Weaving necessity

> Text comes from the Latin *texere,* meaning to weave, and context derives from *contexere,* meaning to weave together or to weave with. (Janssen, 1985, cited in Asdal, 2012)

> [Y]ou should not see that [appraisal process] as too fabricated, I think, because those are actually quite organic, er, organic elements, that discussion, right? (Institute employee 4, interview)

> We have thought too much in recent years about the science of thinking and not enough about the art. (Jacobs, 2017, p. 16)

HOW TO ESTABLISH NECESSITY COLLECTIVELY?

This dissertation was tasked with the quest for a new, more apt, and more usable operationalisation of the necessity criterion in Dutch health care coverage decisions as the process of refining and tightening the establishment of necessity on a collective level was considered to have lagged compared to other criteria. Though not every country employs a formalised necessity criterion, many do utilise qualitative considerations that concern the necessity questionin their deliberations. This question asks: do we think providing this health care technology is necessary on a collective level; or as phrased more specifically in Dutch policy documents and reports, *do we think a claim on the collective solidarity is justified?*

In the Netherlands, decisions concerning the contents of the basic benefits basket are taken by the Minister of Health based on advice formulated by the Dutch National Health Care Institute (in Dutch: *Zorginstituut Nederland*, ZIN, here: the Institute). Many people input into such advice: professionals accumulating and arranging scientific data on effectiveness, cost-effectiveness, and individual severity of illness, building on the work done by many others who have gathered these data; stakeholders who attend and contribute to different meetings that are part of the decision-making process; committees who deliberate and weigh; professionals who write extensive and summative reports. During this multi-cogged process, four criteria are employed so that the final decision may meet all four. Effectiveness and cost-effectiveness of the health care technology in question are first and second, feasibility third, and necessity is the fourth criterion. Using necessity as a criterion has never been as straightforward as the others, and this lack of perceived clarity led to this dissertation.

Over the course of this project (2015-2020), the Institute appears to have moved its position on the necessity criterion and the use of criteria generally. A presentation by an appraisal committee member on an away day (December 2016, field notes #161202) suggested that necessity should be considered an outcome of the whole process rather than input in the form of a criterion. This was followed, a year later, by the publication of the report *Package advice in Practice* (in Dutch: *Pakketadvies in de Praktijk – Wikken en Wegen voor een Rechtvaardig Pakket*) (Zorginstituut Nederland, 2017). Here, the word 'necessity criterion' has largely disappeared in favour of a combination of

severity of illness and own risk considerations. The emphasis of this report lies on argumentations now specifically serving as input for *wikken en wegen* (a Dutch idiom perhaps best translated as a mixture between "deliberation", "weighing", and "hemming and hawing" – the report is now usually referred to at the Institute as the *Wikken en Wegen* report). The report argues that for the acceptance of package advice, it is important that all relevant arguments are visible, that it is clear how they are weighed together, and it is shown which arguments were decisive in the advice and why. "After all," the report states, "new argumentations may always present themselves" (Zorginstituut Nederland, 2017, p. 5). Without concluding that this process of *wikken en wegen* offers a one-to-one replacement for the necessity criterion as previously operationalised, it is important to note the difference in emphasis with the previous *Package management in Practice* reports (College voor Zorgverzekeringen, 2006, 2009; Zorginstituut Nederland, 2013).

My approach to operationalising necessity was influenced by a discrepancy I encountered relatively early on in my research. This discrepancy existed between, on the one hand, the difficulties and critiques displayed by the many scholarly and policy reports that had considered necessity previously and, on the other hand, the relative ease with which Institute employees pinpointed decisions in which necessity was of particular influence. During the early days of my field work, these employees provided a nearly continuous stream of examples: cosmetic surgery, health care aids like walking aids with wheels, maternity care, paracetamol, and vitamin D tablets, physiotherapy, smoking cessation therapies, Viagra, etc., etc. They even often classified these as 'typical necessity questions'. This sensitised me to the difference between argumentations (as present in decision-making practice) and criteria, which seemed mostly visible as such in the decision documents but were stipulated by policy reports to be used in practice. One of the first main questions this project raised in me was: what *is* a criterion exactly? Moreover, I wondered about the fact that many operationalisations had been substantive in terms of descriptions or even checklists (Commissie-Dunning, 1991; Niëns, 2014), even though the process and place of using these argumentations seemed so important in the policy reports (as outlined in the Introduction to this thesis).

From the very beginning – but encouraged by the notable developments in the *Wikken en Wegen* report – I chose to look at the practices of using necessity as a criterion in the form of argumentations, rather than relying on a primarily theoretical approach. The reason was that there seemed to be most clarity in daily practice. The main research question of this dissertation thus became: *how is the necessity criterion used in practice?* The first step entailed the identification of argumentations pertaining to necessity. I purposely widened this first step, the identification of potential argumentations, to scholarly literatures rather than policy reports only. Doing so, I aimed to retrieve a wider variety of argumentations, not only those considered valid enough to become part of a decision justification or rationale. The following steps entailed observing how these argumentations are used in decision-making practice. I did this through a cross-country comparison of practices and decision documents, hoping to gain insight through mirroring 'our' Dutch practices to others', and observations at the appraisal committee meetings enriched by interviews with appraisal committee

members and Institute employees. For the latter two elements, I chose a case approach to gain insight into patterns of social behaviour specifically (Creswell & Poth, 2017; Ragin, 2004).

