer other important questions related to diagnosis. For example, does the

identity of the mock patient affect the series of diagnostic questions or the final diagnosis? Eubanks-Carter and Goldfried (2006) investigated whether the client's sexual orientation affected the probability that therapists would diagnose them with borderline personality disorder (BPD) and found a significant interaction between client gender and sexual orientation regarding the diagnosis of BPD. The interactive interview would allow for the identity of the mock patient to be manipulated to investigate if the series of diagnostic questions or the final diagnosis varies based on the patient's gender, age, or sexual orientation. Conversely, does the identity of the mental health professional affect diagnostic decisions?

Implications

The interactive interview mechanism is a novel approach to better understand diagnostic decision-making. Vignette study designs have been shown to be highly generalizable to real world behaviors among clinicians (Evans et al., 2015). The present design is an improvement to vignette methodology by mimicking the information gathering portion of diagnostic interviewing rather than the limitation of only receiving a historical summary of current and past symptom.

Findings display important clinical and training implications for clinicians. The interactive interview mechanism was successful in capturing moment to moment diagnostic decisions, including the type of questions asked and patterns of both System 1 and System 2 processing. This will enable clinicians to be aware of their decision-making process and the potential biases that may influence their judgments (Lilienfeld & Lynn, 2014). For example, multiple participants engaged in confirmation bias by only asking questions that affirmed their preliminary diagnosis. Research suggests that when mental health professionals engage in confirmatory information seeking, they are less likely to provide an accurate diagnosis,

compared to those who engage in “balanced” information seeking by considering differential diagnoses (Mendel et al., 2011). Other participants engaged in premature closure by only choosing to ask two or three questions before providing a diagnosis. Similarly, these participants are at an increased risk of providing an inaccurate diagnosis by not asking enough questions to confirm their preliminary diagnosis as well as rule out all potential differential diagnoses. Clinicians could benefit from training programs that help them recognize these biases and improve their clinical decision-making skills. By identifying and addressing biases and heuristics used by clinicians, individualized training and consultation can be provided to improve their decision-making process. This approach has the potential to enhance the quality of care provided to patients, leading to better health outcomes. Furthermore, if adopted by doctoral-level graduate programs, the proposed methodology could be used as an assessment and training tool to improve critical thinking and decision-making skills among future clinicians. This would help to address the problem of diagnostic inaccuracy, which has been shown to be related to differences in the way clinicians think, rather than deficits in clinical knowledge. Overall, the proposed paradigm has the potential to enhance the quality of care provided to patients and improve the skills of mental health professionals, making it a valuable tool in the field of public health.

Conclusion

The present study successfully piloted the ability of an interactive interview to operationalize the dual-process model of decision-making. Results identified System 2 activation during the diagnostic interview with a mock patient and revealed the importance and ordering of diagnostic questions. Findings provide insight on diagnostic decision-making on the individual level. Future studies should focus on further validation of the interactive interview mechanism and potential incorporation in training programs.

References

Aarts, A.A., Witteman, C. L. M., Souren, P. M., & Egger, J. I. M. (2012). Associations between psychologists' thinking styles and accuracy on a diagnostic classification task. *Synthese (Dordrecht), 189*(Suppl 1), 119–130. https://doi.org/10.1007/s11229-012-0081-3

Akinhanmi, M. O., Biernacka, J. M., Strakowski, S. M., McElroy, S. L., Balls Berry, J. E., Merikangas, K. R., … Frye, M. A. (2018). Racial disparities in bipolar disorder treatment and research: A call to action. *Bipolar Disorders*, *20*(6), 506–514. https://doi.org/10.1111/bdi.12638

Alba, B. (2018). Factors that impact on emergency nurses' ethical decision-making ability. *Nursing Ethics*, *25*(7), 855–866. https://doi.org/10.1177/0969733016674769

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.).

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.).

American Psychiatric Association. (2013a). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

American Psychiatric Association. (2013b). *Depressive disorders. In diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596.dsm05

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders : DSM-5-TR* (5th edition, text revision.).

Anderson, C. J. (2010). Central limit theorem. In *The Corsini Encyclopedia of Psychology and Behavioral Science* (pp. 1–2). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470479216.corpsy0160

Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90–109. https://doi.org/10.1016/j.cognition.2016.10.014

Bailey, R. K., Blackmon, H. L., & Stevens, F. L. (2009). Major depressive disorder in the African American population: Meeting the challenges of stigma, misdiagnosis, and treatment disparities. *Journal of the National Medical Association*, *101*(11), 1084–1089. https://doi.org/10.1016/S0027-9684(15)31102-0

Bechler, C.J., Tormala, Z. L., & Rucker, D. D. (2021). The attitude–behavior relationship revisited. *Psychological Science*, *32*(8), 1285–1297. https://doi.org/10.1177/0956797621995206

Berk, M., Dodd, S., Callaly, P., Berk, L., Fitzgerald, P., de Castella, A. ., … Kulkarni, J. (2007). History of illness prior to a diagnosis of bipolar disorder or schizoaffective disorder. *Journal of Affective Disorders*, *103*(1), 181–186. https://doi.org/10.1016/j.jad.2007.01.027

Bessonova, L., Ogden, K., Doane, M. J., O'Sullivan, A. K., & Tohen, M. (2020). The economic burden of bipolar disorder in the United States: A systematic literature review. *Clinico Economics and outcomes research: CEOR*, *12*, 481–497. https://doi.org/10.2147/CEOR.S259338

Bieleke, M., Gollwitzer, P. M., Oettingen, G., & Fischbacher, U. (2017). Social value orientation moderates the effects of intuition versus reflection on responses to unfair ultimatum offers: Processing unfair ultimatum offers. *Journal of Behavioral Decision Making*, *30*(2), 569–581. https://doi.org/10.1002/bdm.1975

Billings, J., Biggs, C., Ching, B. C. F., Gkofa, V., Singleton, D., Bloomfield, M., & Greene, T. (2021). Experiences of mental health professionals supporting front-line health and social care workers during COVID-19: qualitative study. *BJPsych Open*, *7*(2), e70–e70. https://doi.org/10.1192/bjo.2021.29

Bowden, C. L. (2005). A different depression: Clinical distinctions between bipolar and unipolar depression. *Journal of Affective Disorders, 84*(2–3), 117–125. https://doi.org/10.1016/S0165-0327(03)00194-0

Cantor, N., Smith, E. E., French, R. D., & Mezzich, J. (1980). Psychiatric diagnosis as prototype categorization. *Journal of Abnormal Psychology*, 89(2), 181-193. https://doi.org/10 .1037/0021-843X.89.2.181

Cloutier, M., Aigbogun, M. S., Guerin, A., Nitulescu, R., Ramanakumar, A. V., Kamat, S. A., … Wu, E. (2016). The economic burden of schizophrenia in the United States in 2013. *The Journal of Clinical Psychiatry*, *77*(6), 764–771. https://doi.org/10.4088/JCP.15m10278

Coulter, C., Baker, K. K., & Margolis, R. L. (2019). Specialized consultation for suspected recent-onset schizophrenia: Diagnostic clarity and the distorting impact of anxiety and reported auditory hallucinations. *Journal of Psychiatric Practice*, *25*(2), 76–81. https://doi.org/10.1097/PRA.0000000000000363

Croskerry, P. (2002). Achieving quality in clinical decision making: Cognitive strategies and detection of bias. *Academic Emergency Medicine*, *9*(11), 1184–1204. https://doi.org/10.1197/aemj.9.11.1184

Croskerry, P. (2008). Cognitive and affective dispositions to respond. In P. Croskerry, K. Cosby, S. Schenkel, & R. Wears (Eds.), *Patient safety in emergency medicine* (pp. 219–227). Philadelphia: Lippincott, Williams & Wilkins.

Croskerry, P. (2009). A universal model of diagnostic reasoning. *Academic Medicine*, *84*(8), 1022–1028. https://doi.org/10.1097/ACM.0b013e3181ace703

Croskerry, P. (2013). From mindless to mindful practice—Cognitive bias and clinical decision making. *New England Journal of Medicine, 368*, 2445-2448

Croskerry, P. (2017). A model for clinical decision-making in medicine. *Medical Science Educator*, *27*(S1), 9–13. https://doi.org/10.1007/s40670-017-0499-9

Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 1: Origins of bias and theory of debiasing. *BMJ Quality & Safety*, *22*(Suppl 2), ii58–ii64. https://doi.org/10.1136/bmjqs-2012-001712

De Neys, W. (2021). On dual- and single-process models of thinking. *Perspectives on Psychological Science*, Advance online publication. https://doi.org/10.1177/1745691620964172

De Neys, W., & Pennycook, G. (2019). Logic, fast and slow: Advances in dual-process theorizing. *Current Directions in Psychological Science*, *28*(5), 503–509. https://doi.org/10.1177/0963721419855658

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, 313, 684–687.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5–18.

Dickerson, B.C., & Eichenbaum, H. (2010). The episodic memory system: Neurocircuitry and disorders. *Neuropsychopharmacology (New York, N.Y.)*, *35*(1), 86–104. https://doi.org/10.1038/npp.2009.126

Diederich, A., & Trueblood, J. S. (2018). A dynamic dual process model of risky decision making. *Psychological Review*, *125*(2), 270–292. https://doi.org/10.1037/rev0000087

Djulbegovic, B., Elqayam, S., & Dale, W. (2018). Rational decision making in medicine: Implications for overuse and underuse. *Journal of Evaluation in Clinical Practice*, *24*(3), 655–665. https://doi.org/10.1111/jep.12851

Elstein, A. S. (1994). What goes around comes around: The return of the hypothetico- deductive strategy. *Teaching and Learning in Medicine, 6*, 121–123.

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology*, *71*(2), 390–405. https://doi.org/10.1037/0022-3514.71.2.390

Eubanks-Carter, C., & Goldfried, M. R. (2006). The impact of client sexual orientation and gender on clinical judgments and diagnosis of borderline personality disorder. *Journal of Clinical Psychology*, *62*(6), 751–770. https://doi.org/10.1002/jclp.20265

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*(1), 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Evans, S. C., Roberts, M. C., Keeley, J. W., Blossom, J. B., Amaro, C. M., Garcia, A. M., Stough, C. O., Canter, K. S., Robles, R., & Reed, G. M. (2015). Vignette methodologies for studying clinicians' decision-making: Validity, utility, and application in ICD-11 field studies. *International Journal of Clinical and Health Psychology*, *15*(2), 160–170. https://doi.org/10.1016/j.ijchp.2014.12.001

Fernando, I., Cohen, M., & Henskens, F. (2012). Pattern-based formulation: A methodology for psychiatric case formulation. *Australasian Psychiatry*, *20*(2), 121–126. https://doi.org/10.1177/1039856212437255

First, M. B. (2014). *DSM-5 Handbook of Differential Diagnosis*. Washington, DC: American Psychiatric Publishing.

Gaebel, W., Stricker, J., Riesbeck, M., Zielasek, J., Kerst, A., Meisenzahl-Lechner, E., … Falkai, P. (2020). Accuracy of diagnostic classification and clinical utility assessment of ICD-11 compared to ICD-10 in 10 mental disorders: Findings from a web-based field study. *European Archives of Psychiatry and Clinical Neuroscience*, *270*(3), 281–289. https://doi.org/10.1007/s00406-019-01076-z

Gara, M. A., Minsky, S., Silverstein, S. M., Miskimen, T., & Strakowski, S. M. (2019). A naturalistic study of racial disparities in diagnoses at an outpatient behavioral health clinic. *Psychiatric Services*, *70*(2), 130–134. https://doi.org/10.1176/appi.ps.201800223

Garb, H. N. (2005). Clinical judgment and decision making. *Annual Review of Clinical Psychology*, *1*(1), 67–89. https://doi.org/10.1146/annurev.clinpsy.1.102803.143810

Genero, N., & Cantor, N. (1987). Exemplar prototypes and clinical diagnosis: Toward a cognitive economy. *Journal of Social and Clinical Psychology*, *5*(1), 59–78. https://doi.org/10.1521/jscp.1987.5.1.59

Ghaemi, S. N., Boiman, E.E., & Goodwin, F.K. (2000). Diagnosing bipolar disorder and the effect of antidepressants: A naturalistic study. *Journal of Clinical Psychiatry*, 61(10), 804-808.

Glazier, K., Swing, M., & McGinn, L. K. (2015). Half of obsessive-compulsive disorder cases misdiagnosed: Vignette-based survey of primary care physicians. *The Journal of Clinical Psychiatry*, *76*(06), e761–e767. https://doi.org/10.4088/JCP.14m09110

Goldberg, J. F., & Truman, C. J. (2003). Antidepressant-induced mania: An overview of current controversies. *Bipolar Disorders*, *5*(6), 407–420. https://doi.org/10.1046/j.1399-5618.2003.00067.x

Goulart, M., Da Costa Jr, N., Santos, A., Takase, E., & Da Silva, S. (2013). Psychophysiological correlates of the disposition effect. *PLoS ONE*, 8(1), 1–4.

Guo, L., Trueblood, J.S., & Diederich, A. (2017) Thinking fast increases framing effects in risky decision making. *Psychological Science. 28*(4), 530-543. https://doi.org/10.1177/0956797616689092

Higgs J., & Jones M. A. (2008). Clinical decision making and multiple problem spaces (pp. 3-18). In J. Higgs, M. A. Jones, S. Loftus & N. Christensen (Eds.), *Clinical reasoning in the health professions* (3rd ed., pp. 3-18). Amsterdam: Butterworth-Heinemann.

Hirschfeld, R. M. A., Lewis, L., & Vornik, L. A. (2003). Perceptions and impact of bipolar disorder: How far have we really come? Results of the national depressive and manic- depressive association 2000 survey of individuals with bipolar disorder. *The Journal of Clinical Psychiatry*, *64*(2), 161–174. https://doi.org/10.4088/JCP.v64n0209

Jensen, J. L., Bienkowski, A., Travers, A.H., Calder, L.A., Walker, M., Tavares, W., & Croskerry, P. (2016). A survey to determine decision-making styles of working paramedics and student paramedics. *Canadian Journal of Emergency Medicine*, *18*(3), 213–222. https://doi.org/10.1017/cem.2015.95

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-291.

Kassirer, J. P., & Kopelman, R. I. (1991). *Learning clinical reasoning*. Baltimore: Williams & Wilkins.

