Spontaneous Generations 13(1), pp. 12–30
DOI 10.xx.yy.zz

The Dimensions of Value-Ladenness in Dementia Measurement

Isabella Braga
University of Western Ontario
ibraga@uwo.ca

ABSTRACT

Dementia staging—the tracking of the progression of dementia—has become a crucial exercise in meeting the healthcare needs of an aging global population with unprecedentedly high life expectancies. Dementia-staging instruments—the tests and/or interviews used to assess the severity of dementia—are the foundation of this exercise, and provide scores used to assess patients and patient data for clinical and research ends. Because dementia-staging scores are the baseline through which disease severity is interpreted, the impact of dementia staging is felt at every level of dementia care, treatment, research, and policy. However, an overlooked aspect of dementia staging is the pervasiveness of values in tracking disease progression. From tool creation to implementation, dementia-staging instruments are entrenched in creator and practitioner values. This paper underscores the subjectivity in dementia diagnosis and tracking through an investigation of these creator and practitioner values, applying Douglas's framework of value-guided scientific methodology for the purpose of demonstrating how ethical, social, and cognitive values have influenced four proposed steps of the dementia-staging process. This provides a foundation for future research into dementia as a pluralistic, value-laden phenomenon, paving the way for analyses of the impact of culture on the phenomenal experience of cognitive decline.

Introduction

The pervasiveness of values in dementia staging begs some rational inquiries. We may ask how the creator of a dementia-staging tool defines each stage, how assessment should occur, and/or what each stage indicates about the level of care for each patient. These inquiries are inextricable from the unique epistemic challenges of creating and implementing staging instruments for neurodegenerative disorders: neither the staging instrument creator nor the assessor has ever had the experience of suffering from all stages of dementia, and so value-influenced choices must be made in determining which criteria to assess, how to assess these criteria, and how to interpret these assessments. Because it is not firsthand experience which grounds these staging tools, these staging instruments rely on observations of symptoms attributed to the disease and, akin to many clinical tools, have been historically refined on an ad hoc basis. This paper outlines the values embedded within the creation and utilization of dementia-staging instruments, for

the ultimate purpose of positing that dementia staging should be acknowledged as a subjective exercise, such that dementia-staging instrument scores can be tempered against the reality of their origination.

After a brief overview of the dementia-staging procedure, I will demonstrate how ethical, social, and cognitive values guide each step of the dementia-staging process, from creator intentionality in making the tool, to the continued use of the instrument in practice. To do so, I will apply Douglas's three-pronged analysis of value-guided methodology in the sciences (2014) and Wijsen, Borsboom, and Alexandrova's analysis of values in psychometrics (2022) to popular dementia-staging instruments to illustrate the dimensions in which dementia-staging instruments are infiltrated by values. In order to dissect this spiderweb of values across a plurality of dementia-staging instruments, I will address each step of the dementia-staging procedure in turn, from the definition of the purpose of staging, to the utilization of the dementia-staging instrument in clinical practice or research. As I will explain, the creation and utilization of each dementia-staging instrument adhere to this four-step template.

- Step 1: A purpose is defined.
  I refer to the motivation for the tool. Why and in what capacity are we tracking dementia?
- Step 2: Qualitative evidence is gathered then categorized and/or quantified.
  The administrator of the staging tool conducts interviews and qualitative assessments of the subject and the subject's context. These qualitative assessments are stratified into categories and numbers for expedient reference by clinicians, researchers, and caretakers.
- Step 3: These categories inform the treatment of the subject and/or subject's data within a clinical and/or research context.
  Clinicians, researchers, and caretakers refer to these quantifications as guidance for treatment and further research.
- Step 4: Over time, the instrument is utilized due to its fulfillment of a desired end.
  The instrument is continually utilized due to its perceived success, often either via "validation" by comparison to a gold standard dementia-rating instrument[1]

[1] Though there is significant overlap between the two terms, it is worth noting the definitional distinction between a dementia-rating instrument and a dementia-staging instrument. The technical difference is in how assessments are quantified and conveyed. A dementia-rating instrument quantifies assessments into scores and conveys a numerical rating. A dementia-staging instrument quantifies assessments into levels and conveys a stage of dementia. Notably, these stages may or may not be numerical. Thus, an instrument might be considered both a dementia-rating instrument and a dementia-staging instrument. Furthermore, a dementia rating might be a component of dementia staging; for example, a patient might receive scores in different

such as the Global Deterioration Scale, a less formalized judgment of the instrument as having accurately captured the subject's phenomenology, and/or practitioner satisfaction with the treatment or research outcomes.

Before addressing how values undergird each of these steps, I will first define values and present Douglas's three aspects of value-infiltrated science.

Douglas establishes values as "normative or emotive commitments people hold" which "can also concern a wide variety of things, from commitments to ethical principles, communal patterns of being, or even to qualities one wants to have in one's knowledge about the world" (2014, p. 163). Douglas further divides these values into ethical, social, and cognitive values. Ethical values focus "on the good or the right" (2009, p. 92), weigh potential benefits versus harms, and are concerned with rights, suffering, and risk. Social values "arise from what a particular society values, such as justice, privacy, freedom, social stability or innovation" (2009, p. 92). While these often overlap with ethical values, there are cases when society's values oppose what is ethical: Douglas points to the prejudicial social values reflected in scientific studies of race and IQ as an exemplar of such values as antithetical to ethical values (2009). Finally, cognitive values are "aspects of scientific work that help one think through the evidential and inferential aspects of one's theories and data" (2009, p. 93). Cognitive values are commitments to facilitation of scientists' cognition of their work, and may include simplicity, explanatory power, scope, and predictive precision, with the ultimate "embodiment of cognitive values" being fruitfulness (as in, improvement in productivity in an area of science).