This discussion starts with the answer to the main research question: how is the necessity criterion used in practice? I will subsequently highlight how using the necessity criterion achieves a societal weighing, conceptualising this societal weighing as *contextualisation practices* while positioning it in relevant Health Services Research (HSR) and Science and Technology Studies (STS) literatures. I will then provide policy recommendations and a reflection on them before ending with the general limitations of this research and a reflection on my research position and role.

ON USING THE NECESSITY CRITERION IN PRACTICE

The unit or form the necessity criterion takes in practice is *argumentations* in favour of, or against, coverage of a health care technology. Argumentations are the explicated reasons pertaining to coverage, generally given on paper or in discussion by anyone reasoning about the potential coverage status of a technology. These necessity argumentations are first and foremost diverse: they may range from scientific calculations on individual severity of illness to more practical considerations such as whether similar treatments are covered or not; from definitions of (non-)illness and what may, or may not, be considered the normal experience to the wider societal impact of coverage (see Chapter 1). Some necessity argumentations are relatively idiosyncratic (as in, specific to the health care technology under consideration) and others more permanent. An example of the first would be the argument that young people should be enabled to step into adulthood with a full set of teeth as this is important for eating apples and daily communication. A (to insiders highly familiar) example of the latter, more permanent type is that as an insurance-based system, to reimburse relatively cheap medicines through health care insurance is considered more expensive than not to do so, meaning that the rule of thumb of 'cheaper than €100 means no coverage' may be applied. In this variability, necessity differs from criteria such as effectiveness and cost-effectiveness, as these generally work with calculations. These may also be, and generally are, under discussion, but the diversity of argumentations and their largely qualitative nature are the first characteristics that make necessity stand out.

The selection of necessity argumentations differs per health care technology under consideration and is crucial to the use of the necessity criterion in practice. As Chapter 1 shows, the reason is that not all necessity argumentations are considered equally valid for the decision on every health care technology. This makes earlier efforts to operationalise necessity in the form of a checklist (Niëns, 2014) relatively difficult to use, as it indeed turned out to be (interviews appraisal committee member 5, Institute employee 6, March 2015). It also sets necessity again in stark contrast to more classic coverage decision criteria such as effectiveness and cost-effectiveness (cf. Stolk & Poley, 2005). Others have noted, over fifteen years ago, that the use of what they termed considerations pertaining to necessity was inconsistent, which made it a "problematic notion" (Hoedemaekers

& Oortwijn, 2003). This dissertation fully supports the conclusion of inconsistency in terms of necessity argumentation use, though expressly not the valuation thereof.

When comparing the use of necessity argumentations across countries, it becomes clear that not every country has formalised necessity in a criterion to the same degree. This does not mean, however, that necessity argumentations are absent elsewhere. In fact, the use of necessity argumentations is widespread, at least in the Western world but also not absent from other middle- and high-income countries such as Argentina, Israel, Japan, and Korea (Chapter 1). The reason it may perhaps appear a quintessentially Dutch notion is, at least in part, the strong tradition of explicating and operationalising these argumentations, as evidenced by the policy reports described in the Introduction but also by scholarly work in this area (Hoedemaekers & Oortwijn, 2003; Stolk et al., 2002). Chapter 2 delves in further to show how in Belgium, England, and Germany, necessity argumentations are both recognised and used by decision makers. The Netherlands has both a comparatively high level of formalisation of necessity and a broad set of necessity argumentation usage. This holds generally across the decisions on the four health care technologies studied but is especially visible in the individual cost consideration. This consideration indicates a lower necessity of coverage of health care technologies with a low price and is unique to the Netherlands in this dataset. The high level of explicated and formalised necessity argumentations confirms Dutch coverage practice as an especially fruitful research site.

From the cross-country comparison, it also becomes clear that the combination or *pattern* of argumentation use is of high importance, as similar patterns of argumentation use lead to similar health care coverage decisions in the countries studied. In Chapter 2, I highlight similarities between argumentation patterns in specific decisions, such as the English and Dutch decisions on nivolumab, benzodiazepines, and smoking cessation therapies. Chapter 1 resonates this conclusion as it not only demonstrates that some argumentations are used primarily in favour of, and others primarily against, coverage but also indicates similarities between the argumentation patterns for cancer drugs and orphan drugs, which are very different from the argumentation patterns for cosmetic surgery, Viagra, infertility treatments, obesity treatments, and smoking cessation therapies. This research thus underlines the previously-noted but not well-studied importance of clustering or patterning of argumentations (Lehoux et al., 2010; Martin et al., 2001; Singer et al., 2000). Notably, these argumentation patterns could well coincide with potential different decision networks with varying robustness (Chapter 4). This indicates that exchange between countries on more qualitative aspects of health care coverage decisions, in addition to ongoing efforts on quantitative aspects, may well be a fruitful endeavour.

Necessity argumentations are used 'around the table', that is, in a *deliberative* setting, in all four countries studied in Chapter 2. In the Netherlands, this primarily takes the shape of the appraisal phase of coverage decisions. The appraisal phase generally follows the assessment phase, in which the relevant scientific input is determined, but it may or may not be a distinct moment in time and space (Oliver et al., 2004; Patera & Wild, 2014; Walley, 2007). The goal of the Dutch appraisal, which is indeed separated from the assessment, is to provide a *societal weighing* of this input, similar to the

former social value judgements in England (Culyer & Rawlins, 2004; NICE, 2008; Shah et al., 2013; Zorginstituut Nederland, 2017). This appraisal is achieved by the appraisal committee (in Dutch: *Adviescommissie Pakket*, ACP), a committee of eight to ten experts with professional backgrounds such as pharmaco-economics, health care ethics, and patient sciences (art. 14, Zorginstituut Nederland, 2016).