Kim, N. S., & Ahn, W. (2002). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology: General*, *131*(4), 451–476. https://doi.org/10.1037/0096-3445.131.4.451

Krishan Aggarwal, N., Rosenheck, R. A., Woods, S. W., & Sernyak, M. J. (2012). Race and long-acting antipsychotic prescription at a community mental health center: A retrospective chart review. *The Journal of Clinical Psychiatry*, *73*(4), 513–517. https://doi.org/10.4088/JCP.11m07161

Lilienfeld, S. O., & Lynn, S. J. (2014). Errors/biases in clinical decision making. In *The Encyclopedia of Clinical Psychology* (pp. 1–9). https://doi.org/10.1002/9781118625392.wbecp567

Jones, K. D. (2010). The unstructured clinical interview. *Journal of Counseling & Development*, *88*(2), 220–226. https://doi.org/10.1002/j.1556-6678.2010.tb00013.x

Maj, M. (2012). Bereavement-related depression in the DSM-5 and ICD-11. *World Psychiatry*, *11*(1), 1–2. https://doi.org/10.1016/j.wpsyc.2012.01.001

Markovits, H., Brunet, M.-L., Thompson, V., & Brisson, J. (2013). Direct evidence for a dual process model of deductive inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(4), 1213–1222. https://doi.org/10.1037/a0030906

Mathur, M. B., & Reichling, D.B. (2019). Open-source software for mouse-tracking in qualtrics to measure category competition. *Behavior Research Methods.* 51, 1987-1997. https://doi.org/10.3758/s13428-019-01258-6

Mendel, R., Traut-Mattausch, E., Jonas, E., Leucht, S., Kane, J. M., Maino, K., Kissling, W., & Hamann, J. (2011). Confirmation bias: why psychiatrists stick to wrong preliminary

diagnoses. *Psychological Medicine*, *41*(12), 2651–2659. https://doi.org/10.1017/S0033291711000808

Miller, C. (2019). Interviewing strategies, rapport, and empathy. In D. L. Segal (Ed.), *Diagnostic Interviewing* (pp. 29–53). Springer US. https://doi.org/10.1007/978-1-4939-9127-3_2

Mitchell, A. J., Vaze, A., & Rao, S. (2009). Clinical diagnosis of depression in primary care: A meta-analysis. *The Lancet*, *374*(9690), 609–619. https://doi.org/10.1016/S01406736(09)60879-5

Mojtabai, R. (2013). Clinician-identified depression in community settings: Concordance with structured-interview diagnoses. *Psychotherapy and Psychosomatics*, *82*(3), 161–169. https://doi.org/10.1159/000345968

Narrow, W. E., Clarke, D. E., Kuramoto, S. J., Kraemer, H. C., Kupfer, D. J., Greiner, L., & Regier, D. A. (2013). DSM-5 field trials in the United States and Canada, Part III: Development and reliability testing of a cross-cutting symptom assessment for DSM-5. *The American Journal of Psychiatry*, *170*(1), 71–82. https://doi.org/10.1176/appi.ajp.2012.12071000

Norman, G. (2005). Research in clinical reasoning: Past history and current trends. *Medical Education*, *39*(4), 418–427. https://doi.org/10.1111/j.1365-2929.2005.02127.x

Norman, G., Sherbino, J., Dore, K., Wood, T., Young, M., Gaissmaier, W., Kreuger, S., & Monteiro, S. (2014). The etiology of diagnostic errors: A controlled trial of system 1 versus system 2 reasoning. *Academic Medicine*, *89*(2), 277–284. https://doi.org/10.1097/ACM.0000000000000105

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology, 76*(6), 972–987. https://doi.org/10.1037/0022-3514.76.6.972

Pelaccia, T., Tardif, J., Triby, E., & Charlin, B. (2011). An analysis of clinical reasoning through a recent and comprehensive approach: The dual-process theory. *Medical Education Online*, *16*(1), 5890–5899. https://doi.org/10.3402/meo.v16i0.5890

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, *80*, 34–72. https://doi.org/10.1016/j.cogpsych.2015.05.001

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, *3*(3), 382– 407. https://doi.org/10.1016/0010-0285(72)90014-X

Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). DSM-5 field trials in the United States and Canada, part II: Test- retest reliability of selected categorical diagnoses. *American Journal of Psychiatry*, *170*(1), 59–70. https://doi.org/10.1176/appi.ajp.2012.12070999

Reuter, T. R., Sharp, C., Kalpakci, A. H., Choi, H. J., & Temple, J. R. (2016). Sexual orientation and borderline personality disorder features in a community sample of adolescents. *Journal of Personality Disorders*, *30*(5), 694–707. https://doi.org/10.1521/pedi_2015_29_224

Richards, D.W., Fenton-O'Creevy, M., Rutterford, J., & Kodwani, D. G. (2018). Is the disposition effect related to investors' reliance on system 1 and system 2 processes or their strategy of emotion regulation? *Journal of Economic Psychology*, *66*, 79–92. https://doi.org/10.1016/j.joep.2018.01.003

Rodriguez-Seijas, C., Morgan, T. A., & Zimmerman, M. (2021). Is there a bias in the diagnosis of borderline personality disorder among lesbian, gay, and bisexual patients? *Assessment*, *28*(3), 724–738. https://doi.org/10.1177/1073191120961833

Rogers, R. (2001). *Handbook of diagnostic and structured interviewing*. New York, NY: Guilford Press.

Sampaio, M., Haro, M. V. N., De Sousa, B., Melo, W. V., & Hoffman, H. G. (2021). Therapists make the switch to telepsychology to safely continue treating their patients during the COVID-19 pandemic. Virtual reality telepsychology may be next. *Frontiers in Virtual Reality*, *1*. https://doi.org/10.3389/frvir.2020.576421

Schmidt, H., Norman, G., & Boshuizen, H. (1990). A cognitive perspective on medical expertise: Theory and implications. *Academic Medicine*, *65*(10), 611–621.

Schneider, W., & Shiffrin, R.M. (1977). Controlled and automatic human information processing: Detection, search, and attention. *Psychological Review*, *84*(1), 1-66.

Segal, D.L., & Hersen, M. (Eds.). (2010). Diagnostic interviewing (4th ed.). New York: Springer.

Shaver J. (2022). The state of telehealth before and after the COVID-19 pandemic. *Primary care*, *49*(4), 517–530. https://doi.org/10.1016/j.pop.2022.04.002

Shefrin, H. & Statman, M. (1985), The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance, 40*: 777-790. https://doi.org/10.1111/j.1540-6261.1985.tb05002

Shirzadifard, M., Shahghasemi, E., Hejazi, E., Naghsh, Z., & Ranjbar, G. (2018). Psychometric properties of rational-experiential inventory for adolescents. *SAGE Open*, *8*(1), 215824401876721. https://doi.org/10.1177/2158244018767219

Stark, M., & Fins, J. J. (2014). The ethical imperative to think about thinking - diagnostics, metacognition, and medical professionalism. *Cambridge Quarterly of Healthcare Ethics*, *23*(4), 386–396. https://doi.org/10.1017/S0963180114000061

Sung, S. C., Dryman, M. T., Marks, E., Shear, M. K., Ghesquiere, A., Fava, M., & Simon, N. M. (2011). Complicated grief among individuals with major depression: Prevalence, comorbidity, and associated features. *Journal of Affective Disorders*, *134*(1–3), 453–458. https://doi.org/10.1016/j.jad.2011.05.017

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3), 107–140. https://doi.org/10.1016/j.cogpsych.2011.06.001

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323.

van den Berg, B., de Bruin, A. B. H., Marsman, J.-B. C., Lorist, M. M., Schmidt, H. G., Aleman, A., & Snoek, J. W. (2020). Thinking fast or slow? Functional magnetic resonance imaging reveals stronger connectivity when experienced neurologists diagnose ambiguous cases. *Brain Communications*, *2*(1), fcaa023–fcaa023. https://doi.org/10.1093/braincomms/fcaa023

Veloski, J., Tai, S., Evans, A.S., & Nash, D.B. (2005) Clinical vignette-based surveys: A tool for assessing physician practice variation. *American Journal of Medical Quality*. 20(3):151-157. Doi:10.1177/1062860605274520

Vermani, M., Marcus, M., & Katzman, M. A. (2011). Rates of detection of mood and anxiety disorders in primary care: A descriptive, cross-sectional study. *The Primary Care Companion for CNS Disorders*, *13*(2), PCC.10m01013. https://doi.org/10.4088/PCC.10m01013

vonNeumann, J., & Morgenstern, O. (1953). Theory of games and economic behavior (3rd ed.). Princeton, NJ: Princeton University Press.

Wakefield, J. C. (2013). The DSM-5 debate over the bereavement exclusion: Psychiatric diagnosis and the future of empirically supported treatment. *Clinical Psychology Review*, *33*(7), 825–845. https://doi.org/10.1016/j.cpr.2013.03.007

Wakefield, J. C., & First, M. B. (2012). Validity of the bereavement exclusion to major depression: Does the empirical evidence support the proposal to eliminate the exclusion in DSM-5? *World Psychiatry*, *11*(1), 3–10. https://doi.org/10.1016/j.wpsyc.2012.01.002

Walco, D. K., & Risen, J. L. (2017). The empirical case for acquiescing to intuition. *Psychological Science*, *28*(12), 1807–1820. https ://doi.org/10.1177/0956797617723377

Weller, J.A., Levin, I.P., & Denburg, N.L. (2011). Trajectory of risky decision making for potential gains and losses from ages 5 to 85. *Journal of Behavioral Decision Making*, *24*(4), 331–344. https://doi.org/10.1002/bdm.690

Witteman, C., van den Bercken, J., Claes, L., & Godoy, A. (2009). Assessing rational and intuitive thinking styles. *European Journal of Psychological Assessment: Official Organ of the European Association of Psychological Assessment*, *25*(1), 39–47. https://doi.org/10.1027/1015-5759.25.1.39

World Health Organization. (2020). *International statistical classification of diseases and related health problems* (11th ed.). https://icd.who.int/

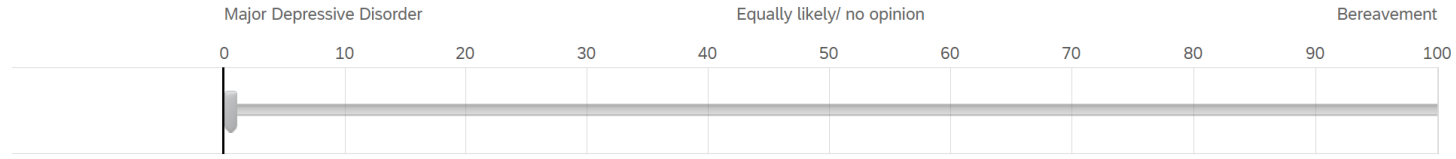
Wu, EQ., Birnbaum, HG., Shi, L., Ball, DE., Kessler, RC., & Moulis, M. (2005) The economic burden of schizophrenia in the United States in 2002. *Journal of Clinical Psychiatry, 66*:1122–1129. doi: 10.4088/JCP.v66n0906

Zalzala, A.B. & Gagen, E. C. (2023). Metacognitive reflection in supervision: the role of supervision in addressing health inequities. *Journal of Contemporary Psychotherapy*, *53*(1), 109–115. https://doi.org/10.1007/s10879-022-09561-8

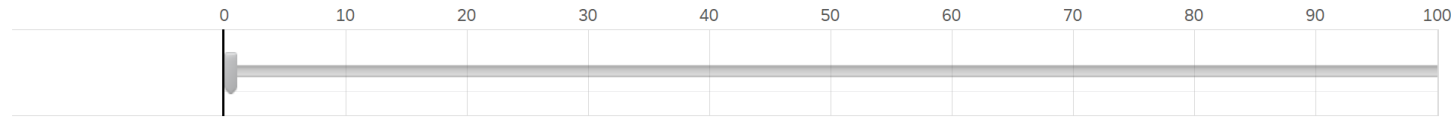
Appendix A

Diagnosis and Confidence Rating Scales

Please rate your certainty in leaning towards a preliminary diagnosis for Monica.



How confident are you with your preliminary diagnosis for Monica?



Appendix B

Demographic Questionnaire

1) What is your age? (Free-response question)
2) What gender do you identify as?
    - Cisgender Male
    - Cisgender Female
    - Transgender Male
    - Transgender Female
    - Non-Binary
    - Not listed (specify)
    - Prefer not to answer.
3) What is your racial/ethnic background? Please choose all that apply.
    - White/Caucasian
    - Hispanic/Latinx
    - Black/African American
    - Native American/American Indian
    - East Asian/ Pacific Islander
    - South Asian/ Indian
    - Arabic
    - Other (specify)
4) What is your profession?
    - Psychologist
    - Psychiatrist

- Counselor
- Social Worker
- Psychiatric Nurse
- Other (specify)

5) What is the highest degree you obtained?
   - Bachelor's
   - Master's
   - PsyD
   - PhD
   - EdD
   - MD
   - Other
6) How many years of experience do you have providing mental health services (include years of professional training)? (Free-response question)
7) Do you provide direct supervision to other mental health professionals delivering services?
   - Yes
   - No
   - Prefer not to answer.
8) Please use the rating scale slider, ranging from not at all familiar (0) to very familiar (100) for your familiarity with diagnosing the following mental disorders?
   a. Adjustment Disorder
   b. Major Depressive Disorder
   c. Persistent Depressive Disorder

d. Prolonged Grief Disorder

e. Post-traumatic Stress Disorder

9) What is your typical caseload per week?

- 1-5
- 6-10
- 11-15
- 16-20
- 21-25
- 26-30
- More than 30
- Prefer to not say.

10) What is the typical age range of your clients? Please choose all that apply.

- 11 years old and under
- 12-17 years old
- 18-24 years old
- 25-39 years old
- 40-60 years old
- Over 60 years old

11) In what kind of practice setting do you work? Please check all that apply.

- Outpatient
- Inpatient (non-psychiatric) Hospital
- Psychiatric Hospital
- University Setting

- Private Practice
- Rehabilitation Facility
- Nursing Home
- Telehealth
- Other
- Prefer to not say.

12) What percentage of your week is devoted to the following responsibilities? (Free-response question)

a. Psychotherapy

b. Assessment

c. Teaching

d. Research Activities

e. Mentoring

f. Clinical Supervision

Total percentage: ____

Appendix C

Interactive Interview

Opening Prompt

Monica just scheduled an intake interview with you. You meet her at the front desk of the clinic and she follows you back to the therapy room. You can see that she is walking somewhat slowly, and she seems to have a downcast look on her face. After you sit down, you ask her, "What brings you in today?" She responds, "I've been feeling really down recently."

Questions and Answers

Major Depressive Disorder

How long have you been feeling this way?	It has been about 2 months.
Tell me a little more about what you mean when you say you feel down.	It is a profound feeling of emptiness, like there is a hole in my life.
Have you been feeling down consistently during this time or has the feeling come and gone?	It has been coming and going, depending upon what I'm doing, but it feels like it is bad most of the time.
How has your sleep been?	I have been having a real hard time falling asleep. It seems like I am lying awake half the night, despite wanting to fall asleep.
How is your energy level during the day?	I feel completely exhausted. Most days, I have to lay down and take a nap.

Have you noticed or has anyone else said that you seem to be moving slower than usual?	I am sluggish all day. My kids also see that I just don't have the energy to drive them around to practices like I use to.
Any difficulties concentrating? Reading a book or following a movie?	I have really been struggling to concentrate at work. I use to read for 30 minutes before bed and I can't even do that anymore. I read a chapter and forget what happened by the end.
How have your symptoms impacted you at work?	I can't seem to focus and I am falling really far behind. I've been making simple mistakes I normally would catch.
Does how you are feeling stop you from enjoying things you normally like doing?	I don't really have any hobbies. The things I do in my spare time have always been what my husband wants to do.
*What do [did] you do with your husband?	We primarily would go out to bars, watch sports, and hang out with his friends.
*Have you stopped doing these activities?	Yes.
*What has led you to stop doing these things?	*tearfully* My husband died in a car crash 2 months ago.
Have there been any changes in your appetite?	No, my appetite has been pretty much the same but I have never been one to eat very much.

Have you noticed any significant weight gain or weight loss?	No, I don't think so.
Have you had any thoughts about wanting to kill yourself [or wanting to die too]?	No, never.
How do you feel about the future?	I know things will get better, and I will eventually figure out how to move on, but it still so hard right now.
*Do you have any feelings of guilt?	All the time. I know there is nothing I could have done, but I feel like I should have been there. Maybe things would have gone differently if I were in the car with him.
Have you been drinking or using any other drugs during this time?	No. I've never been much of a drinker anyway. Maybe a glass of wine at a wedding or something. Never done drugs either.
Have you ever had a period of time where you felt much better than usual, like you were really excited or energetic?	No.
Have you been diagnosed with any medical conditions or are you taking medication for anything?	I only take medication for high blood pressure.

Have you started developing beliefs that others may view as odd or seen/heard things that others were not able to see or hear?	No, never anything like that.

Normal Grief Response

Have any big changes happened in your life recently?	Yes, my husband passed away about 2 months ago.
*I'm so sorry to hear that. How did he pass away?	He died in a car crash.
*May I ask what happened? Were you with him?	No. I asked him to run an errand for me. It was a rainy night, and he lost control of the car and rolled off a steep embankment. Thankfully, no one else was hurt.
*Do you feel guilty about what happened?	I know it was not my fault, but I still feel that if I had just not asked him to go, or gone myself, none of this would have happened.
How have things been with your family and friends?	They are all very supportive, but I get the sense that they are walking on eggshells around me because they do not want to upset me.