Similarly, Elliott states, "A value is something that is desirable or worthy of pursuit" (2017, p. 12). Elliott draws a line between epistemic values, which "are regarded as contributing to the goal of gaining knowledge" and non-epistemic values, which "do not consistently help us gain knowledge" (2017, p.12). Though Elliott acknowledges that the popular conception of values is typically non-epistemic, Douglas makes the stronger claim that epistemic values—concerned with truth finding and preservation—are not values at all but are, rather, criteria for science. Douglas argues that while values "are aspects of science for which to strive", epistemic values act in a negative capacity, as "[w]ithout meeting these criteria, one does not have acceptable science" (2009, p. 94). In contrast, cognitive values seek to improve the possibilities of scientific research, but do not actually further knowledge acquisition itself.

When confronted with the question of what differentiates natural and social sciences, the general populace might reasonably consider the influence of values to be the dividing line. The natural sciences are objective whereas the social sciences are subjective. The

areas of assessment which will factor into their overall stage. The key difference is that dementia rating produces a score while dementia staging produces a level.

chemical structure of hydrogen peroxide is uninfluenced by the molecule's social context; economics, in contrast, is inextricably bound within its context and the values of said context. Or so goes the common canon.

In her critique of this characterization of the natural sciences, Douglas states that there are three aspects which expose the values embedded within science. First, values influence "the direction and selection of research, the decision about which research to do" (2014, p. 163). Next, values influence the inferences derived from evidence. Finally, values impact the language used to describe findings, coloring "natural" evidence through a social lens. Douglas argues that social sciences have been falsely characterized as distinctive from natural sciences in their "value contamination", while in actuality, "the practice of science, whether natural or social, is shot through with values" (2014, p. 163).

It is indisputable that ethical, social, and cognitive values influence the ways in which science is conducted. The existence of ethics review boards is evidence enough of this. Yet to claim that science itself is value-laden pushes these normative or emotive commitments onto the shape of the measurements themselves, and our investigation becomes less a matter of ethics and more an evaluation of the ontology of this kind of "science". For the ultimate purpose of indicating the subjectivity of dementia-staging procedures, I will now investigate how values have been interwoven with the nature of dementia-staging instrumentation.

Step 1: A purpose is defined

Prior to the delineation of any dementia-staging procedure is the motivating purpose for the procedure. I refer both to the broader issue of dementia staging and to the specific purposes of each dementia-staging instrument.

Whether they are to be utilized in a clinical, research, or other setting, dementia-staging instruments are broadly incited by the same root cause: dementia as a source of harm, and a desire to mitigate the impacts of this harm through treatment and research. Per Douglas, values influence the direction and selection of scientific research—and in the case of dementia staging, Douglas's assertion holds most obviously in the sense that an ethical and social valuation of wellbeing motivates us to track dementia. As dementia staging is one aspect of a larger ethical and societal public health effort to track and thus mitigate the effects of dementia, there is always the underlying motivation to improve the wellbeing of patients and, by extension, communities at large.

To this, one might reasonably object that wellbeing is a value that motivates all tools utilized in clinical and research settings, and, furthermore, that the desire to improve wellbeing motivates science, medicine, and society. One might thus claim that dementia staging is no more shot through with values than an electrocardiogram. However, dementia is a unique affliction in that it irreparably threatens the integrity of one's life

narrative, personal relationships, and capacity to self-actualize in a manner more associated with aging than psychopathy, disorder, or illness, evoking values that may not be implicit in the creation of all clinical tools or psychometric instruments. For example, due to the nature of the disease, the decision to stage dementia is motivated by the social valuation of dignity, cognitive integrity, mobility, rational capacity, social relationships, legacy, and control, among others.

It may seem unusual to single out "values" which appear so obviously fundamental to human flourishing. Yet we must remain conscious that other cultural contexts may prioritize other values in creating dementia-staging tools, motivated by different cultural relationships with aging, memory loss, physical incapacity, and death. If we are to examine the values guiding the creation of dementia-staging instruments, we cannot assume a natural kind of wellbeing. In a theoretical society where narrative continuity is not valued and the demented individual is seen as an aging, evolving manifestation of a pluralistic self, scientists may not be motivated to track the subject's memory as a progression of the disease. Scientists in this theoretical society might instead view the severely demented patient as being of a different, changed natural kind rather than suffering from a progressed version of a disease. The theoretical measurement tool in this theoretical society might only mark the patient's symptoms as evidence of disease if the symptoms manifest in physical incapacitation. The feasibility of such a cultural positionality indicates that we cannot take for granted even the most "obvious" motivating values for dementia-staging instruments. This will be relevant when examining how the quantification of qualitative evidence leads to a characterization of dementia progression as a matter of degree but not kind.

Fortunately for our purposes, in explanations of the methodology of creation of dementia-staging tools, researchers will often indicate values motivating the purpose of the tool. They often, however, present these values under the guise of objective goals. To exemplify this, I will contrast the purposes of the Global Deterioration Scale (GDS) and Clinical Dementia Rating (CDR), the two most ubiquitous dementia-rating scales, often used as the gold standard by which to validate other staging tools. I will additionally address other commonly used instruments, such as the Functional Assessment Staging Tool (FAST), Quick Dementia Rating System (QDRS), and General Practitioner Assessment of Cognition (GPCOG).