To describe how necessity is constructed and what constitutes societal weighing expertise therein, I the appraisal committee meeting in detail in Chapters 3 and 4. From this, I conclude that societal weighing expertise has two primary facets in this dataset. First, it takes the form of receiving input and allowing this input (in the form of explicit criteria presented by Institute employees and contributions by patient (representative)s) to steer and challenge the process to increase the quality of the decision. This dynamic of perceived increased quality of decisions due to the presence and contribution of patient (representative)s, I would suggest, has two primary explanations. First, unlike other criteria noted above, necessity is the type of consideration that may more easily also be contributed by laypersons and those with experiential expertise (Chapter 1). In this way, the variety and number of argumentation types may be increased and the resulting decision potentially strengthened (Chapter 4). The committee members themselves expressed the second explanation, namely how, primarily for expensive medicines with a relatively unfavourable cost-effectiveness ratio, the committee was forced to consider matters extra carefully and formulate their negative coverage decision advice especially diligently, when faced with patients and/or their representatives (Chapter 3).

The second facet of societal weighing expertise is the deliberation and the formulation of the advised decision and recommendations. This deliberation is where the pattern of necessity argumentations that is to become part of the final coverage decision is established. The appraisal committee does so by bringing in new argumentations, derived from many sources ranging from scientific reports to newspapers and beyond, and *weaving the argumentations together*. This weaving serves as a metaphor for the verbal combining of different case-specific and actively-integrated argumentations considered pertinent to the case at hand (Chapter 2). Different argumentations may be woven together in different ways, of which I conceptualise three in Chapter 4. First is the articulation of links between argumentations, whereby different argumentations are brought together as both relevant to a positive or negative decision. Examples include linking the coverage of maternity care not to preparing *beschuit met muisjes* but to detecting risky situations. Using black-boxed links is the second. Black-boxed links are previously-formulated connections, of which the negative decision based on an unfavourable cost-effectiveness ratio appears the most common, described by some as "simply stamping the file" (appraisal committee member 6, personal communications). The third and final way of combining argumentations is broadening the scope of networks, i.e., including previously unconnected matters to a decision under construction. This includes comparing it to a previous decision, in the sense that 'entering adulthood with good front teeth' should be covered just as 'entering adulthood without an unwanted pregnancy' is. Broadening the scope of networks also comprises the formulation of specific recommendations. These recommendations

may be aimed at actors as diverse as the Minister of Health and treatment expertise centres abroad. The intended end-product, an advised decision plus recommendations, I conceptualise as a robust decision, meaning able to withstand pressure 'out there'.

In sum, Chapters 1 and 2 describe how the necessity question – is a claim on the collective solidarity justified? – is answered in practice by using argumentations. These argumentations are numerous and diverse, and their perceived validity differs per decision. In some countries, like the Netherlands, these argumentations are formalised to a high degree into a criterion, whereas in other countries these argumentations are used without such a high level of formalisation. In the Netherlands, a high level of formalisation does coincide with a high diversity of argumentation types. Moreover, necessity argumentations are used in patterns, with similar patterns leading to similar decisions across countries. Finally, which argumentation types to use is determined in deliberation, 'around the table', which generally occurs in the Netherlands in the appraisal phase with the explicit purpose of achieving a societal weighing. Zooming in on the Dutch appraisal, Chapters 3 and 4 describe what such a societal weighing comprises: allowing explicit criteria and patient (representative)s' contributions to steer and challenge the process, and subsequently, bringing in new argumentations, combining them in different ways, (namely through linking argumentations, using black-boxed links, and broadening the scope of decisions), making the decision and formulating recommendations. This shows how argumentations and other decision elements are combined into a robust decision.

SOCIETAL WEIGHING AS CONTEXTUALISATION PRACTICES

The societal weighing that happens in appraisal, identified under the previous heading as the primary place of use of necessity argumentations, serves to set the decision-making process and outcome in context, to contextualise it. I will show in this section that societal weighing entails first, 'bringing the outside in', and second, placing the decision 'back into society'. I will then set these findings in the wider Health Services Research and Science and Technology Studies literatures relevant to coverage decision-making practice, explicating my contribution by describing societal weighing as *contextualisation practices*.

Bringing the outside in

In the process of societal weighing, 'the outside' is actively brought in. This happens in two primary ways, first, through the patients, patient representatives, and/or other stakeholders that are present in the room (or have been present in the scoping session), and second, through argumentations brought in by the appraisal committee members. I will discuss these in turn.

First, the patient, patient representatives, and/or other stakeholders strongly impact the processes of societal weighing as they represent the outside. They arrive in the scoping or appraisal committee meeting literally from the outside: they use the visitors' entrance, sit in the audience section (see

Chapter 3), and are here for this decision only. In this dataset on appraisals, these groups were almost exclusively visible in the eculizumab case, where a mother described the roller coaster their family had experienced since their daughter's diagnosis with aHUS. However, the effect of patient (representative)s on the deliberations were a primary topic in the interviews with the committee members. Their contributions are considered to steer or guide the deliberative process in terms of challenging the committee to substantiate their decision well. Moreover, other stakeholders contributed in several cases at other moments, such as pharmacists in the paracetamol-vitamin D case drawing attention to vulnerable groups, and the association of dentists who set the front teeth replacement therapy decision on the agenda and inputted later on as well.