How have you been feeling physically?	I feel exhausted. I can't sleep and when I can I have really vivid dreams of my husband.
*Is it painful to think of memories of your husband?	At times, especially when I think about how he died, but most of my memories about him are happy ones.
*Have you done anything to avoid thinking of the accident or your husband because it's too difficult?	At first, the first few weeks, I attempted to avoid the road where my husband had lost control but not anymore. It's really difficult at times to think of my husband but I do not avoid it.
*How have you been coping with the loss?	Not very well. I feel like I am coming apart at the seams. I can't believe he is gone.
*Have you lost other important people in your life before?	No, this is the first time.
*How often do you find yourself thinking about your husband?	A lot of the time. At any moment something might remind me of him. It's comforting even if I get sad.
*What are your beliefs about death and what happens when we die?	I don't know. I used to be believe in God but I am not so sure anymore.
*Has this experience changed your outlook on life?	I just don't see what the point is anymore.

*Have you been able to do anything to honor your husband's life?	We had a funeral but it wasn't what he would have wanted. It was terrible.
*Do you still speak to your husband? Or even think he is still here?	I still talk to him out loud when I'm upset or when something reminds me of something he would think is funny. But I know he isn't really here. I mean he is, but you know what I mean.

*These questions are dependent on previous questions being asked first.

Appendix D

Rational-Experiential Inventory–40 (Pacini & Epstein, 1999)

Instructions: Using the following scale, please rate the extent that these items refer to you.

1	2	3	4	5
Definitely true of myself				Definitely not true of myself

Rationality scale

Rational Ability

1) I'm not that good at figuring out complicated problems*

2) I am not very good at solving problems that require careful logical analysis*

3) I am not a very analytical thinker*

4) Reasoning things out carefully is not one of my strong points*

5) I don't reason well under pressure*

6) I am much better at figuring things out logically than most people

7) I have a logical mind

8) I have no problem thinking things through carefully

9) Using logic usually works well for me in figuring out problems in my life

10) I usually have clear, explainable reasons for my decisions

Rational Engagement

11) I try to avoid situations that require thinking in depth about something*

12) I enjoy intellectual challenges

13) I don't like to have to do a lot of thinking*

14) I enjoy solving problems that require hard thinking

15) Thinking is not my idea of an enjoyable activity*

16) I prefer complex problems to simple problems

17) Thinking hard and for a long time about something gives me little satisfaction*

18) I enjoy thinking in abstract terms

19) Knowing the answer without having to understand the reasoning behind it is good enough for me*

20) Learning new ways to think would be very appealing to me

Experientiality scale

Experiential Ability

21) I don't have a very good sense of intuition*

22) Using my gut feelings usually works well for me in figuring out problems in my life.

23) I believe in trusting my hunches

24) I trust my initial feelings about people

25) When it comes to trusting people, I can usually rely on my gut feelings

26) If I were to rely on my gut feelings, I would often make mistakes*

27) I hardly ever go wrong when I listen to my deepest gut feelings to find an answer

28) My snap judgments are probably not as good as most people's*

29) I can usually feel when a person is right or wrong, even if I can't explain how I know

30) I suspect my hunches are inaccurate as often as they are accurate*

Experiential Engagement

31) I like to rely on my intuitive impressions

32) Intuition can be a very useful way to solve problems

33) I often go by my instincts when deciding on a course of action

34) I don't like situations in which I have to rely on intuition*

35) I think there are times when one should rely on one's intuition

36) I think it is foolish to make important decisions based on feelings*

37) I don't think it is a good idea to rely on one's intuition for important decisions*

38) I generally don't depend on my feelings to help me make decisions*

39) I would not want to depend on anyone who described himself or herself as intuitive(-)

40) I tend to use my heart as a guide for my actions

Note: Labels should be removed and items randomized prior to administration. Items marked with an asterisk (*) should be reverse coding prior to scoring. Subscale scores are computed by averaging the 10 composite items.

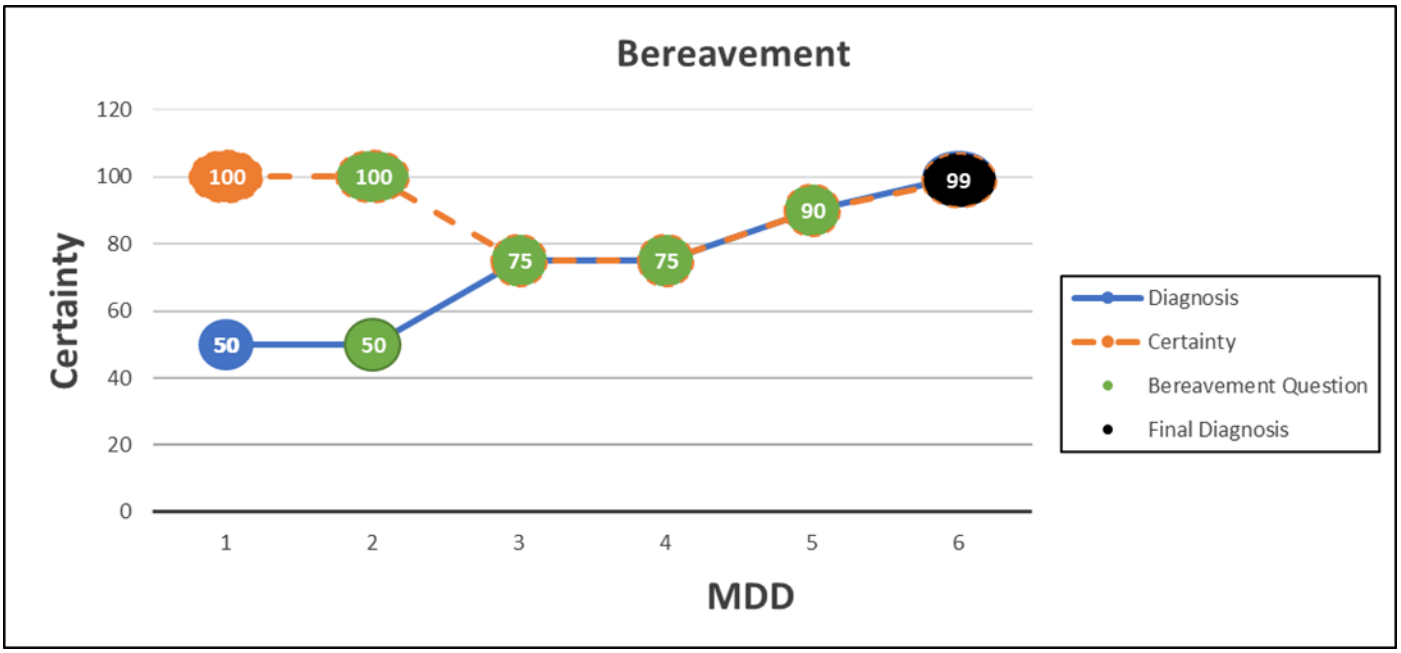
Appendix E

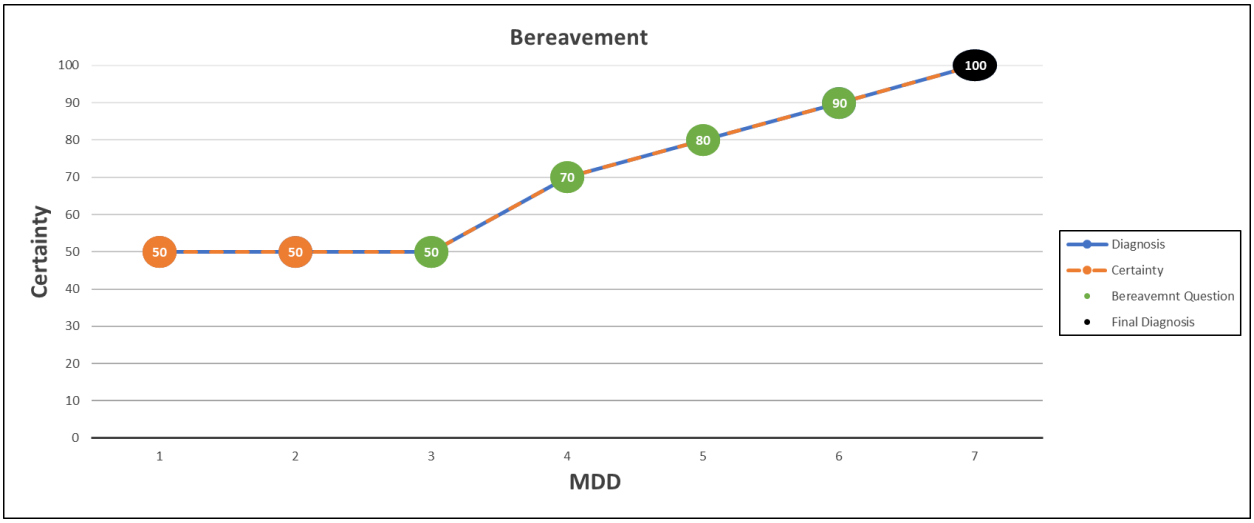
Clinical Relevance Questions

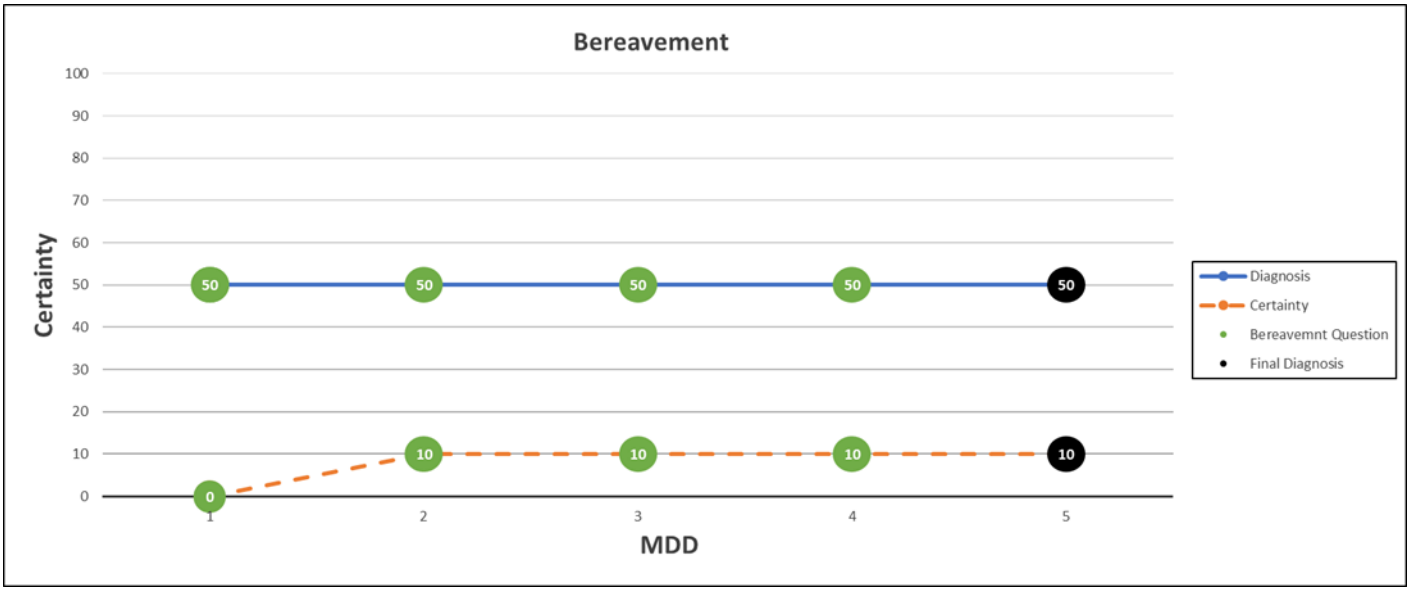
1) What diagnostic features were most indicative for you to arrive at a diagnostic decision? (Free-response question)
2) What diagnostic features were absent which would have better supported the differential diagnosis? (Free-response question)
3) What diagnostic materials (e.g., DSM, ICD) or other tools (e.g., SCID) guided your decision-making process when selecting diagnostic questions to ask? (Free-response question)
4) Do you agree that Major Depressive Disorder and normative grief exist on a spectrum?
    a. Yes
    b. No
    c. Unsure
    d. Specify

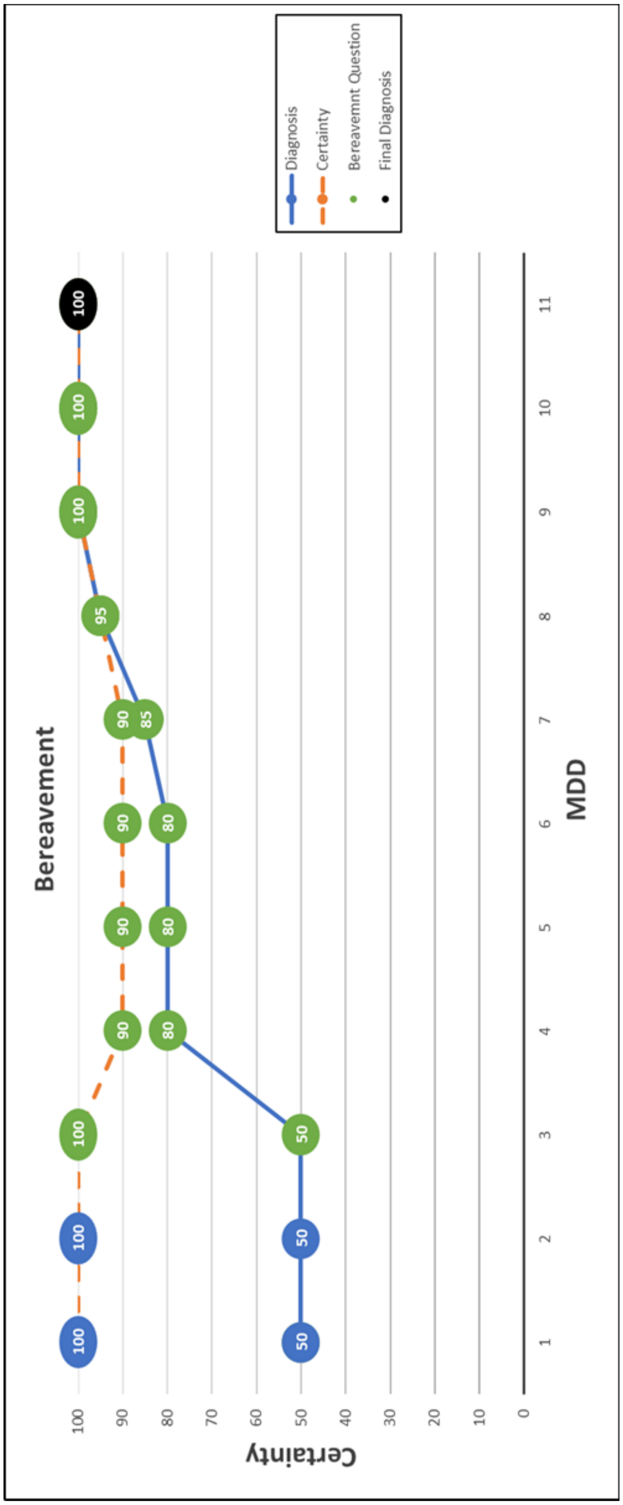
Appendix F

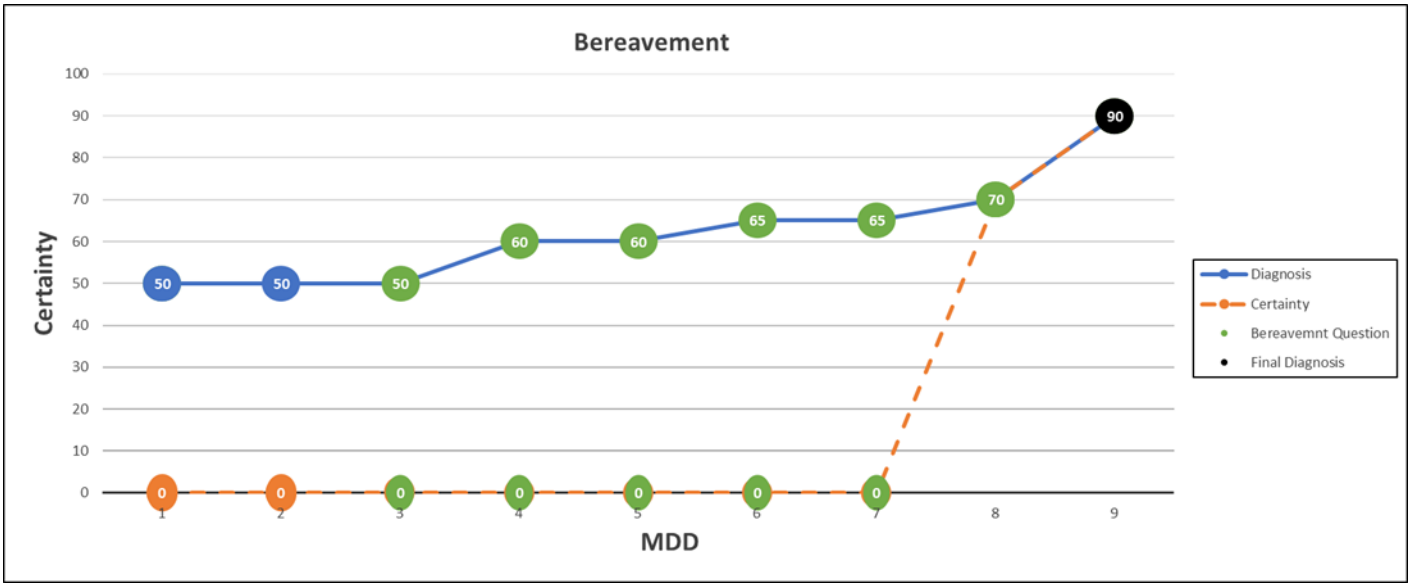
Participants With Limited System 2 Activations