In 1977, Reisberg et al. (1982) reported their first use of the Global Deterioration Scale (GDS). Their flagship report on the dementia-staging instrument claimed the measurement tool had been for five years "validated against behavioral, neuroanatomic, and neurophysiologic measures in patients with primary degenerative dementia" (Reisberg et al. 1982, p. 1136). The team cited "significant relationships" between changing GDS scores and anatomic brain changes in patients with primary degenerative

dementia. The GDS was developed for wide utilization in diagnosing primary degenerative dementia, and to adjust for demographic factors while accurately mapping onto empirical observations, such that "behavioral, neuroanatomic, and neurophysiologic investigations [...] indicate a continuum of decline throughout the stages of the Global Deterioration Scale" (Reisberg et al. 1982, p. 1139). Patients are categorized in one of seven stages on the scale, which range from (1) "no cognitive decline" to (7) "very severe cognitive decline", based on a combination of scores on a "Mental Status Questionnaire," "Guild memory test," and "WAIS" (Wechsler Adult Intelligence Scale) (1982). The GDS stands as a well-known staging instrument used as a standard measure of the progression of dementia (Auer & Reisberg, 1996).

From the instrument creators' own report, we can glean a picture of the values laden within the GDS's development. These most obviously include epistemic values, such as the desire for validity—defined by Reisberg et al. as the tool accurately tracking both behavioral and physiological changes in subjects—and the prioritization of empiricism, which, per Douglas, may not be values at all, but, rather, the baseline for good science. More relevant to this argument are Reisberg et al.'s motivations for seeking such epistemic values: strongly non-epistemic ethical values such as a desire for wide utilization and equity, and, interestingly, the social valuation of cognitive decline as a measure of disease progression. While at face the measurement of cognitive decline may seem empirically motivated and thus an epistemic value, it is socially motivated in that it shades the experiences of dementia patients through the lens of loss along a linear scale. Recall the example of a society that may not quantify mental change as decline.

These values of the GDS become stark when contrasted with the Clinical Dementia Rating (CDR), another historical gold standard of dementia rating. Also published in 1982, the CDR "rates cognitive performance in six domains: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care" (Morris et al., 1997, p. 1508). Though the subject is ranked along a scale of cognitive decline much like in the GDS, the CDR segregates these results to domains as metrics of flourishing. This allows for a diverse characterization of the subject's capabilities, unconfined to a singular number.

The form and function of the Clinical Dementia Rating are indicative of the values foundational to the staging instrument's creation. Developed to assess the severity of DAT (dementia of the Alzheimer type), the CDR is known for being lengthy and thus unsuitable for brief dementia screenings; its thorough process is instead optimized for clinical trials and studies (Morris, 1997). The founding purpose of the Clinical Dementia Rating was to provide a multifaceted account of cognitive decline, with the assessor typically not in a caretaking role and thus not subject to efficiency constraints. While the CDR holds similar epistemic priorities to the GDS in that it aims to conduct "good science", its non-epistemic

values are exemplified by its creators' willingness to prioritize considerations of the subject's flourishing over assessor convenience and efficiency. The CDR is time-consuming but allows for a more holistic picture of the subject as a complex individual.

Similarly considerate of multifaceted disease progression is the Functional Assessment Staging Tool (FAST). Also created by Reisberg, primary designer of the GDS, the FAST was created to assess the subject's ability to complete activities of daily living (ADLs) rather than the subject's cognitive decline (Sclan and Reisberg, 1992). With seven stages and sixteen overall subcategories, the FAST was made such that it can be used in tandem with the GDS, with stages tracking onto each other. The used ADLs range from complex tasks such as traveling, meal planning, and paying bills, to the basic ability to hold up one's head. However, in contrast to the CDR, FAST categories are not subdivided into domains of flourishing—like the GDS, these activities are used to grade the subject along a single linear scale. While the FAST takes into consideration the ways in which an individual's life might prosper outside of their intellectual capacities, it does not allow for a plurality in metrics of wellbeing like the CDR.

Desiring to capture the multifaceted nature of the CDR but acknowledging that its time-consuming nature was a barrier to access, Galvin developed the Quick Dementia Rating System (QDRS) (Galvin, 2015). A 10-item questionnaire completed by the informant without a need for a rater, the QDRS offers an overall score of cognitive decline along a 0–30 scale while also ranking the individual within 10 subcategories including decision-making and problem-solving abilities, activities outside the home, behavior and personality changes, and mood. While Galvin's project is motivated by a desire for expediency, the undergirding impetus for the QDRS is not the speed itself but the social value of widespread access and the cognitive value of simplicity. Of course, though the QDRS intends to mirror the categories of the CDR while increasing equity of access, this does not necessarily entail that the QDRS is successful in its emulation, nor does it mean that the QDRS commits to epistemic values such as validity. Regardless, this section only seeks to illustrate the value-ladenness of motivation for the creation of dementia-staging tools. This desire for balance represents a cognitive value in and of itself: the viability of instruments in a clinical context increases the fruitfulness of dementia staging as a scientific enterprise.

All of this is to say that dementia-staging instruments are not created equal. Their creation is propelled not just by social valuations of the harms of dementia, nor just by the ethical desire to rectify these perceived harms. They are angled for particular purposes by their creators, whose values determine which aspects of dementia are evaluated. When viewing dementia-staging scores, it is thus crucial to remember that these tools are from the outset made to fulfill different functions. Imbued with these various causes, these

staging instruments stand as representations of their creators' perceptions of what matters in incapacity.

Step 2: Qualitative evidence is gathered then categorized and/or quantified

I will next explore the values pervasive in how dementia-staging tools gather, categorize and/or quantify qualitative evidence.