The second way in which the outside is brought in is through the committee members themselves, who derive argumentations from a wide variety of 'outside sources'. These sources include but are not limited to newspapers (as in the case of maternity care, where reference was made to *beschuit met muisjes smeren*) and previous decisions (as in the cases of front teeth replacement therapy, where it was linked to the coverage status of contraceptives for under-18s, and eculizumab, with the earlier decision for PNH patients). Naturally, argumentations may also come from the contributions noted above, as the pharmacists who contributed during the scoping session for paracetamol-vitamin D tablets, whose argumentations were repeated by an appraisal committee member.vv The committee, then, may add an argumentation from the outside to a (collective) decision network and in this way ascertain and, if successful, ensure its validity in this case.

> *In societal weighing, the outside is actively brought in through the presence and contributions of patients, their representatives, and/or other stakeholders, which steer or guide the deliberations, and through the committee members themselves as they bring in new argumentations from outside sources and add them to decisions.*

Placing the decision back outside

Placing the decision back outside is likewise achieved through two interlinked ways. The first element is that the committee seeks to make a *robust* decision, conceptualised by Rip as being able to withstand pressure in particular outside settings, achieved through careful bringing together of argumentations into a solid justification or rationale. The second element is the addition of recommendations.

The first element, making a robust decision, entails for the committee to bring argumentations and other decision elements together with the societal context in mind, which means here to make reasonings explicit. Specifically, the explication of argumentation types for reasons of providing a justification or rationale is important here. Such explication into networks of argumentation types cannot be performed randomly, seen first in the fact that many argumentation types appear to be used for either positive or negative coverage decisions. They tend to come in patterns, as shown by the similarities between several argumentation patterns in the cross-country comparison, such as the English and Dutch decisions on nivolumab and smoking cessation therapies. It is also shown

by the ways alternative decision networks are explicated, such as the appraisal committee member who professed to be able understand that *beschuit met muisjes smeren* should not fall under the collective solidarity, as well as the committee member who posited that safeguarding the disposable income of vulnerable groups is not the prime responsibility of the basic benefits basket. Both were done to benefit the reason-giving: to provide a justification or rationale displaying that these matters had been thought about, though dismissed.

The second way in which the committee places the decision back outside is through recommendations to other stakeholders. These recommendations comprise, in this dataset, continued work on indication criteria (e.g., the maternity care case, but also the English and Dutch decisions on benzodiazepines and the German decision on walking aids with wheels), price negotiations (nivolumab case), but also cooperation with other research centres (eculizumab case) and for the Minister to "to think carefully about the prescription rule" (paracetamol-vitamin D case). Importantly, giving recommendations is not part of the formal remit of this committee; they actively branch out when they give such recommendations. These practices also serve to make the decision more robust, even specifying the outside settings in which the decision is to have an effect.

> *Societal weighing thus secondly entails placing the decision back outside through making careful justifications or rationales for decisions and making them explicit to benefit those outside, and through formulating recommendations for specific outside settings.*

Contextualisation practices

In this section, I would like to concretise how conceptualising societal weighing as contextualisation practices builds on, and contributes to, the HSR and STS literatures.

First, as outlined in the Introduction, scholars in HSR have described committees like the Dutch appraisal committee as taking decisions flexibly, humanely, with sensitivity towards emotions and preferences of the recipient(s) of the health technology under consideration as well as those making the decision (Hughes & Light, 2002; Mechanic, 1997; Russell & Greenhalgh, 2014). This is combined with adhering to decision criteria and procedures. These committees are especially good at doing both (Hughes & Doheny, 2011; Jenkings & Barber, 2004; Russell, 2017). Russell has designated this combination as experts displaying *pragmatic rationality* (Russell, 2017, following a.o. Aristotle; Russell & Greenhalgh, 2014).

HSR's conceptualisation of pragmatic rationality as human (responsive to patients and others) and rational (responsive to criteria and procedures), this dissertation shows, has a distinct outworking when it comes to societal weighing. In societal weighing, these patients, their representatives, and/or stakeholders are representing the societal *context* in the deliberative setting. During the deliberations, the appraisal committee feels they need to keep their distance (cf. Moreira, 2012) but that the presence of patient (representative)s does give them "handles" for substantiating their position. In this sense, this outside that is brought in by patient (representative)s being present, this context, follows STS scholar Asdal as it "conditions or enables a specific utterance to happen"

(Asdal, 2012, p. 388). As such, it is primarily the people that are present that have this type of impact (cf. Wallenburg et al., 2019), and this is perceived to increase the quality of the decision outcome. Notably, this links the presence of the patients, their representatives, and/or other stakeholders, who represent the outside, directly to the quality of the decision justification or rationale, which is for the benefit of the outside. It is a double form of contextualisation. Deliberations, decisions, and justifications or rationales being shaped by these actors in this way I thus denote as contextualisation practices.

Second, STS has classically described pragmatic rationality as taking into account circumstances relevant to that unique situation in order to achieve robust outcomes in particular settings (Rip, 1985, 1992). Rip specifies that pragmatic rationality is crucial in the absence of scientific and socio-political consensus, in line with more recent studies that highlight such rationality in situations of high uncertainty (Calnan et al., 2017; Moreira, 2011). Within STS, one-off controversy has been deemed an important area of study. The reasoning is that many implicit argumentations become explicated in confrontational settings especially (Callon et al., 2009; Moreira, 2011, 2013).