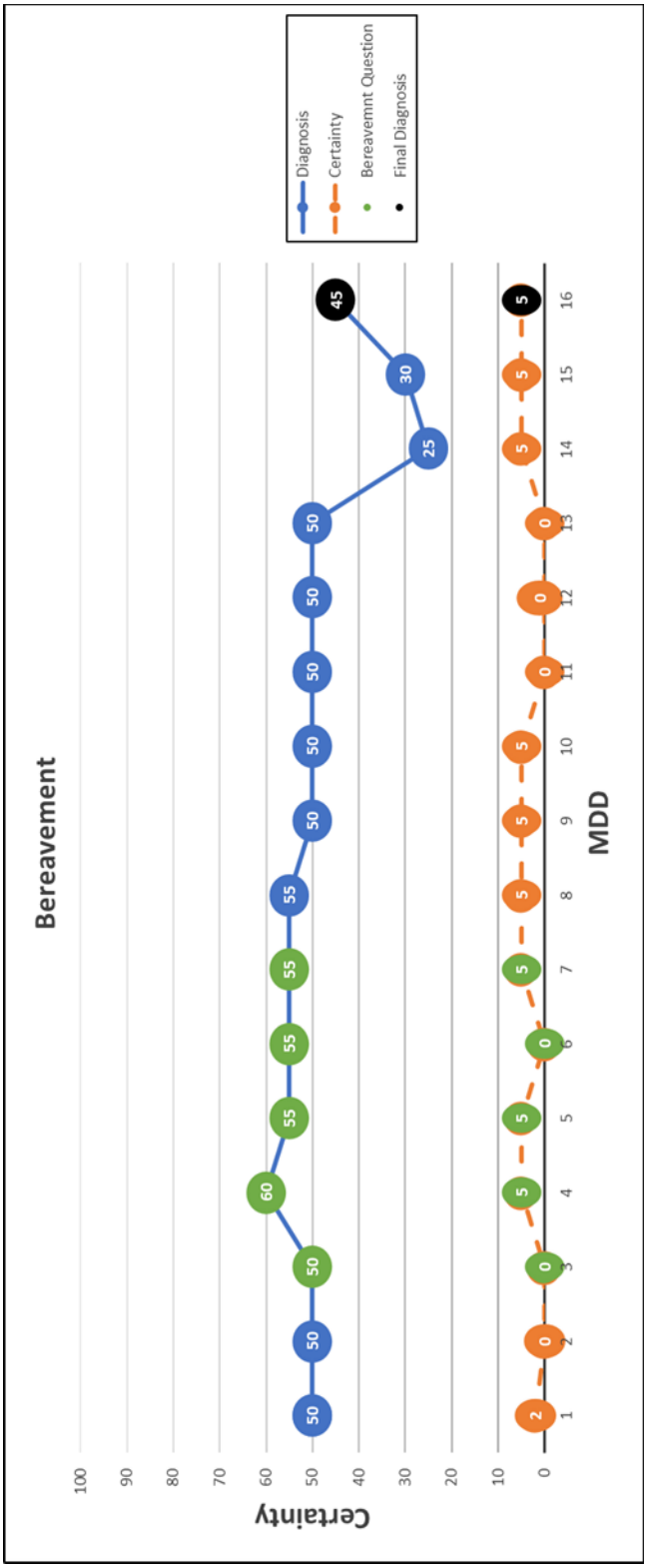


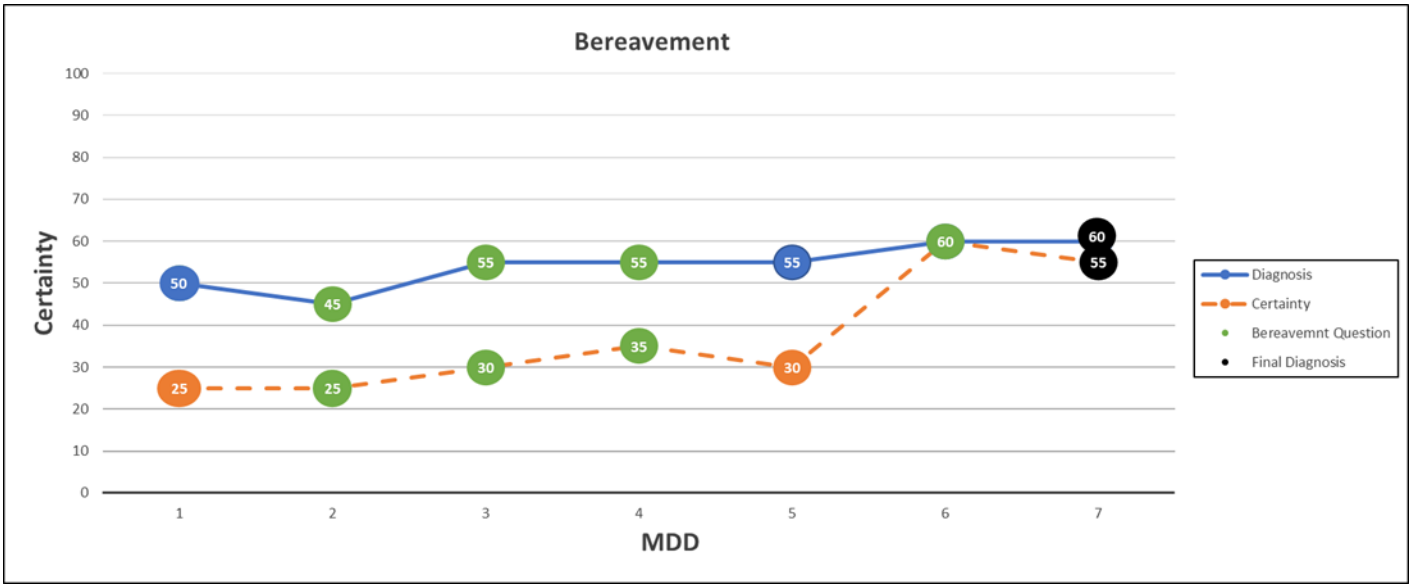

Bereavement
Certainty
MDD
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
50
50
50
70
80
90
100
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis

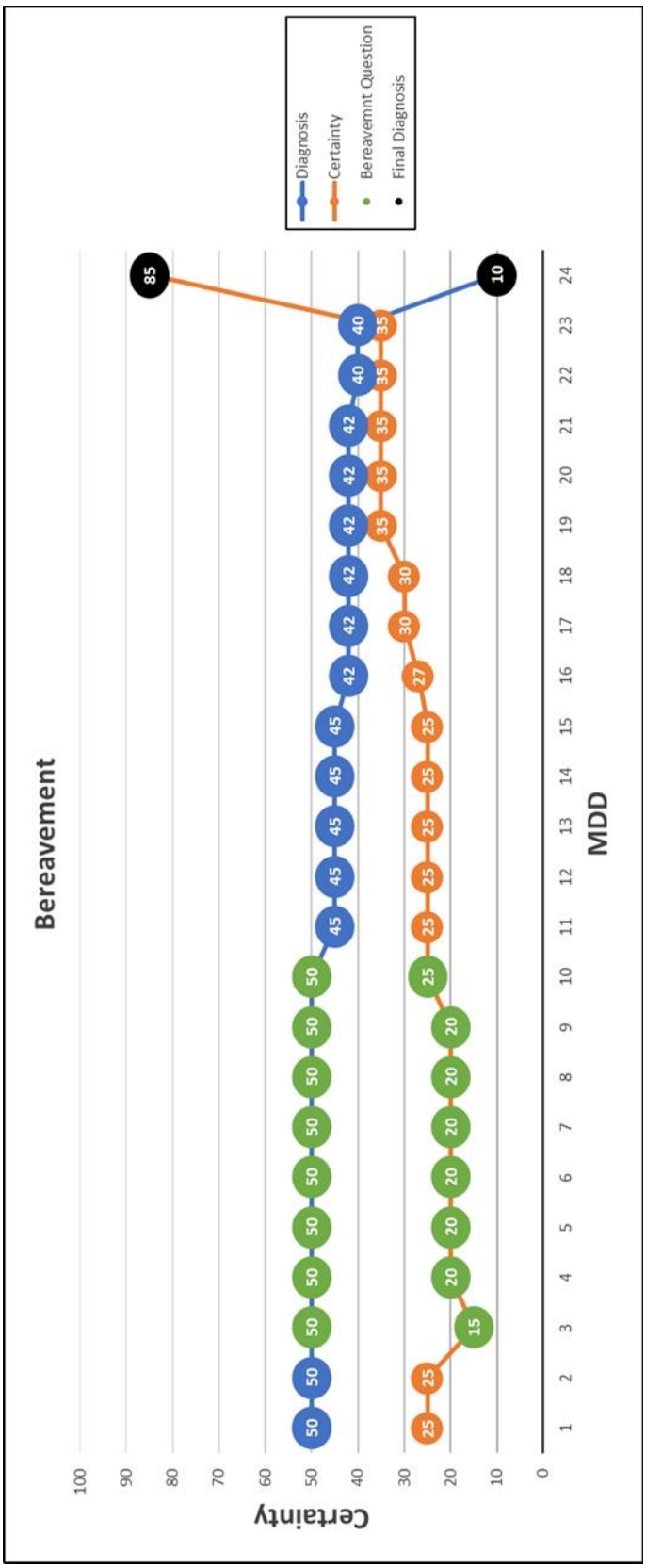

Bereavement
100
90
80
70
60
50
40
30
20
10
0
Certainty
50
50
50
50
50
10
10
10
10
0
1
2
3
4
5
MDD
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis

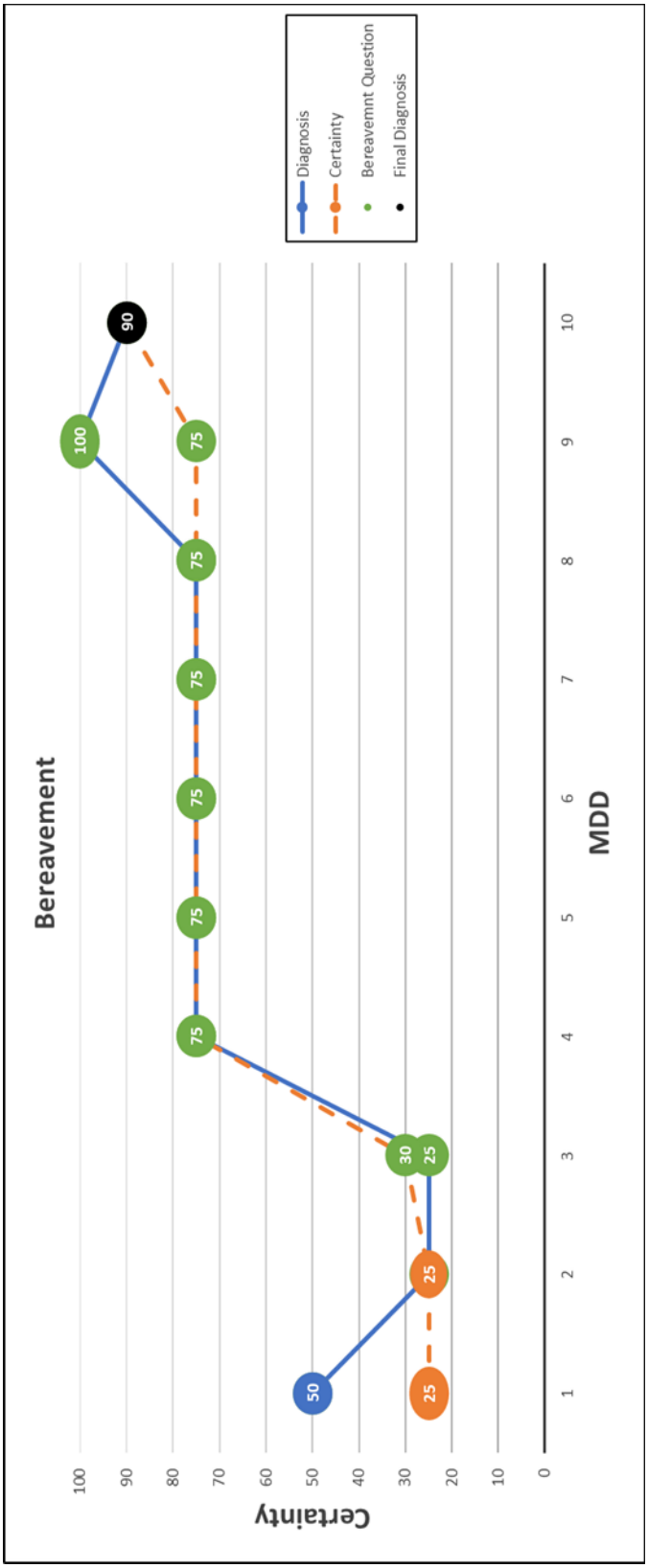

Bereavement
Certainty
MDD
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis

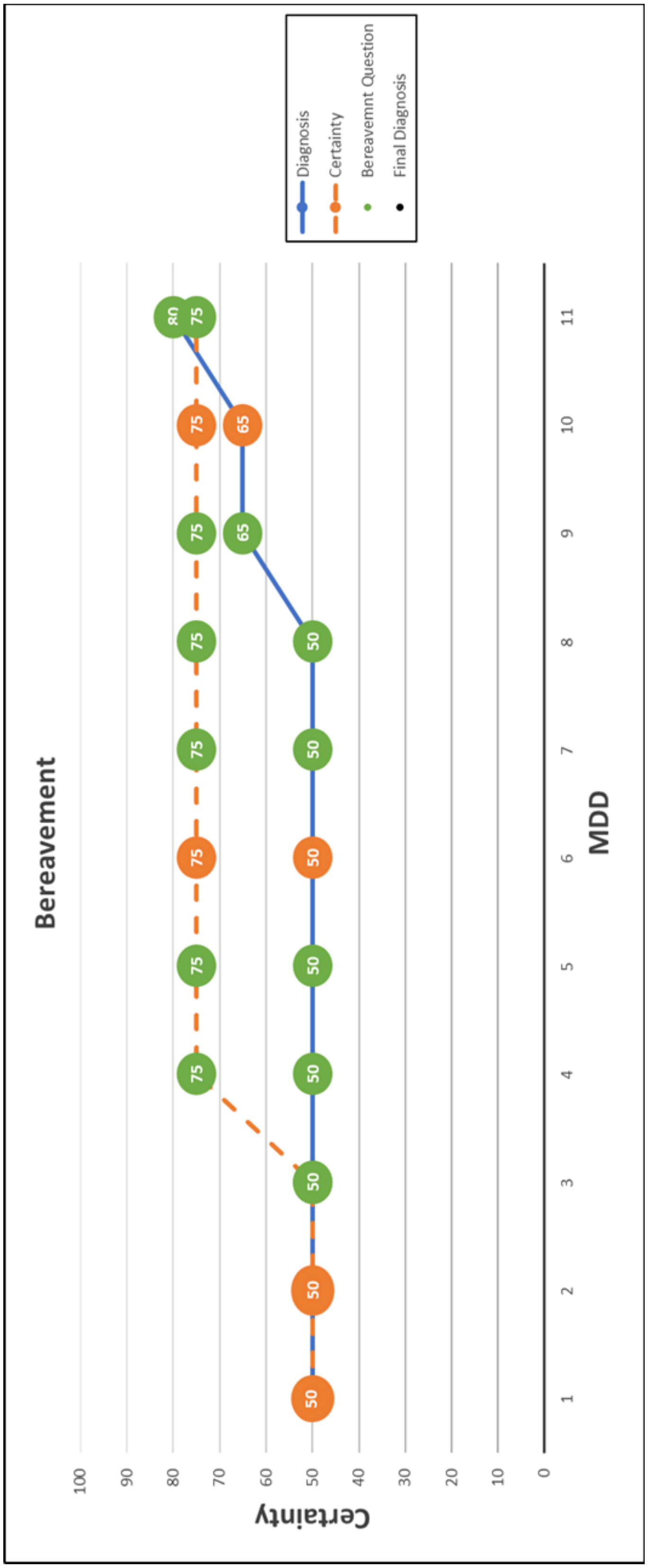

Bereavement
Certainty
MDD
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
50
50
50
60
60
65
65
70
90
0
0
0
0
0
0
0
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis

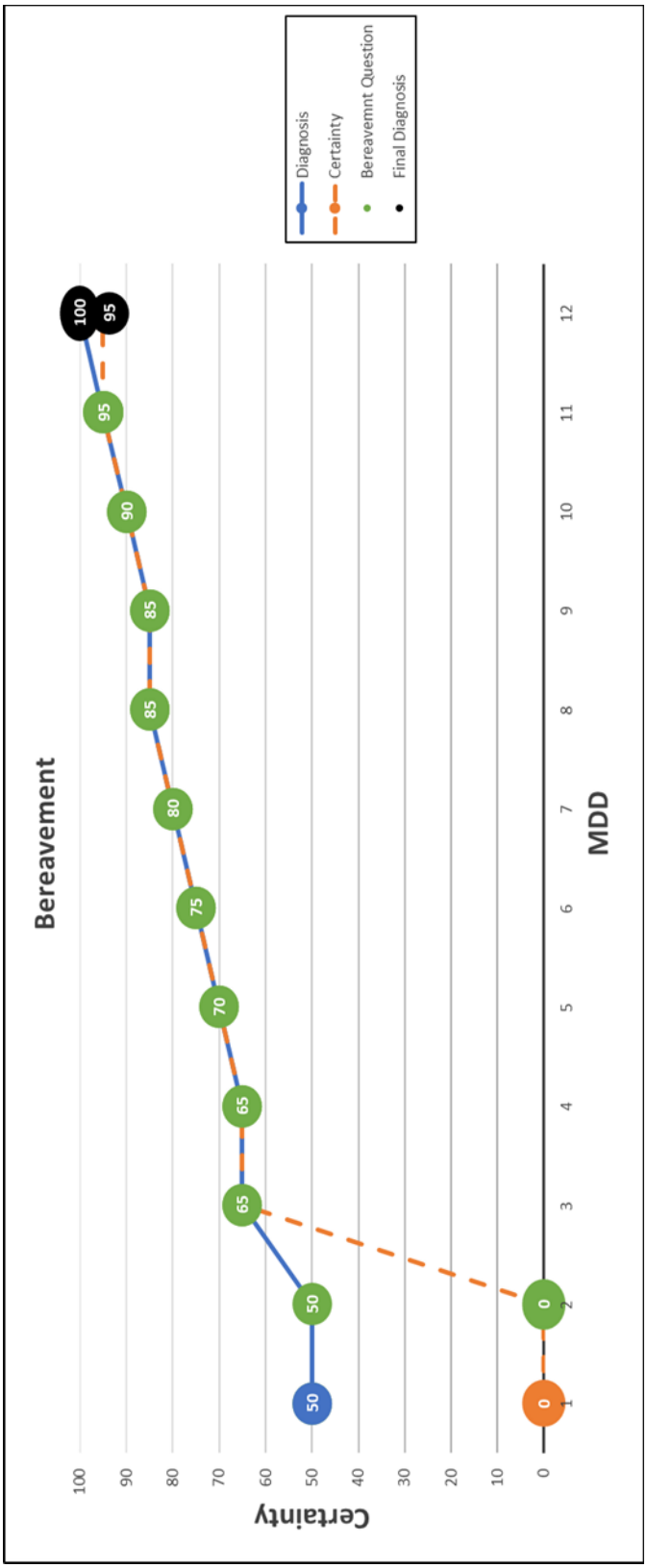

Bereavement
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
Certainty
MDD

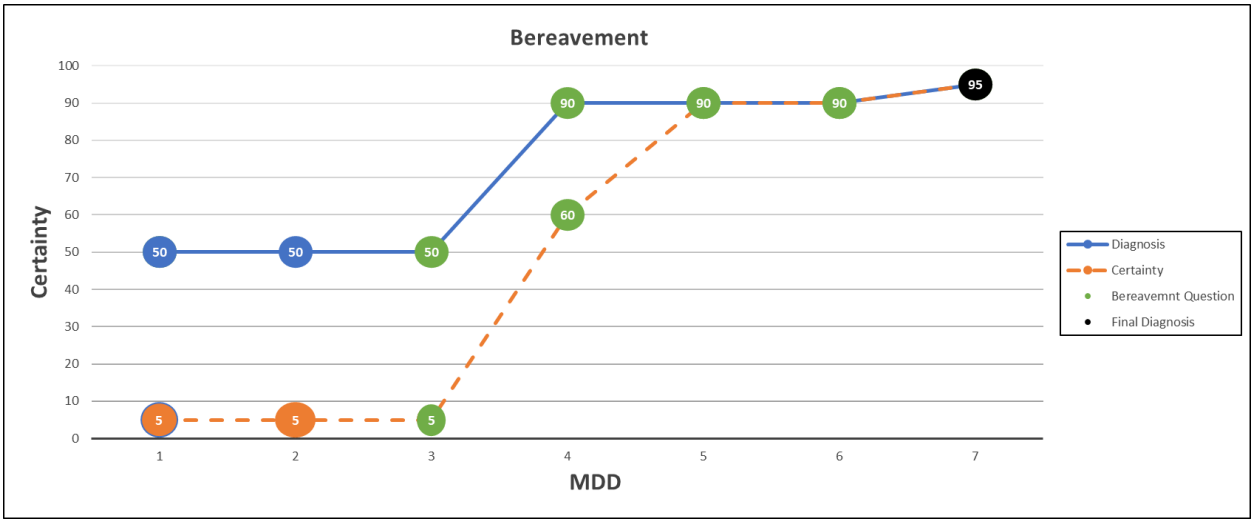

Bereavement
Certainty
MDD
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
50
45
55
55
55
60
60
25
25
30
35
30
55
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis

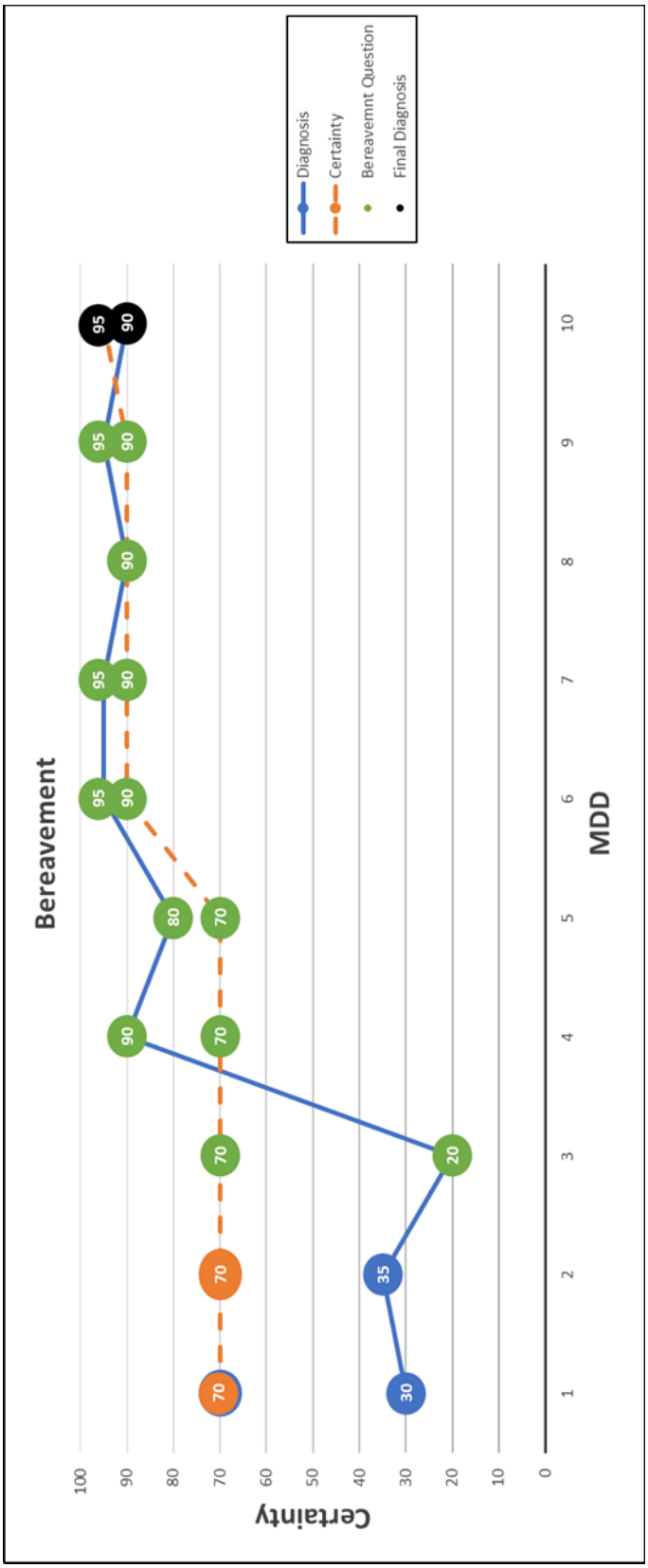

Bereavement
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
Certainty
MDD
85
10


Bereavement
Certainty
MDD
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
90
100
75
75
75
75
75
75
30
25
25
25
50


Bereavement
Certainty
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
11
MDD
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis

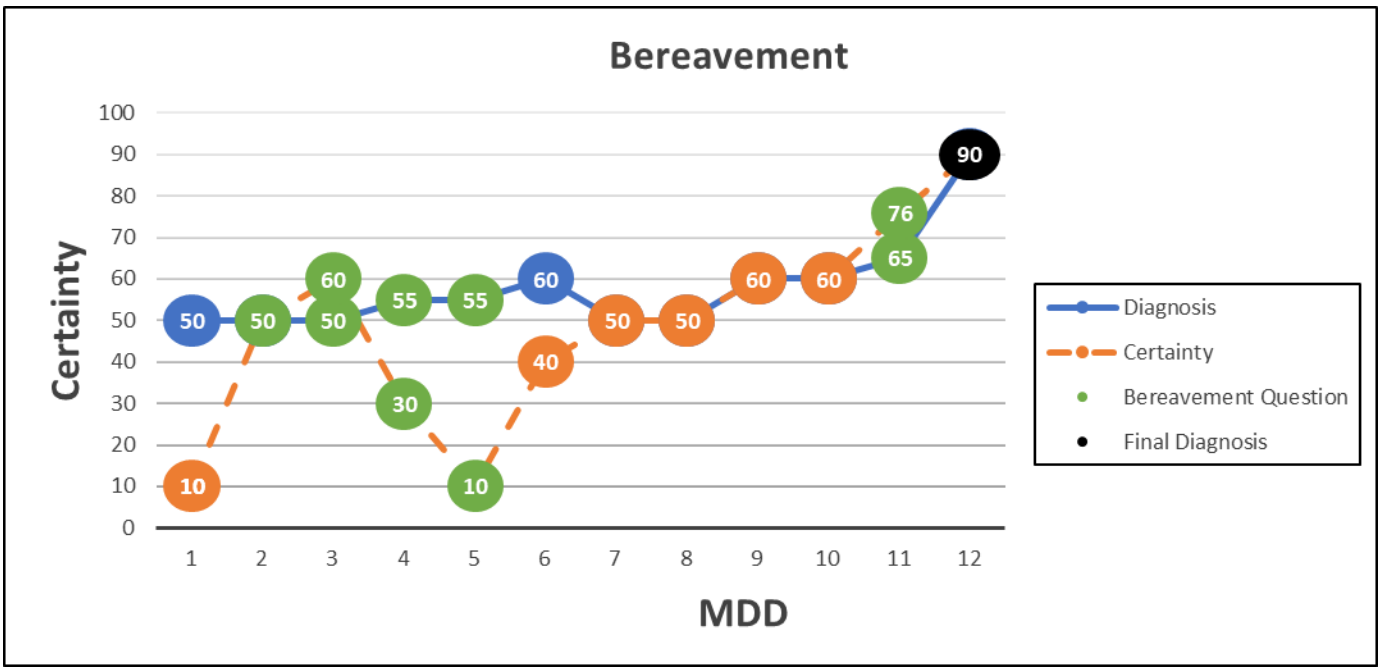

Bereavement
Certainty
MDD
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
100
95
95
90
85
85
80
75
70
65
65
50
50
0
0

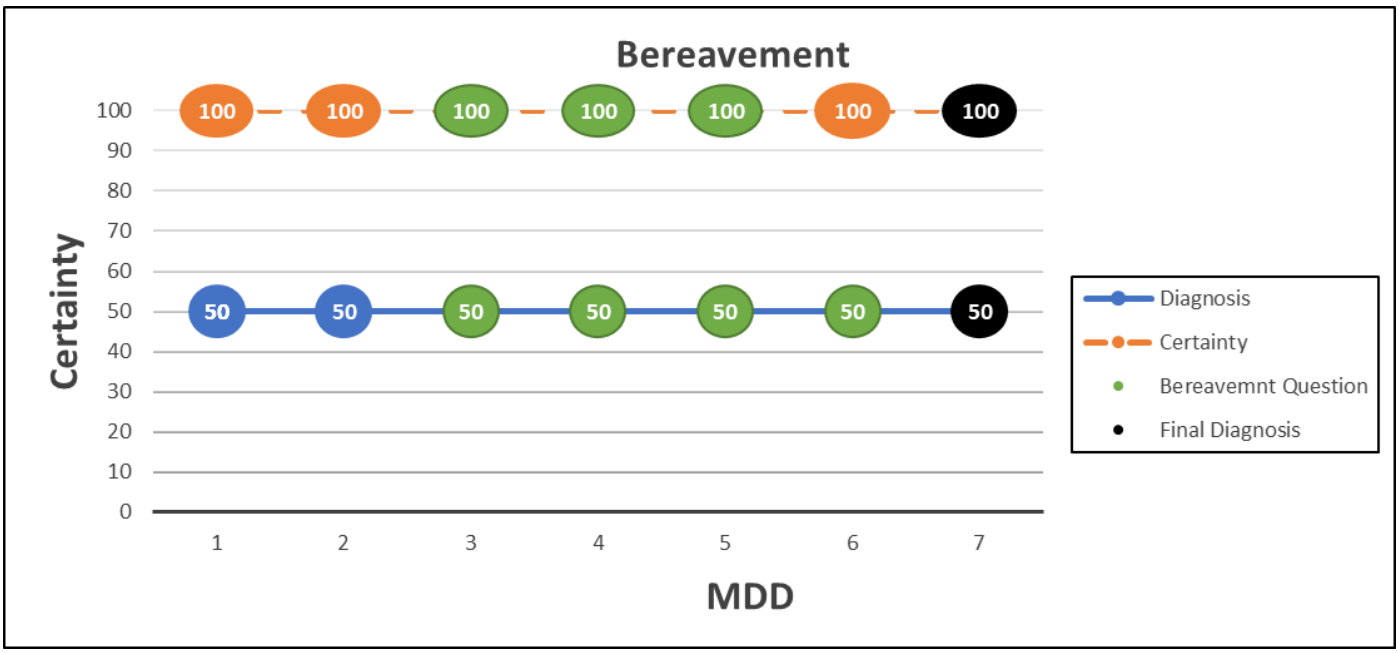

Bereavement
Certainty
MDD
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
50
50
50
90
90
90
95
5
5
5
60
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis

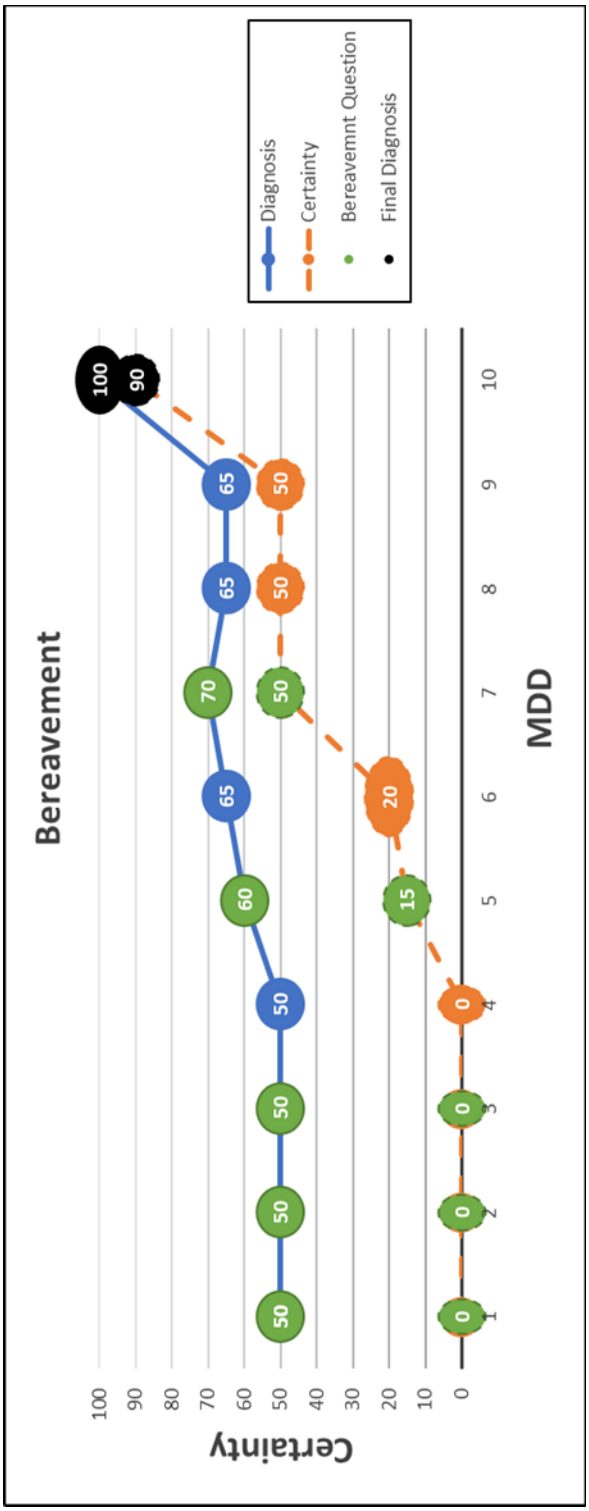

Bereavement
Certainty
MDD
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
95
90
80
70
35
30
20

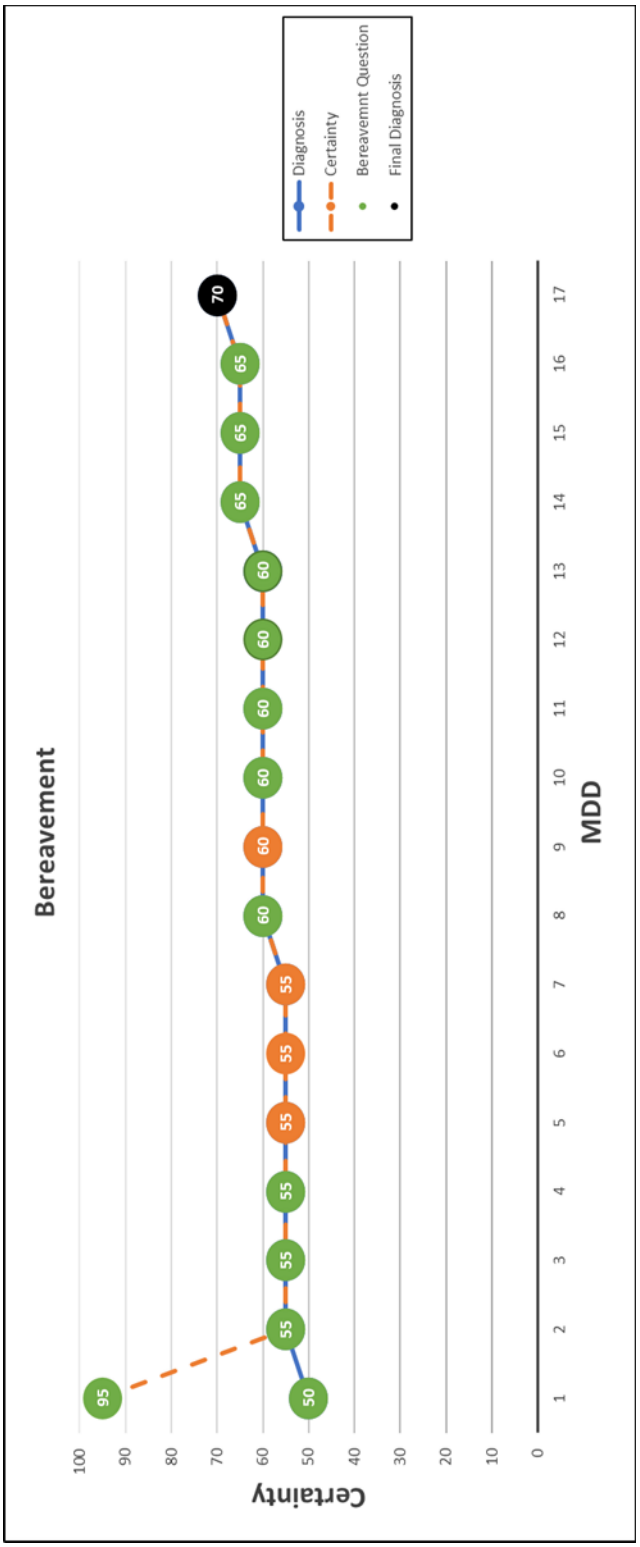
Appendix G

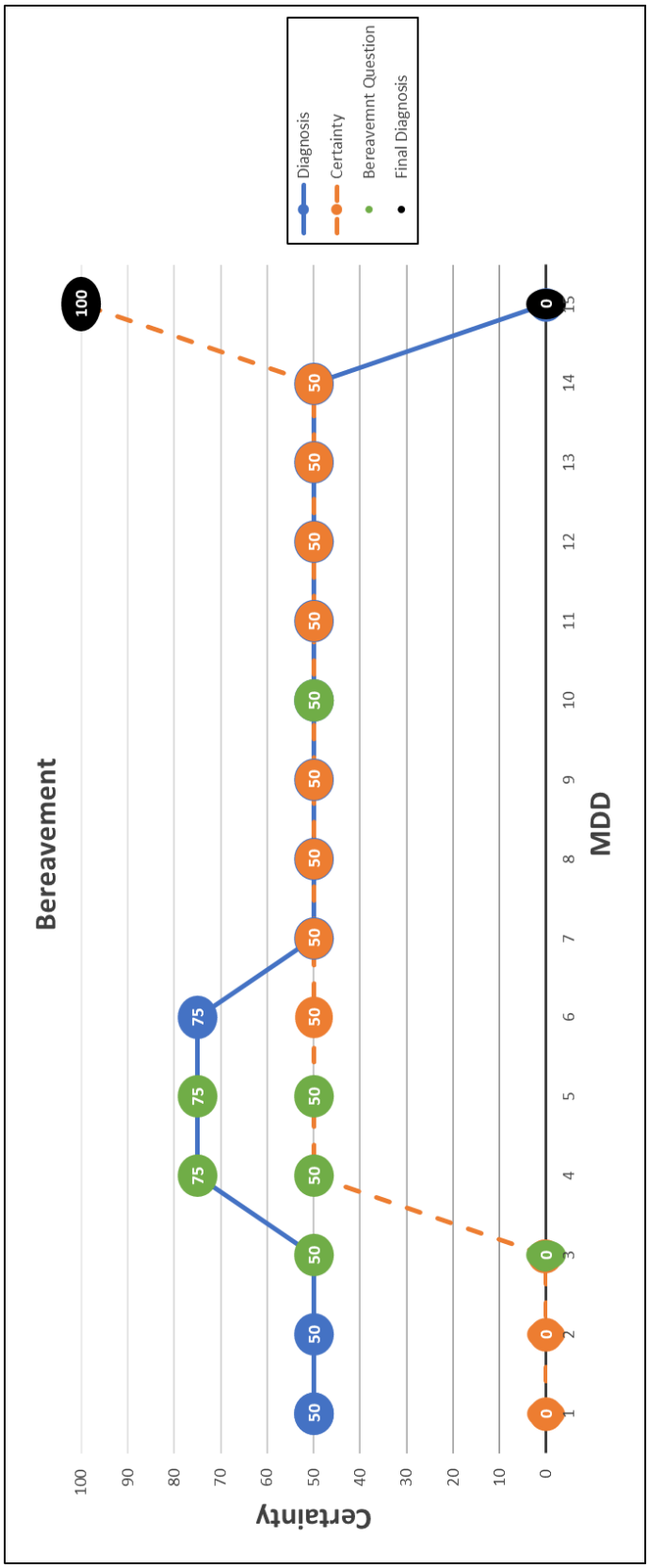
Participants With Multiple System 2 Activations

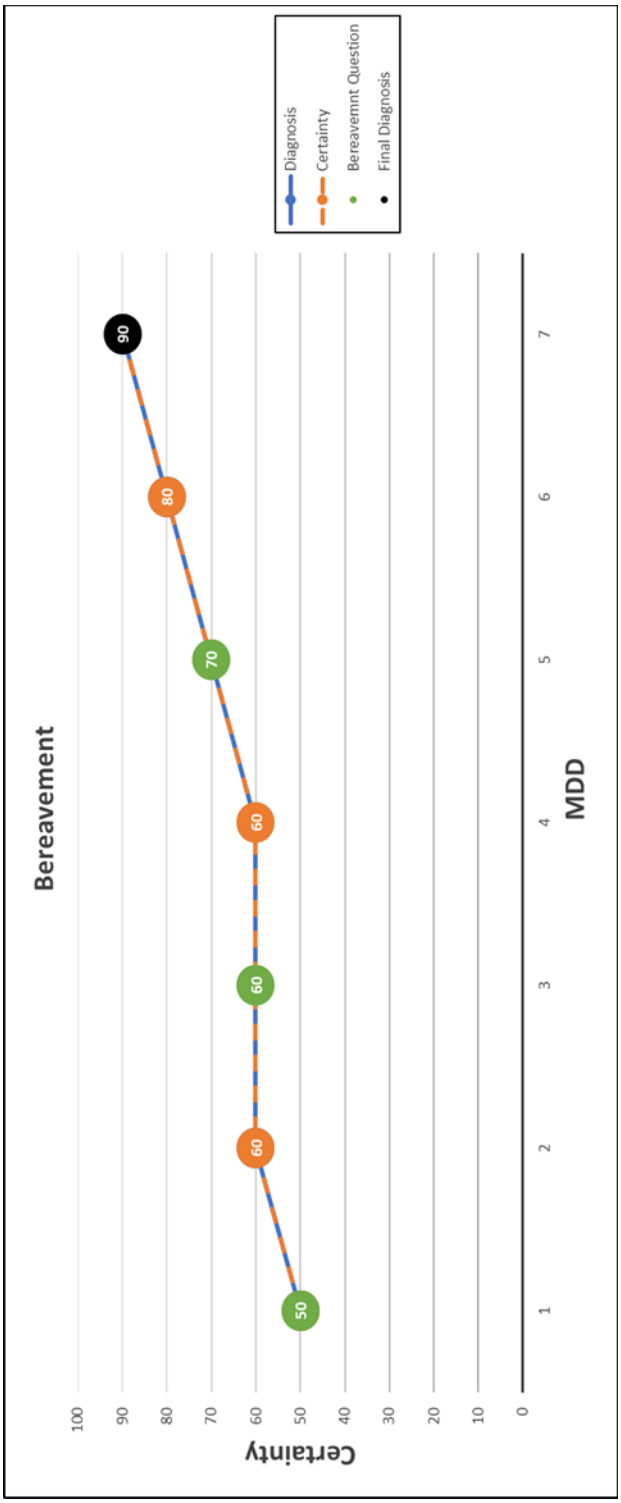


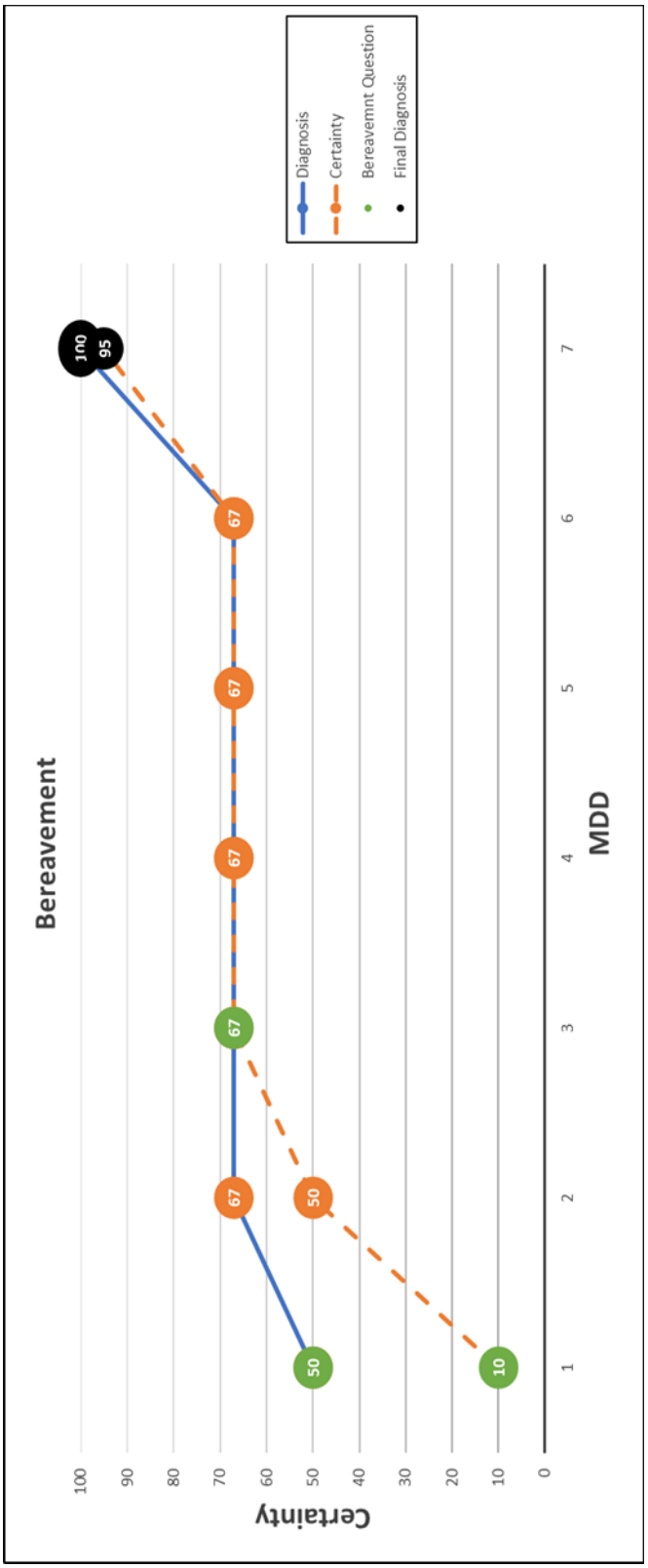

Bereavement
100
90
80
70
60
50
40
30
20
10
0
Certainty
100
100
100
100
100
100
100
50
50
50
50
50
50
50
1
2
3
4
5
6
7
MDD
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis

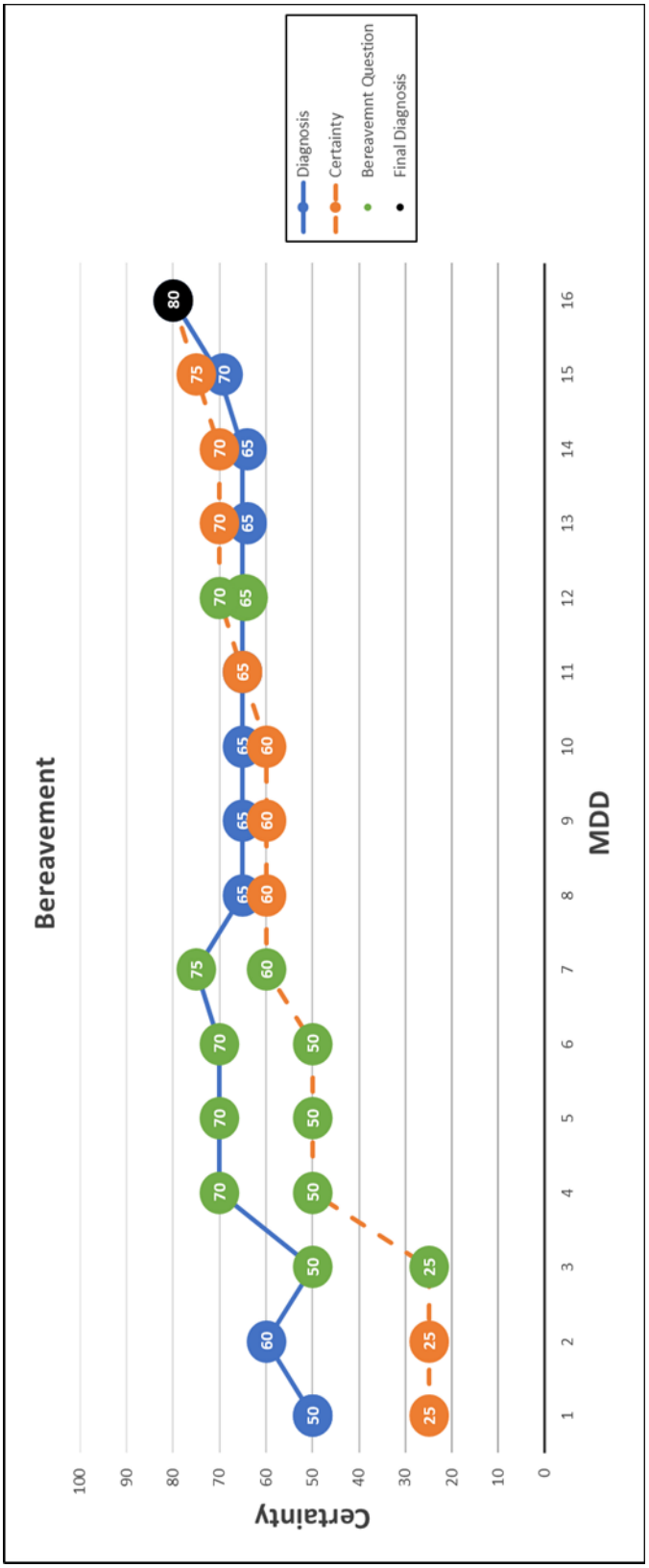

Bereavement
MDD
Certainty
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
100
90
65
65
70
65
60
50
50
50
50
50
50
20
15
0
0
0
0
0
0
0
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10

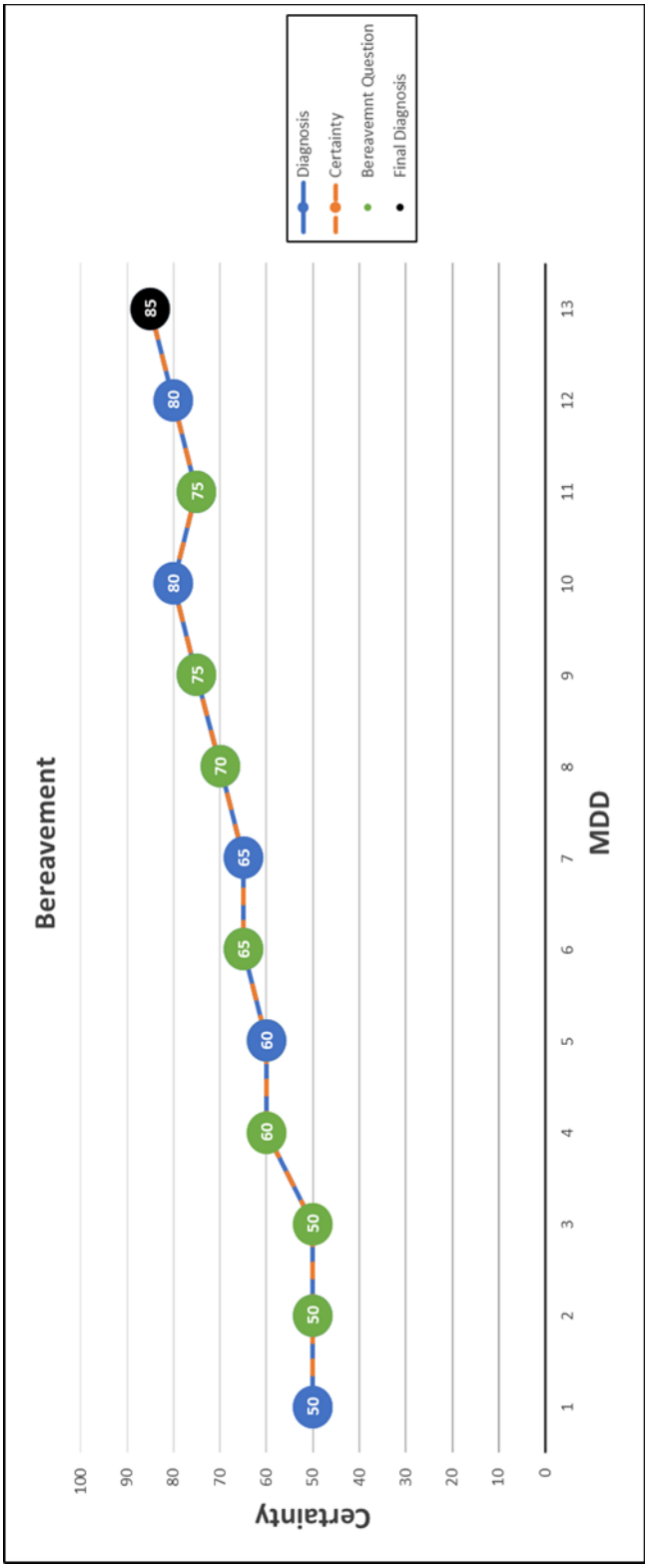

Bereavement
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
Certainty
MDD
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
95
55
55
55
55
55
55
60
60
60
60
60
60
65
65
65
70

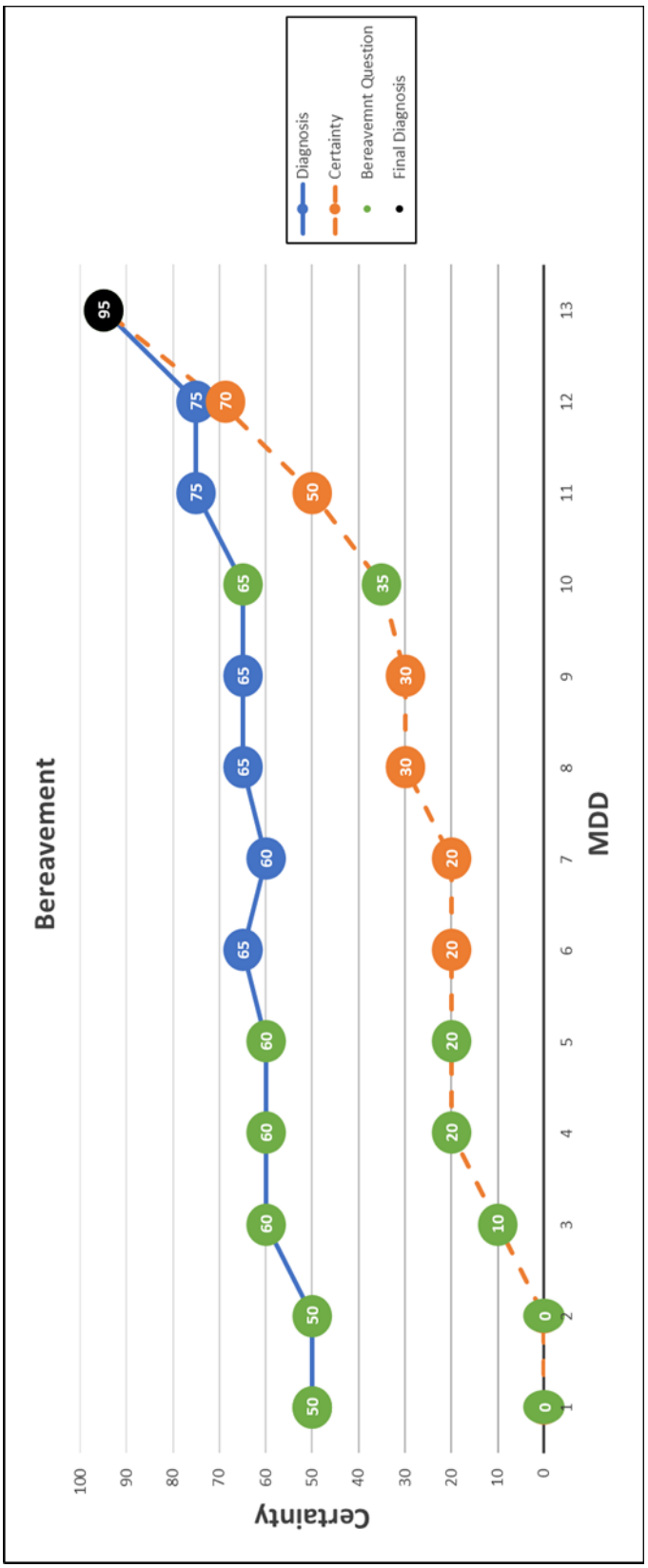

Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
Bereavement
MDD
Certainty
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

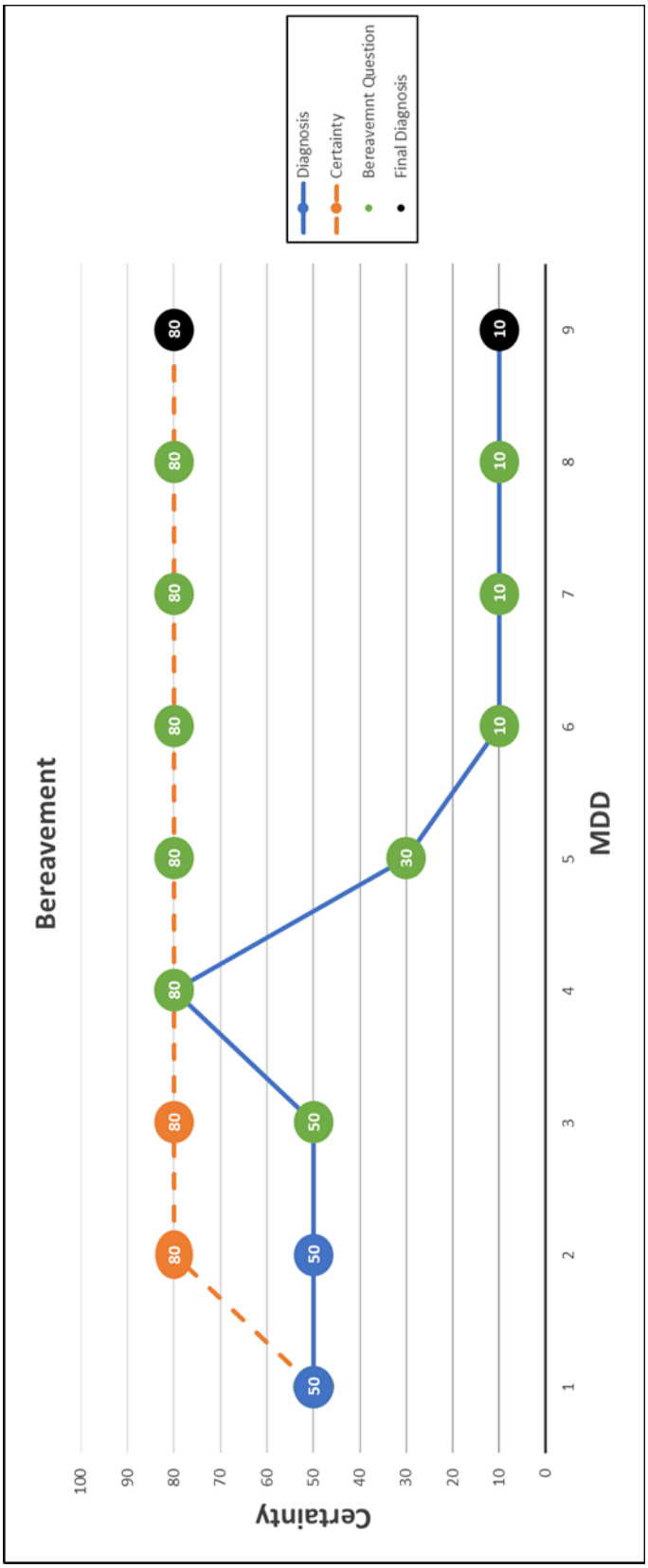

Bereavement
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
Certainty
MDD
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
90
80
70
60
60
60
50

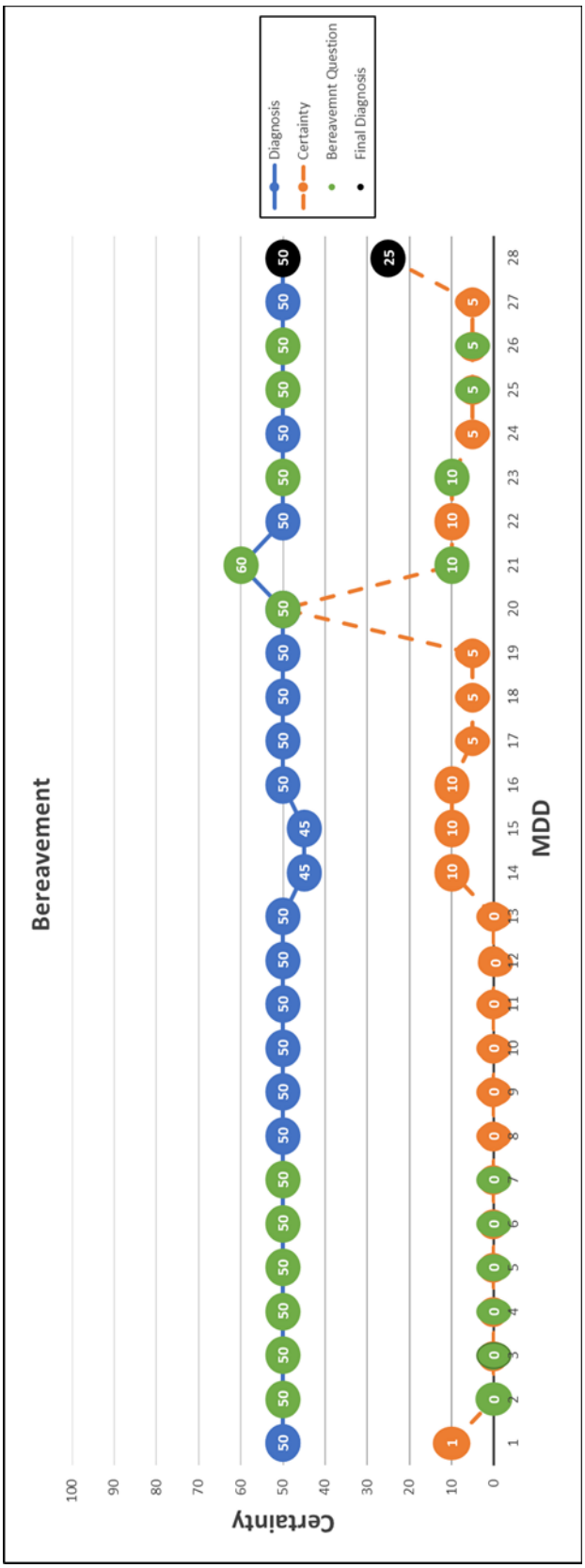

Bereavement
Certainty
MDD
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
100
95
67
67
67
67
50
50
10