There are a variety of methods by which qualitative evidence is gathered, but they are all constructed with the intention of being quantified. It is never enough for a dementia-staging instrument to gather the reported experiences of subjects to inform care. The prioritization of quantification is an aspect of dementia-staging instruments' ends: to act as a shorthand for assessment and treatment of dementia patients. If absolved of the cognitive value of simplicity and, thus, the need for quantification of this qualitative data, dementia-staging instruments could feasibly instead be structured as freeform interviews with general guidelines of care based on assessor intuition.

Of course, practitioners have resource constraints, as well as cognitive values such as concern for consistency in practice. The very nature of dementia-staging instrumentation is thus motivated by a desire for quantification and objectivity, which are dubbed values themselves by Wijsen et al. (2022). Conversely, de-prioritization of qualitative evidence gathering as an inherently valuable enterprise is indicative of dementia-staging instruments' infiltration by the following: social values prioritizing expedience in medical care, ethical values prioritizing resource maximization to increase access to care, and cognitive values prioritizing scientific flourishing, such as simplicity, explanatory power, and consistency in representing the phenomenological experiences of patients.

What cannot be overstated in the gathering of qualitative evidence is the impact of language. Douglas indicates that language embedded in scientific processes influences the findings of science itself (2014). We see this exemplified in both the GDS and CDR in the form of survey questions centered around loss, with degrees of severity attributed to perceived loss, rather than a more open-ended collection of the respondents' lived experiences. This contrasts with tools such as the Clinical Global Impression of Change (CGI-C), which are open-ended and evaluate overall improvement in wellbeing. We can thus see that subjectivity transcends the process of evidence gathering, quantification, and categorization—through language, subjectivity informs how findings are processed and communicated.

Another aspect of note is that qualitative evidence in dementia can vary in kind. These include clinicians' reported experiences with patients, patient self-evaluation, and loved ones' accounts of disease progression. No matter the source, qualitative evidence captures

the progression of dementia by attempting to relay the phenomenological experiences of the subject. Due to this variation in kind, however, the evidence-gathering procedure takes a plurality of forms in staging tools. The CDR, for example, entails an interview touching on every conceivable facet of functioning affected by dementia, with topics including basic personal information, general knowledge, hygiene, recall of major life events, self-reported personal care habits, and memory testing. The GDS and FAST, in contrast, require the assessor to report the status of the subject in a straightforward informant checklist. The QDRS simply requires the subject to fill out a checklist. The GPCOG, interestingly, is brief—consisting only of six questions graded along a correct/incorrect binary—but covers a variety of topics, including recall, time orientation, and clock drawing, followed by an informant interview evaluating the individual only in relation to their cognitive baseline from years prior.

All these variations represent value decisions made by staging-tool creators. For example, wheresoever the staging tool takes the form of a "yes/no" checklist, the assessor has to quickly determine on which side of the binary the patient falls. This indicates the creators' prioritization of cognitive values (such as simplicity) over epistemic values (such as construct, content, and criterion validity). Longer interviews indicate the opposite prioritization. Even prior to the generation of a dementia-staging score, the evidence itself has been colored by the desires of the instrument creators. This coloration occurs when the creators' values influence the language of evidence gathering, method of evidence gathering, and type of evidence gathered. At this step of dementia staging, we see the expression of the tool creators' beliefs about what constitutes effective, valid evidence for phenomenological changes in cognitive decline. From the outset, evidence gathering is an expression of subjective commitments to what qualifies as an indication of dementia.

This value-influence is exacerbated when the evidence is quantified into scores. Following (or in the process of) evidence collection, dementia-staging instruments convert qualitative expressions of lived experiences into quantitative numerals and intervals indicating a stage of dementia progression.

Here, we might hearken back to Douglas's claim that values influence the inferences derived from evidence in science. While this is relevant to dementia-staging instruments in the classical sense of the interpretation of quantitative evidence for treatment purposes (Step 3), it is also relevant to this step, Step 2, in that the conversion of qualitative evidence into quantitative evidence is riddled with value-influenced inferences.

The most obvious indication of this is the value-ladenness inherent in the assessor's role of quantifying the qualitative evidence they observe. Insofar as Step 2 is impossible without a human intermediary, there is an inherent subjectivity in quantification. Simply put, a human must determine if the patient's symptoms warrant a "2" or a "3". However, there are other, less obvious values at play in the process of quantification. Tool creators

must make value judgments when determining categories and methods with which to quantify evidence. The creator of the staging tool must draw inferences about the significance of the qualitative evidence to stratify observations into categories of progression. This stratification relies on value judgments about 1) which indicators qualify as evidence of disease progression and 2) how disease progression will be represented. For example, the tool creator must make judgment calls about the role of comorbidities in evaluating the indicators of disease progression, and also about how dementia itself progresses. If the tool creator does not specify this, then the value judgment falls to the assessor. Consequently, the creator and assessor's chosen quantifications of disease progression will manifest in differing conceptions of the bounds and nature of dementia.

Though this addresses how values might manifest in the process of quantification, Wijsen et al. posit that quantification and objectivity themselves constitute values. While Douglas presents three aspects to indicate values in science in general, Wijsen et al. specifically delineate "the conceptualization of individual differences as having a quantitative (not qualitative) structure, the aim for objective measurement [...] and the preference of utility above truth" as "values that permeate psychometric decision-making" (Wijsen et al. 2022, p. 792). The first two values are especially relevant to Step 2 of dementia staging while the last is especially relevant to Step 4.