As noted in Chapter 4, this study shows that such explication also happens, at least in part, in less controversial decisions. Moreover, it is not just explication of argumentations; the committee carefully brings them together by bringing the outside in (deriving from newspapers and the like) and adding recommendations to stakeholders, who are in daily life positioned outside. Argumentations to be factored into the decision are fragile during the process and needs to be solidified or linked in some way, which happens in the decision-making situation (Callon et al., 2009; Nowotny, 2003; Rip, 1992). Part of the contextualisation of the decision thus involves deciding which argumentation types are, and which are not, taken along in the final decision: around the table is where the active *integration* of the outside argumentations happens. Moreover, it is where the recommendations are formulated, which is a specification of actively making a decision robust *in certain outside contexts*. Recommendations specify who should do what to 'make this work', to heighten the quality of care. As indicated earlier, previous work has not engaged much with giving recommendations, and this is therefore an important nuance: recommendations make a decision robust in certain contexts. Here we thus also see a double contextualisation movement: active integration of outside argumentations, combined with active formulation of recommendations to benefit outside contexts.

Concluding, conceptualising societal weighing as contextualisation practices focuses our gaze on two elements. First, it highlights the impact of patients, their representatives, and/or other stakeholders as representing the outside in terms of the quality of the decision justification, which is for the benefit of those outside; second, it shows the expertise of the committee in terms of choosing the right outside argumentations, combining them, and formulating recommendations, all deriving from or aimed at this same outside, i.e. the societal contexts as conceived of for that specific health care technology and coverage decision.

Societal weighing is achieved through contextualisation practices, which specifically conceive of patients, their representatives, and/or other stakeholders as representing outside contexts which

condition or enable utterances to happen and thereby increasing the quality of the decision justification or rationale for the benefit of those outside. Second, they comprise actively integrating argumentations deriving from outside contexts, showing that explication also happens in less controversial decisions, and adding recommendations to stakeholders which are positioned in certain outside contexts.

EXPERTS AS GUARANTORS OF DECISION QUALITY

In this section, I will briefly explore the role of the committee in terms of experts as guarantors of decision quality in view of the different 'types' of societal weighing that this dissertation has described, and how this differs from the traditional types of guarantors described in the Introduction, namely through adherence to procedural or substantive criteria.

Though all societal weighing processes, characterised by contextualisation practices, display elements of bringing the outside in and placing the decision back outside, the way they do so varies significantly. Societal weighing expressly does not look the same for every decision. The societal weighing takes a different shape every time and this is most obvious in how the committee assumes different roles and by extension in how the collective solidarity through the basic benefits basket assumes different shapes. Examples that stand in stark contrast include: "we want to make something possible here" (eculizumab case) or "simply stamping the file", meaning that the committee functions as an enabler of price negotiations (Dutch and English nivolumab decisions). Another apparently contradictory set would be 'not responsible for the disposable income of certain groups' (paracetamol-vitamin D case, as there are other responsible organisations, but also the Dutch walking aids with wheels decision) contrasted with 'responsible for children and young people' (maternity care and front teeth replacement therapy cases). Many may conceive of these differences as problematic. I will proceed to argue the opposite, namely that choosing the right role and thus demarcating the extent of the collective solidarity is exactly what the appraisal committee is supposed to do.

The robustness of a decision is dependent on the setting(s) in which the decision is to play a role, and the particulars of these settings are different every time. The societal context in this dissertation is expressly not used as a non-specific explanatory resource (Asdal, 2012; Asdal & Moser, 2012): *the societal context varies per decision as it is differently conceived of for that specific health care technology and coverage decision*. Let me demonstrate this. The examples above showcase different patients of differing ages, differing clinical pictures, but also differing ways in which the care is provided, different stakeholders and divisions of responsibilities, different initiatives and possibilities in terms of research but also in terms of price negotiations. Naturally, the effectiveness and cost-effectiveness of a health care technology differ every time as well but the role these criteria may play and their perceived validity do, in principle, not vary to the same extent across decisions. The committee, then, does well to be aware of the particulars of a decision's specific societal contexts and to take

them along actively in the decision-making process. In this active taking-along, this integration, it is important to note that the people present in the appraisal committee meeting, the patient (representative)s, other stakeholders, Institute employees, and the committee members themselves, all contribute to the fact that *not anything goes*. Instead, delineating the collective solidarity in the right way for a particular health care technology requires expertise in terms of societal weighing, evidenced by this specific set of contextualisation practices: bringing the outside in and placing the decision back outside in a careful, appropriate, and well-legitimated manner.

> *Not all societal weighing processes look the same: in fact, they vary significantly. This is shown by the fact that the committee takes different roles in the deliberations, and that correspondingly, the extent of the collective solidarity is demarcated differently. This, I argue, is not a problem – rather, it is exactly how it should be. Not only varies the health care technology per decision but also the contexts that are perceived as relevant are significantly different each time. This means that in expert contextualisation practices, through which a high-quality societal weighing is achieved, these different contexts are taken into account well. This results in carefully made, appropriate, well-legitimated, robust decisions.*

WEAVING NECESSITY

Above, I have described contextualisation practices as the processes of using argumentations to achieve an expert societal weighing. I have chosen the metaphor of weaving to denote these contextualisation practices and I will spend some words on this metaphor here. The contextualisation practices described are evidently more than muddling through, even elegantly; they require expertise to achieve well. They achieve more than rationality, even pragmatic rationality; they achieve a decision that is considered well-rounded and well-grounded. The metaphor of weaving, evoking notions of a structured, organised craft, with differing substrates in terms of colours and thickness but clear boundaries in terms of what may be achieved, I trust, helps bring out the art of making health care coverage decisions. Previous operationalisations of the necessity criterion, whether in the form of checklists or broadly-defined considerations, seem to have missed this aspect. I would argue they have thought more about the 'science' of decision making, and less about the art (Jacobs, 2017).