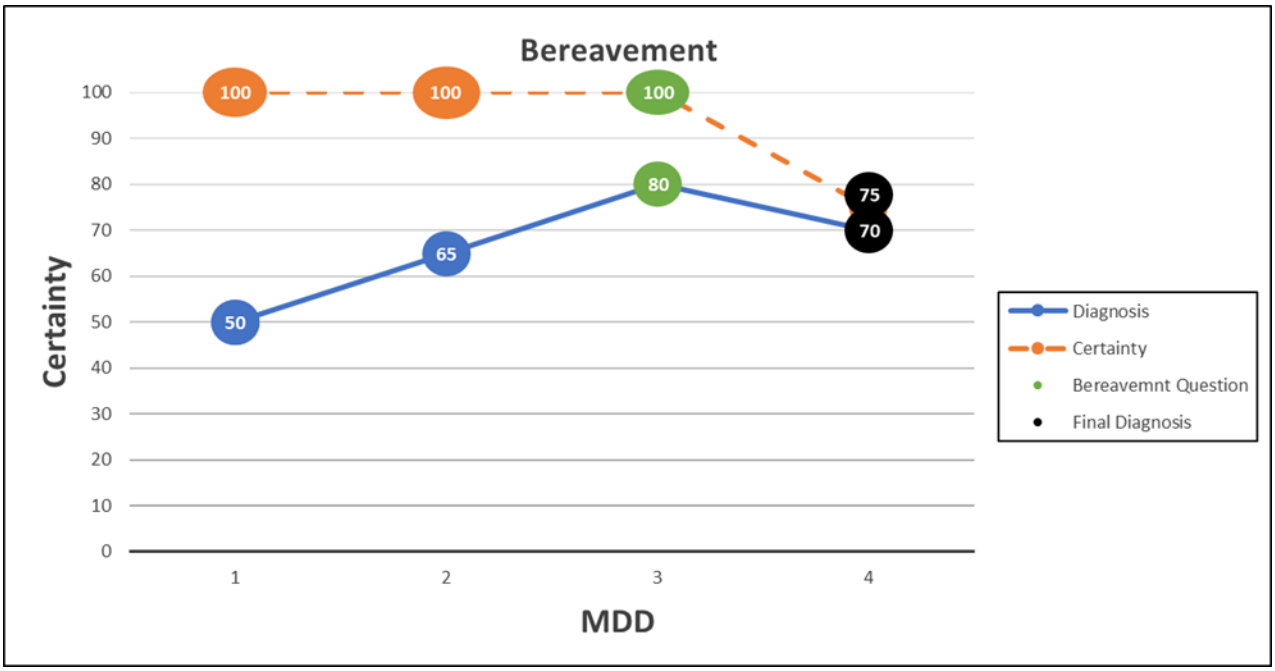

Bereavement
Certainty
MDD
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis

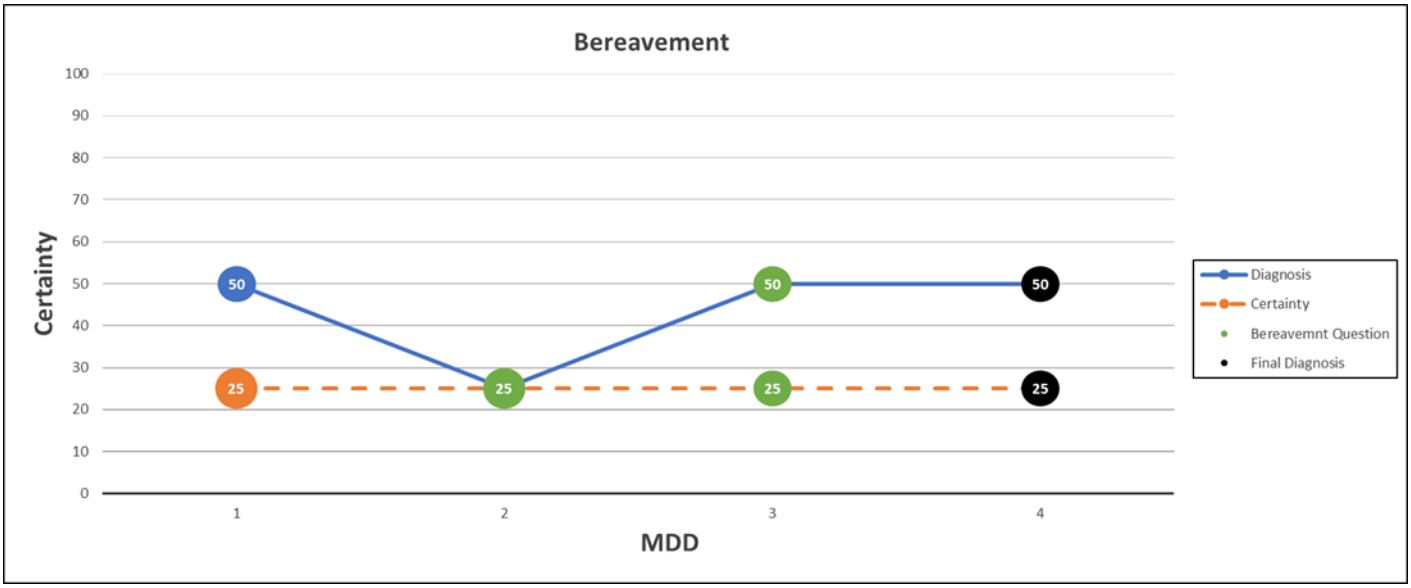

Bereavement
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
Certainty
MDD
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
11
12
13
85
80
75
80
75
70
65
65
60
60
50
50
50

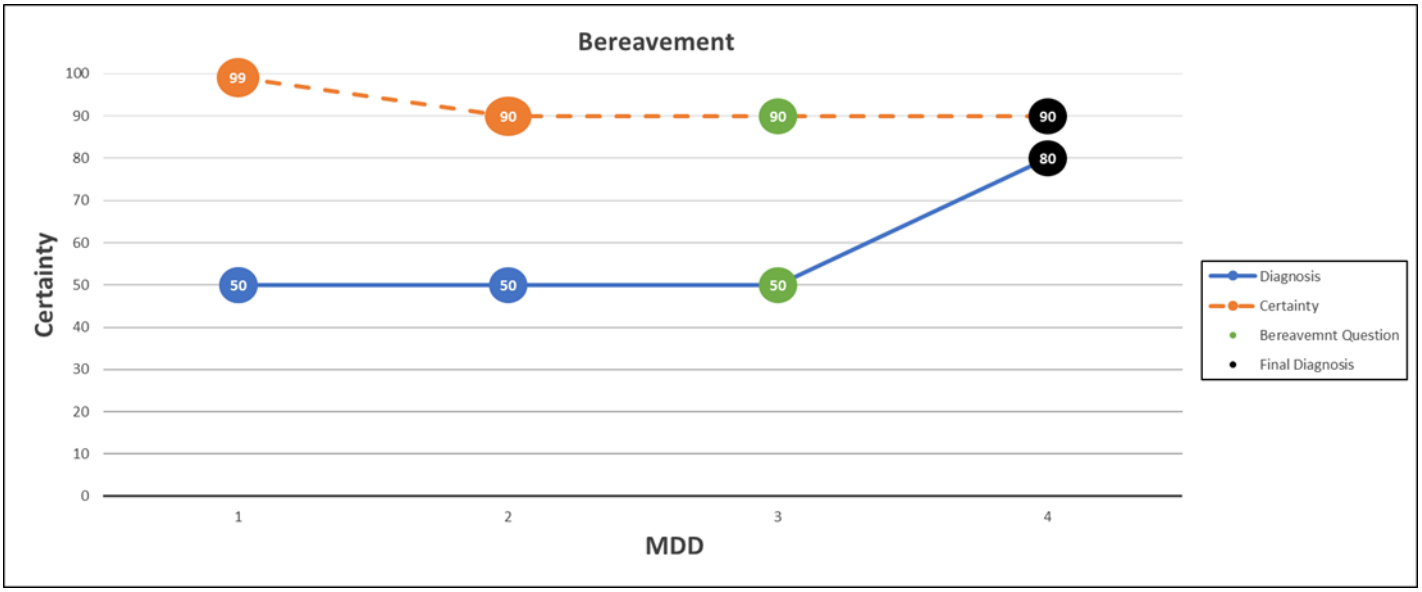

Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
Bereavement
MDD
Certainty
100
90
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
9
10
11
12
13
95
75
75
65
65
65
60
65
60
60
60
50
50
70
50
35
30
30
20
20
20
20
10
0
0

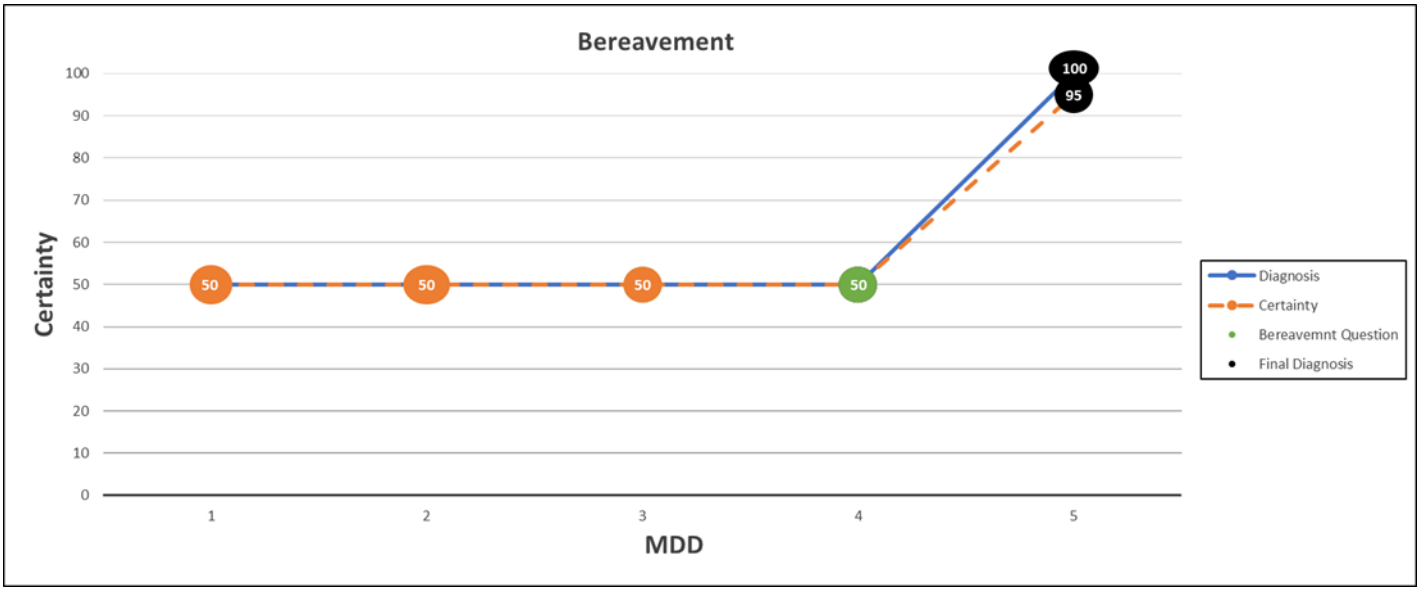

Bereavement
Certainty
MDD
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
80
50
30
10

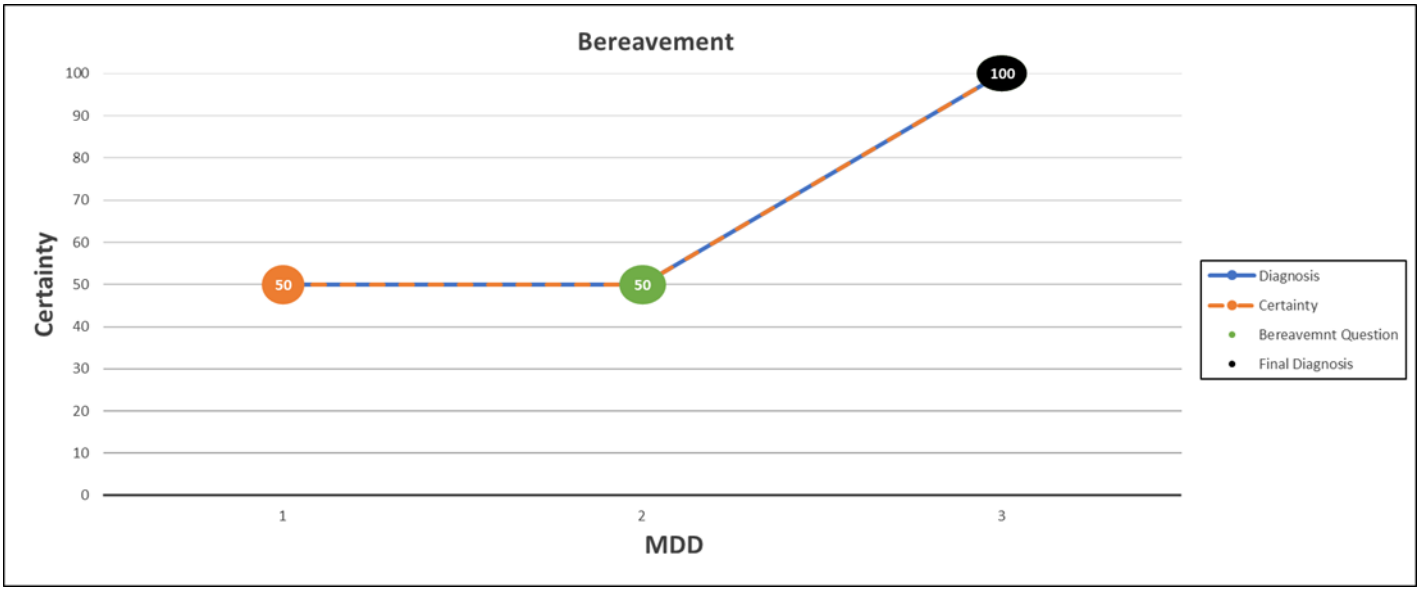

Bereavement
Certainty
MDD
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis


Bereavement
Certainty
MDD
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
100
100
100
80
75
70
65
50
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis


Bereavement
100
90
80
70
60
50
40
30
20
10
0
Certainty
50
25
25
50
25
50
25
1
2
3
4
MDD
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis


Bereavement
Certainty
MDD
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
99
90
90
90
80
50
50
50
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis


Bereavement
Certainty
MDD
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
50
50
50
50
100
95
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis


Bereavement
Certainty
MDD
100
90
80
70
60
50
40
30
20
10
0
1
2
3
50
50
100
Diagnosis
Certainty
Bereavemnt Question
Final Diagnosis