As discussed, in dementia staging, the conceptualization of differences as quantitative and the aim for objective measurement are inherent to the structure of dementia-staging instruments. Quantification presents as a value in dementia-staging instruments insofar as it is undergirded by social and cognitive value motivations. Due to a valuation of simplicity, access, and, therefore, speed, clinicians and scientists necessitate shorthand, a means of tracking the progression of dementia without needing to dedicate hours to interpreting personalized accounts of dementia progression, a process that could arguably bottom out in its own form of value-laden arbitrariness but that could offer a more descriptive and nuanced understanding of the subject's needs.

Wijsen et al. corroborate this practical motivation for quantification, stating that "the commitment to being 'practical' is another crucial component of contemporary psychometrics", as "[d]ecisions based on qualitative data are typically hard to trace" (Wijsen et al., 2022, p. 793). In defense of quantification as a value, they reference Porter's characterization of quantification as aiding greater communication about data and as minimizing the need for intimate knowledge of the subject prior to comprehension (Porter, 1995, p. ix, as cited in Wijsen et al., 2022, p. 793). As we will explore later, this practical need for shorthand is also closely related to Wijsen et al.'s assertion that psychometrics values utility above truth—what is sacrificed in accuracy to the experiences

of the subject is gained in the utility of having a tool that allows quick treatment or data categorization.

The attempt to achieve objective measurement in Step 2 of dementia staging is exemplified by the intensive efforts of practitioners to validate dementia-staging instruments against one another and, particularly in the case of the GDS, against neuroanatomic indicators. We can see the valuation of objectivity on a larger scale in popular clinical reviews of dementia tools such as the Dementia Outcomes Measurement Suite (DOMS), which rates tools based on categories such as reliability, validity, consistency, responsiveness, and generalizability (Bentvelzen et al., 2017).

Regarding psychometrics broadly, Wijsen et al. (2022) further argue that scientists' prioritization of quantifiability and objectivity exemplify values in and of themselves, and that these values entail two other values, namely 1) an unfounded valuing of only quantitative knowledge as purely scientific and 2) a commitment to a sense of equality in "that differences between human beings are only of degree and not kind" (2022, p. 792). By the same token, dementia-staging instruments go to great lengths to minimize room for qualitative interpretation (see the DOMS)—for, yes, practical reasons (and the associated values), but also for the broader, more basic, and more socially motivated reason that science and medicine which rely on subjective accounts of qualitative experiences are not valued as science or medicine at all. Additionally, because garnering this kind of scientific legitimacy requires dementia-staging instruments to track subjects' experiences to numbers (which are then placed in intervals), quantification does not allow for variation in kind. Though the qualitative evidence allows for variations in kind, in the process of quantification, these variations in kind are often disregarded, filtered out as being inessential to the characteristics of the disease itself insofar as they do not fit within the quantifiable bounds of the instrument's categories. The valuing of objectivity and quantified knowledge thus tailor dementia-staging instruments—somewhat ironically yielding characterizations of the disease that may not cohere with characterizations by other dementia-staging instruments.

To this end, as there is no naturally occurring standard of what constitutes a valid measurement of the progression of dementia (for example, a hypothetical section of the brain with indicators that unfailingly track a linear conception of dementia progression), those who create and use staging tools must in Step 2 decide in which cases the instrument is functioning, valid, and representative of the qualitative evidence. Subjects who do not cohere to the standards of the instrument may be categorized as not possessing the disease or as having a differing stage of the disease when compared to other staging instruments; and because the instruments themselves set the standard for whom qualifies as presenting at a certain stage of dementia, our justification for the staging instruments becomes evidently circular.

We may look to the GDS and CDR as examples of how quantification of qualitative evidence is undergirded by values.

Because the GDS was created for clinical and caregiving settings, it was not designed as an exhaustive numerical representation of the subject's lived experiences. This is evident in the straightforwardness of the seven stages, which range from "no" to "very severe" cognitive decline" (Reisberg et al., 1982). Placement in a stage is subject to the methodology, perception, and inferences of the assessor, which, as discussed, will color the subject's score. The responses of the subject, no matter how extensive, manifest in a single number along a single interval scale.

Reisberg et al. allege that the GDS bears several advantages over other rating tools, as it is able to "incorporate the clinician's assessment of varying educational, cultural, socioeconomic, and other biases into determining the stage of primary degenerative dementia" (Reisberg et al. 1982, p. 1139). Reisberg et al. additionally posit that the GDS is unique in its ability to delineate "all" stages of the progression of dementia, a sentiment echoed by Reisberg et al. (2002), who claim that the "the GDS and the CDR [...] are generally compatible except that the GDS is more detailed and specific and identifies two stages that the original CDR staging does not" (p. 143), referring to the GDS's final two stages which address progressed incapacity beyond the CDR's description of the most severe stage of incapacity as the subject "requir[ing] much help with personal care" (Hughes et al. 1982, p. 568).

However, Eisdorfer et al. critique the rating system of the GDS. They claim the authors of the GDS "implicitly postulate a single, relatively consistent progression of decline in Alzheimer's disease without empirically verifying this hypothesis", presuming homogeneity in neurodegeneration "whereas a growing literature supports the hypothesis that it may be a heterogeneous disease" (1992, p. 194). Eisdorfer et al. additionally warn that the GDS is hazardous in that, because "the scaled score provides information in a quasimathematical form, it can appear to convey more information about the patient's condition than it actually does" (1992, p. 194). Also particularly vulnerable to these critiques are the FAST, QDRS, and GPCOG, which provide similarly singular "quasimathematical" progressions.