The use of the necessity criterion in practice, I would like to characterise as 'weaving necessity'. Necessity argumentations, which are many, case-specific, actively integrated during deliberations around the table, and varied in level of formalisation but used in patterns, ensure that decisions are not made without reference to relevant societal contexts. Weaving necessity, then, entails using these carefully-selected argumentations in a way that achieves a societal weighing (in terms of both bringing the outside in and placing the decision back outside), delineating the collective solidarity

in a way that is well-legitimated and sensitive to the case at hand, in a way that yields a decision that is robust in contexts that are considered relevant.

> *Weaving necessity is a metaphor I use to evoke the art of decision making, showcasing that necessity argumentations, which are many, case-specific, and patterned, are actively integrated into decisions through these contextualisation practices. Using necessity argumentations thus aids achieving a high-quality societal weighing through delineating the collective solidarity in an appropriate, robust manner.*

POLICY RECOMMENDATIONS

I would like to provide three specific policy recommendations and a brief reflection on all three. Before I begin, it is clear that policy practice is also on the move while research happens (Zorginstituut Nederland, 2017), and it may well also point to how the two worlds intertwine. This may happen through informal conversations and more formal presentations; the recent *Wikken en Wegen* report is written by my primary point of contact at the Institute, who also co-authored Chapter 2. Three specific elements of this report as describe above resonate strongly with this dissertation. First, the renewed interest for reference to the societal context. Second, the non-exhaustiveness of a list of criteria: additional relevant argumentations are always possible. Three, the importance of a strong justification or rationale to back up the advised decision. In line with the above, I would advise the Institute to continue on this path by following three recommendations.

Contextualisation practices entail first, bringing the outside in, in which patients, their representatives, and/or other stakeholders play a vital role. Consequently, my first recommendation is to take steps to *invite more different perspectives* into the deliberative process in the appraisal. This tallies with Moes' reasoning that this values especially patients in their capacity as knower (Moes, 2019) and the scoping sessions currently being institutionalised at the Institute (Zorginstituut Nederland, 2017). In my words, this entails bringing more different and more different types of argumentations into the decision-making process. To this end, I would encourage reflection on how to achieve that and whom to involve, as well as exchanges between HTA agencies on more qualitative aspects of health care coverage decisions. I would add that having more actors present in the decision-making setting might also contribute to the realisation of the rest of the decision through linking argumentations together into decision networks. This would also thus potentially facilitate a more rigorous testing step, as the appraisal committee currently tests these decision networks alone. This would enable research on how different forms of expertise and experiential knowledge work differently in terms of 'opening up the decision network (a suggestion which I owe to Professor Tiago Moreira). As a corollary, I suggest *institutionalising an improved appeals procedure*, as it will open opportunities to contribute to not only ongoing decisions but also to

those taken in the past, perhaps redressing at least part of the power differential inherent to these decisions.

Second, contextualisation practices comprise setting the decision back outside. That includes, as stated above, the involvement of others not only in the scoping phase but also in the testing during the appraisal phase. Moreover, as the justification or rationale is essential herein, I would suggest that the final advised coverage decision would benefit from *including alternative decisions that were considered but not chosen*. I would suggest that if for example both a positive and a negative decision were considered, or two variations on a positive decision, both potential decisions are displayed in the final advised decision including the argumentations that were considered to back up these alternatives. The reason is that this inclusion of alternative decisions would helpfully show that the concluding decision was not consensual, spelling out rather than obscuring differing points of view (Mouffe, 1999, 2011; Rip, 1986). Although a potential controversy may, of course, still follow, a relatively robust decision outcome will benefit from the transparency gained by the inclusion of alternative decisions that were also considered.

Third, I would advocate for the Institute to consider ways to *actively find controversy and invite it in* (again not just in the appraisal phase but more generally), as an additional way to bring the outside in. Callon et al. suggest:

> [Controversies] should be encouraged, stimulated, and organized. There are overflows everywhere. They produce the fabric of our individual and collective lives. (Callon et al., 2009, p. 257)

They see overflows, that is, places of controversy, everywhere. This yields another recommendation for the Institute: actively finding controversy that may be brewing 'in society', and drawing it in. This is another, more proactive form of contextualisation, and would perhaps take the form of horizon scanning not just for expensive medicines but for controversy in the making. Examples might include organising a conversation on the inclusion of menstrual cups in the basic benefits basket as suggested by a recent petition (https://petities.nl/petitions/menstruatiecup-in-de-basisverzekering, accessed 19 November 2019). This actively inviting in controversy in the making also yields a task for researchers: designing and experimenting with methods to identify (hidden) controversy, in addition to existing methods to map them (e.g., Marres, 2015; Munk et al., 2016).

All three recommendations may appear to rely on a somewhat idealistic and rationalistic description of deliberation, and as such, clash with the largely organic process that I describe in terms of coming to a decision through weaving. Regarding the second recommendation for example, in the contextualisation process, one potential outcome takes precedence over another, the committee considers it more worth investing 'weaving time' into, and as such, the decision that is not taken will never be as carefully woven as the decision that is. On the first and third recommendation, this weaving process does by no means guarantee all elements to be taken along by those who weave – and thus far, the experts at contextualisation have been the appraisal committee, which has given the deliberative situation another dimension to the already notable power differential in the room.

Finally, it does not explicitly acknowledge the antagonism that is and will be present, in part due to the inherent tension between the individual and the collective (Moreira, 2011), an antagonism that is not to be obscured by consensus (Mouffe, 1999, 2011). Thus far, the weaving process and robust outcomes of decisions have resonated with and strengthened one another, but here these conceptualisations do clash. I would argue that robustness takes precedence here. This means that even though this alternative decision is less carefully woven, which will probably take the form of being less precise, less extensive, and less well-formulated, it is worth including it in the decision outcome nevertheless, for reasons of transparency. These recommendations, then, ultimately hope to provide concrete steps to lessening the power differential somewhat, risking antagonism for the relatively robust and transparent outcomes it may produce.

STRENGTHS AND LIMITATIONS

This study is the first to produce an overview of argumentations pertaining to necessity. It is strong in terms of showing how these argumentations are used, as it does so through analysing observations/audio recordings, documents, and interviews in the Netherlands but also abroad. It interprets a relatively large number of cases (four Dutch appraisals and sixteen further decision outcomes) and finds commonalities in argumentation use across those. Naturally, because of time constraints, some in-depth familiarisation may have been lacking.

Moreover, my efforts have largely been focused on the appraisal phase and the documentation surrounding it, but contextualisation may be conceived of as much more than just the appraisal. This is visible, for example, in the scoping session and the Institute employees who write up the discussion documents. Many argumentations that were prepared were, however, verbalised, explicated, in the appraisal meeting, making it a prime locus for this study. Within the appraisal meetings, I have focused on explicit rather than implicit meaning-giving. This was, as reflected on below, formed by both my personal research interests as well as the practicalities of using explicated argumentations as sensitising concepts. It does remain likely that in this way, I have put less emphasis on more implicit forms of contextualisation. Similarly, I have had a relatively narrow focus on necessity. How these different contextualisation practices intertwine and what more implicit forms look like would benefit from further research.

REFLECTION ON PERSONAL RESEARCH POSITION AND ROLE

In answering the research questions, it is important to be aware of the epistemological angle I have approached them from and how this may have impacted the data collection and interpretation.

My reflections are on epistemology and specifically on how it impacts doing research in terms of methodology, conceptualisations, and conclusions. With my background in the natural sciences

(biomedical research to be precise), I am used to showing *care* towards my research subject(s). Counterintuitive as this may sound to some, cultured cells require diligent care (cleaning, feeding, and the like), and stem cells, in particular, are fussy research subjects, necessitating the researcher to give up Saturday afternoons to care-giving tasks. Naturally, this imperative of taking care holds even more for those with human research subjects and, according to Latour, especially of those who adopt a constructivist stance, as I have done in this research (Latour, 2004). To specify, by a constructivist stance, I understand that 'facts', scientific or otherwise, are made, knowledge is generated, in the sense that knowledge is *mediated* in some way.

This means first and foremost that my academic work has been mediated: I have been as involved in and as responsible for data generation as the people and practices I studied, and likely even more so. This is visible in the construction of the twenty necessity argumentation types, which are mine alone (but checked with a co-reader). It is no less true for the conclusions of Chapters 2 and 3, even though these were member-checked rigorously. Law and Singleton state that research means selecting, and that by writing up data into chapters, I betray part of the data that were available to me (Law & Singleton, 2013). I would have to agree that other conclusions, or at least other nuances, would indeed have been possible. I have two examples of how my personal preferences have shaped my research. First, as a researcher, I am interested in investigating what I consider tangible: I find it difficult to convincingly show how external factors may implicitly affect decisions (Asdal, 2012; Asdal & Moser, 2012). This has impacted the way I have opted to follow the argumentation types across the different observed meetings and documents (Chapters 2-4), rather than follow interviewees' ideas of how certain contexts, such as the presence of innovative pharmaceutical industry or David Cameron's desire to be re-elected, both of which were mentioned by interviewees, may have had an indirect impact on certain coverage decisions. Second, I have been shaped by a book called 'How People Think' (Jacobs, 2017), where Alan Jacobs convincingly argues (following Daniel Kahneman and Jonathan Haidt, amongst others) that people make decisions intuition-first. This is followed by a rationalisation of this intuition by providing reasons for why this intuition is right. I also believe, however, that this process of providing reasons is important for the legitimacy of public decisions especially. I think that decisions that are made by a direct appeal to a Minister of Health, who does not have to account for his or her decisions in the same way, are less valid – as visible in Chapter 4 especially. Overall, during its coming together and when it is finished, academic work is performative, it *does* something. I am therefore not surprised to see ongoing parallels between my work and the policy developments at the Institute (Law & Singleton, 2013).

On taking care as a constructivist, Latour specifies that:

> The critic is not the one who debunks, but the one who assembles. The critic is not the one who lifts the rugs from under the feet of the naive believers, but the one who offers the participants arenas in which to gather (…) one for whom, if something is constructed, then it means it is fragile and thus in great need of care and caution. (Latour, 2004, p. 246)

From this, I gather that to believe that the research subject is constructed and to show how it is constructed, how it is made, is to expose its vulnerabilities. For me personally, this only truly acquired meaning after a particularly vivid member check, in which I appeared to have misinterpreted the deliberations at the appraisal committee I studied. Sayer gives hands and feet to this idea of taking care when he says in his book 'Why Things Matter to People' that it is possible for social scientists to over-theorise to such an extent that the people studied *do not recognise themselves* (Sayer, 2011). He argues that social scientists should aim to remain close to the interpretation of those studied. These two ideas, of showing care and of remaining close to the interpretation of the research subjects, have helped me find direction for my dissertation. These ideas have impacted the methodology, some of which was chosen in collaboration with the research subjects (Chapter 2). These ideas may also have, perhaps, impacted the theoretical concepts employed to give meaning to the data gathered. After the deconstruction phase of studying the appraisal committee meetings and the advised decisions, I have chosen to conceptualise the *robustness* of health care coverage decisions. I believe this does justice to the everyday work at the Institute, and I believe it may potentially contribute constructively to future efforts.