Much like the GDS, the CDR aims to address a similar decline in the subject's cognitive capacities. Yet the CDR offers a more thorough, heterogenous quantification. Following a patient interview, each domain is independently ranked along a five-point scale (0, 0.5, 1, 2, or 3), from 0 meaning no impairment to 3 meaning severe impairment. Each domain's rank is averaged into a composite score indicating the severity of dementia. Perhaps the most obvious difference between the GDS and CDR is the CDR's accounting for heterogeneity through its ranking of cognitive performance in various domains. In contrast to the GDS, the CDR assigns numerals to aspects of the subject's cognitive decline

beyond that which might be immediately relevant to a caretaker—for example, the status of the individual's relationships. To this end, the CDR values the social and relational positioning of the subject in a manner the GDS cannot address due to its linear quantification of the subject's experiences.

Ultimately, it is dire that we acknowledge the subjectivity inherent in dementia-staging scores themselves. Because we have been socialized to value objectivity—and, furthermore, to associate numeral representations with objectivity—we are naturally inclined to accept this quantification process as gospel. However, it is imperative that practitioners and philosophers acknowledge that quantification is both a subjective interpretation of value-laden evidence and a value in and of itself. If this quantification is how we guide and define our understanding of the etiology and ontology of cognitive decline, philosophers and medical practitioners alike must accordingly operate with appropriate caution when allowing these scores to inform their actions.

Step 3: These categories inform the treatment of the subject and/or subject's data within a clinical and/or research context

Following the quantification of qualitative evidence into numbers and intervals, the subject's dementia-staging scores are then used to determine treatment. Some dementia-staging tools offer brief, general recommendations, such as the GPCOG, which instructs "If patient scores 0–3 [out of 6], cognitive impairment is indicated. Conduct standard investigations" (Brodaty et al., 2002, p. 534). Other dementia-staging instruments, such as the GDS, recommend explicit patient care guidelines, such as "Patient can no longer survive in the community without part-time assistance" and "Patient requires continuous care" (Reisberg et al., 1998, p. 15).

As referenced in the discussion of Step 2 of dementia staging, Douglas's claim that values influence the inferences derived from evidence is especially relevant to Step 3. In this step, inferences are made about the best course of action to enhance wellbeing and mitigate the effects of dementia. These inferences are dependent on understandings of the progression of dementia which are in turn dependent on the creator and assessor's interpretation of the scores produced by Step 2. These inferences are also dependent on the creator and assessor's determination of how a practitioner might best respond to an individual at the subject's stage of dementia. Inferences by necessity occur in Step 3 due to the epistemic limitations of clinicians, researchers, assessors, and instrument creators in evaluating the progression of dementia, having never experienced dementia themselves. Creators and wielders of dementia-staging instruments thus make implicit judgment calls when presented with the results of the staging instruments, especially in the early phases of staging tool development and implementation. Tool creators,

assessors, and practitioners are in the fraught position of having to correlate numbers with recommendations for treatment in culturally diverse communities, with subjects who may not respond positively to specific treatment categorizations. Ethical and social value judgments must then be made about the extent to which the "objective" test scores factor into treatment for different individuals.

While these inferences are derived by a desire to enhance wellbeing, they are limited in their capacities to account for pluralism, like the linear intervals of GDS scoring. In assigning a single kind of recommendation per score category, these prescribed responses may have difficulty accounting for different experiences of dementia, a disparity exacerbated by the value-influence which has accumulated in every step prior. As theorists, we must consider that these guidelines are not just utilized in a clinical setting. More broadly, they frame social and political attitudes toward the needs of those experiencing cognitive decline. Yet there are many different treatment guidelines for many different tools which measure different symptoms to different degrees based on different inferences about what different patients want, need, and deserve—in general with a limited capacity to account for how these patients themselves might be different. There is thus a great amount of subjectivity inherent in treatment recommendations.

Step 4: Over time, the instrument is utilized due to its fulfillment of a desired end

This last step of dementia staging has less to do with tool construction itself, and more to do with subjectivity in clinician and researcher selection of dementia-staging instruments. With use, review, and validation against gold-standard tools such as the GDS (which itself purports to be validated against physiological changes), dementia-staging instruments earn reputations for fulfilling specific niches in dementia care. Though practitioners disclose their selection of instruments based on their fulfillment of epistemic values, non-epistemic values also feature heavily in the selection process. For example, tools with a brief administration time (like the GPCOG and QDRS) might be better suited for screening a large, theoretical population of senior citizens known to be affected, say, by lead exposure in their youths. This choice is value-laden insofar as ethical and social values compel us to reach the largest population possible without excessive expenditure of community resources.

Step 4 is thus the embodiment of Wijsen et al.'s assertion that psychometric measures are value-laden in that they preference utility above truth. Wijsen et al. illustrate this point by describing the history of psychometric measures as having evolved from eugenics to data-analysis for the solving of societal problems (2022, pp. 797–799), adding that psychometrics is not the business of truth-seeking but is rather used to answer questions

and derive research dependent on "what is considered valuable in a specific day and age" (2022, p. 799). Because of its contextual dependency, the instrumentalist nature of psychometrics is a moral value (Wijsen et al., 2022). Dementia staging conveys instrumentalist values in that staging tools are selected for explicit clinical and research ends which demand efficiency, uniformity, and indicators specific to certain drugs and/or treatment goals. We see this in the practice of dementia-staging instrument selection in that ethical and social values compel practitioners to use whichever instrument can fulfill clinical and research ends most efficiently and without exorbitant resource usage.

The problem with value-ladenness in tool selection is that it creates gaps in convention for dementia treatment, disadvantaging patients navigating between centers of treatment and testing, even within healthcare systems. In short, different centers use different tools. This lack of convention may seem to contradict the undergirding value of dementia staging: to improve wellbeing by mitigating the harms of dementia progression. Nearly two decades ago, the European Alzheimer's Disease Consortium noted this discrepancy, and expressed "a need for a consensus in the use of assessment tools for dementia in Alzheimer's centres" due to the "large differences in the tools [...] to evaluate patients with dementia in Europe" (Ramírez Díaz et al., 2005, p. 744). The referenced tools included 20 assessment instruments, including the CDR and two versions of the GDS. The Consortium alleged that large differences between the tools and gaps in understanding between professionals conferred decreased health outcomes for patients. The undergirding issue is that medical practitioners, researchers, and other utilizers of these tools fall into their own conventions about the efficacy of each instrument based on value judgments.

Another dimension of this is that some tools are specifically developed for certain ends and so continue to be utilized in narrow capacities, limiting the opportunity for knowledge transfer. For example, the FDA Clinicians' Interview-Based Impression of Change is largely used in drug trials, while the GDS was created by and for caregivers of dementia patients. Though with the advent of digital communication technology there has been progress in knowledge transfer regarding dementia-staging instruments—minimally, because the internet allows practitioners to research every available tool on a whim—there still is not a central convention for what constitutes a successful dementia-staging instrument. Even though some tools, such as the CDR, are popularly utilized due to successful international adaptation and standardization (Lotfi et al., 2015), these tools remain laden with value tradeoffs that do not lend themselves to ubiquity. For example, the CDR heavily prioritizes epistemic values (Douglas's "good science") over ethical values such as resource expenditure and equity of access.

One might argue that this lack of convention is beneficial. In the DOMS, Bentvelzen et al. describe the Rowland Universal Assessment Scale (RUDAS), "designed for people

from non-English speaking backgrounds living in English speaking countries" (2017, p. 829). The RUDAS boasts excellent generalizability due to its many translations. Bentvelzen et al. also describe the KICA-Cog, a dementia detection tool designed "for indigenous Australians living in rural and remote communities", which boasts high diagnostic efficiency and value for other indigenous populations (2017, p. 830). The niche applicability of these tests allows for greater access to dementia care for marginalized populations. The creation of a central dementia-staging convention might harm populations better served by more tailored tests, as a convention could potentially be constructed on a majority consensus of values which fails to accommodate the values of these marginalized groups. Even if the convention were issued with caveats accommodating these niche tools, there remains the concern that a convention establishes a standard of what is valid, correct, and natural, with these niche tools representing offshoot or abnormal instruments of dementia classification, whose results might then not be taken as seriously by the psychiatric community.

Discrepancies in convention further motivate the theorist to inquire as to how these dementia-staging instruments could possibly be validated. If many subsets of the global population seem to necessitate unique frameworks of staging dementia, based on local cultural frameworks of the interplay between cognitive decline, memory loss, neurodegeneration, and wellbeing, it seems to follow that understandings of dementia progression are bald-facedly value-laden and culturally reliant.

Even when convention is not a point of debate, dementia-staging instrument selection and use continues to be motivated by cognitive values. This is exemplified by ongoing tensions over the utilization of even the most popular tools. For example, Bentvelzen et al. attest that the GDS is highly responsive and can be valued for its brevity and detail, even at the expense of sensitivity and comprehensiveness (2017, p. 830), while Eisdorfer et al. worry that the GDS can lack construct validity in that it was built on assumptions about the "occurrence, rate of change, and interrelationships of cognitive impairment, behavioral dysfunction, impairment in activities of daily living, and psychiatric symptoms", rather than deferring to statistical data about the indicators of cognitive decline (1992, p. 193). In sum, even the gold standard dementia-staging instrument is up for value-laden critique—and the values foundational to the staging instrument's creation may not always cohere with the values implicit in its continued utilization.

Conclusion

In this paper, I have sought to indicate values embedded in the creation and utilization of dementia-staging instruments. Through an application of Douglas's framework of value-guided scientific methodology, I have demonstrated how ethical, social, and cognitive values have influenced four proposed steps of the dementia-staging process.

Due to the growing number of dementia patients in aging populations across the world, dementia staging will only increase in importance. Dementia staging informs clinical practice, research, and policymaking. Given the importance of dementia staging in tracking dementia progression, this paper sought to indicate the pervasive dimensions in which this process is subjective due to its reliance on the values of tool creators and practitioners.

In demonstrating how dementia-staging tools are value-laden, I have illustrated how dementia staging is neither culturally static nor free of subjective influence. On the bioethical front, embracing the subjective nature of dementia staging would constitute a step toward maximizing beneficence for patients in dementia care and research. If practitioners were to operate with a wariness for the malleability of dementia-staging scores, patient outcomes might improve. Future bioethics research might utilize these dimensions of value-ladenness to evaluate how we might create a value-driven dementia-staging instrument that could better serve as a convention across culturally pluralistic populations like those of the United States and Canada. A value-driven dementia-staging instrument might entail the primary delineation of certain values as desiderata, with the dementia-staging instrument then being constructed specifically to fulfill those desired values. Importantly, the construction of a value-driven dementia-staging instrument necessitates prior analysis of which dementia-staging procedures correspond to the fulfillment of certain values. Thus, knowledge of the dimensions of value-ladenness in dementia measurement is significant.

Though acknowledgment of value influences is just the first step in this research, it is especially imperative to do so at this moment in medical and scientific history. As societies across the globe engage in renewed debates about medical aid in dying for sufferers of advanced neurodegeneration, dementia staging will emerge as a crucial component of clinical procedure to verify capacity of consent to die. Evaluating the subjective influences of dementia-staging instruments will better equip us to anticipate the ethical dimensions at the intersection of medical and legal debates about the rights of dementia patients. Careful consideration of these issues is especially important when treating dementia patients, who are especially vulnerable due to the loss of cognitive functioning inherent to the disease's progression. It is therefore my hope that this deconstruction of values in dementia staging will motivate efforts to enhance patient welfare.

Acknowledgments

I would like to thank my colleagues, peers, and good friends Alexander Stooshinoff and Samuel Carlsson Tjernström for their invaluable comments on an early draft of this paper. Furthermore, I would like to thank Dr. Eran Tal at McGill University for leading the incredible seminar on the philosophy of measurement in which this paper was written.

Dr. Tal's mentorship, expertise, and commentary were instrumental to this paper's creation.

References

Auer, S. R., & Reisberg, B. (1996). Reliability of the Modified Ordinal Scales of Psychological Development: A cognitive assessment battery for severe dementia. *International Psychogeriatrics*, 8(2), 225–231.

Bentvelzen, A., Aerts, L., Seeher, K., Wesson, J., & Brodaty, H. (2017). A comprehensive review of the quality and feasibility of dementia assessment measures: The dementia outcomes measurement suite. *Journal of the American Medical Directors Association*, 18(10), 826–837.

Brodaty, H., Pond, D., Kemp, N. M., Luscombe, G., Harding, L., Berman, K., & Huppert, F. A. (2002). The GPCOG: a new screening test for dementia designed for general practice. *Journal of the American Geriatrics Society*, 50(3), 530-534.

Douglas, H. E. (2009). *Science, Policy, and the Value-Free Ideal*. University of Pittsburgh Press.

Douglas, H. (2014). Values in Social Science. In N. Cartwright & E. Montuschi, *Philosophy of Social Science: A New Introduction* (pp. 162-182). Oxford University Press, USA.

Eisdorfer, C., Cohen, D., Paveza, G. J., Ashford, J. W., Luchins, D. J., Gorelick, P. B., Hirschman, R. S., Freels, S. A., Levy, P. S., & Semla, T. P. (1992). An empirical evaluation of the Global Deterioration Scale for staging Alzheimer's disease. *The American Journal of Psychiatry*, 149(2), 190–194. https://doi.org/10.1176/ajp.149.2.190

Elliott, K. C. (2017). *A Tapestry of Values: An Introduction to Values in Science*. Oxford University Press.

Galvin, J. E. (2015). The Quick Dementia Rating System (QDRS): a rapid dementia staging tool. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 1(2), 249-259.

Hughes, C. P., Berg, L., Danziger, W., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *The British Journal of Psychiatry*, 140(6), 566–572.

Lotfi, M. S., Tagharrobi, Z., Sharifi, K., & Abolhasani, J. (2015). Diagnostic accuracy of Persian version of Clinical Dementia Rating (P-CDR) for early dementia detection in the elderly. *Journal of Rafsanjan University of Medical Sciences*, 14(4), 283–298.

Morris, J. C. (1997). Clinical dementia rating: A reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *International Psychogeriatrics*, 9(S1), 173–176.

Morris, J. C., Ernesto, C., Schafer, K., Coats, M., Leon, S., Sano, M., Thal, L. J., & Woodbury, P. (1997). Clinical dementia rating training and reliability in multicenter studies: The Alzheimer's Disease Cooperative Study experience. *Neurology*, 48(6), 1508–1510.

Ramírez Diaz, S.P., Gil Gregório, P., Manuel Ribera Casado, J., Reynish, E., Jean Ousset, P., Vellas, B., & Salmon, E. (2005). The need for a consensus in the use of assessment tools for Alzheimer's disease: The Feasibility Study (assessment tools for dementia in Alzheimer Centres across Europe), a European Alzheimer's Disease Consortium's (EADC) survey. *International Journal of Geriatric Psychiatry*, 20(8), 744–748.

Reisberg, B. (2007). Global measures: Utility in defining and measuring treatment response in dementia. *International Psychogeriatrics*, 19(3), 421–456.

Reisberg, B., Franssen, E. H., Souren, L. E. M., Auer, S., & Kenowsky, S. (1998). Progression of Alzheimer's disease: Variability and consistency: ontogenic models, their applicability and relevance. In H.-J. Gertz & Th. Arendt (Eds.), *Alzheimer's Disease—From Basic Research to Clinical Applications*, 9–20. Springer.

Reisberg, B., Auer, S. R., Monteiro, I., Boksay, I., & Sclan, S. G. (1996). Behavioral disturbances of dementia: An overview of phenomenology and methodologic concerns. *International Psychogeriatrics*, 8(S2), 169–182.

Reisberg, B., Ferris, S. H., De Leon, M. J., & Crook, T. (1982). The Global Deterioration Scale for assessment of primary degenerative dementia. *The American Journal of Psychiatry*, 139(9), 1136–1139.

Reisberg, B., Ferris, S. H., Kluger, A., Franssen, E., Wegiel, J., & De Leon, M. J. (2008). Mild cognitive impairment (MCI): A historical perspective. *International Psychogeriatrics*, 20(1), 18–31.

Sclan, S. G., & Reisberg, B. (1992). Functional assessment staging (FAST) in Alzheimer's disease: reliability, validity, and ordinality. *International Psychogeriatrics*, 4(3), 55-69.

Wijsen, L. D., Borsboom, D., & Alexandrova, A. (2022). Values in psychometrics. *Perspectives on Psychological Science*, 17(3), 788–804.

World Health Organization. (2023, March 15). Dementia. https://www.who.int/news-room/fact-sheets/detail/dementia#:~:text=Key%20facts,injuries%20that%20affect%20the%20brain.