REFERENCES

Asdal, K. (2012). Contexts in Action—and the Future of the Past in Sts. *Science, Technology, & Human Values, 37*(4), 379-403.

Asdal, K., & Moser, I. (2012). Experiments in Context and Contexting. *Science, Technology, & Human Values, 37*(4), 291-306.

Callon, M., Lascoumes, P., & Barthe, Y. (2009). Acting in an Uncertain World: An Essay on Technical Democracy (G. Burchell, Trans.): Cambridge, MA: MIT Press.

Calnan, M., Hashem, F., & Brown, P. (2017). Still Elegantly Muddling Through? Nice and Uncertainty in Decision Making About the Rationing of Expensive Medicines in England. *International Journal of Health Services, 47*(3), 571-594.

College Voor Zorgverzekeringen. (2006). Pakketbeheer in De Praktijk. Diemen.

College Voor Zorgverzekeringen. (2009). *Pakketbeheer in De Praktijk 2*. Diemen: College voor Zorgverzekeringen.

Commissie-Dunning. (1991). Kiezen En Delen. *Rapport van de Commissie Keuzen in de Zorg*.

Creswell, J. W., & Poth, C. N. (2017). *Qualitative Inquiry and Research Design: Choosing among Five Approaches*. USA: Sage publications.

Culyer, A. J., & Rawlins, M. D. (2004). National Institute for Clinical Excellence and Its Value Judgements. *British Medical Journal (BMJ), 329*, 224-227.

Hoedemaekers, R., & Oortwijn, W. (2003). Problematic Notions in Dutch Health Care Package Decisions. *Health Care Analysis, 11*(4), 287-294. doi: 10.1023/B:HCAN.0000010057.43321.b2

Hughes, D., & Doheny, S. (2011). Deliberating Tarceva: A Case Study of How British Nhs Managers Decide Whether to Purchase a High-Cost Drug in the Shadow of Nice Guidance. *Social Science & Medicine, 73*(10), 1460-1468.

Hughes, D., & Light, D. (2002). *Rationing: Constructed Realities and Professional Practices*. Oxford, UK: Wiley-Blackwell.

Jacobs, A. (2017). *How to Think: A Survival Guide for a World at Odds*: Currency.

Jenkings, K. N., & Barber, N. (2004). What Constitutes Evidence in Hospital New Drug Decision Making? *Social Science & Medicine, 58*(9), 1757-1766.

Latour, B. (2004). Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern. *Critical inquiry, 30*(2), 225-248.

Law, J., & Singleton, V. (2013). Ant and Politics: Working in and on the World. *Qualitative Sociology, 36*(4), 485-502.

Lehoux, P., Denis, J. L., Rock, M., Hivon, M., & Tailliez, S. (2010). How Medical Specialists Appraise Three Controversial Health Innovations: Scientific, Clinical and Social Arguments. *Sociology of health & illness, 32*(1), 123-139.

Marres, N. (2015). Why Map Issues? On Controversy Analysis as a Digital Method. *Science, Technology, & Human Values, 40*(5), 655-686.

Martin, D., Pater, J., & Singer, P. (2001). Priority-Setting Decisions for New Cancer Drugs: A Qualitative Case Study. *The Lancet, 358*(9294), 1676-1681. doi: https://doi.org/10.1016/S0140-6736(01)06714-9

Mechanic, D. (1997). Muddling through Elegantly: Finding the Proper Balance in Rationing. *Health Affairs, 16*(5), 83-92.

Moes, F. (2019). *Experts and Idiots: An Ethnography of Epistemic Practices in Dutch Social Health Insurance.* Maastricht.

Moreira, T. (2011). Health Care Rationing in an Age of Uncertainty: A Conceptual Model. *Social Science & Medicine, 72*(8), 1333-1341.

Moreira, T. (2012). Health Care Standards and the Politics of Singularities:Shifting in and out of Context. *Science, Technology, & Human Values, 37*(4), 307-331. doi: 10.1177/0162243911414921

Moreira, T. (2013). *The Transformation of Contemporary Health Care: The Market, the Laboratory, and the Forum*. USA: Routledge.

Mouffe, C. (1999). Deliberative Democracy or Agonistic Pluralism? *Social research*, 745-758.

Mouffe, C. (2011). *On the Political*: Routledge.

Munk, A. K., Meunier, A., & Venturini, T. (2016). *Data Sprints: A Collaborative Format in Digital Con-*

troversy Mapping: Princeton University Press Princeton, NJ.

Nice. (2008). *Social Value Judgements - Principles for the Development of Nice Guidance*. https://www.nice.org.uk/Media/Default/about/what-we-do/research-and-development/Social-Value-Judgements-principles-for-the-development-of-NICE-guidance.pdf.

Niëns, L. (2014). Affordability in Health Care: Operationalizations and Applications in Different Contexts.

Nowotny, H. (2003). Democratising Expertise and Socially Robust Knowledge. *Science and public policy, 30*(3), 151-156.

Oliver, A., Mossialos, E., & Robinson, R. (2004). Health Technology Assessment and Its Influence on Health-Care Priority Setting. *International Journal of Technology Assessment in Health Care, 20*(1), 1-10. doi: 10.1017/S026646230400073X

Patera, N., & Wild, C. (2014). Assessment – Appraisal – Decision. Lbi-Hta Decision Support Document Nr.: 72. Wien: Ludwig Boltzmann Institut für Health Technology Assessment.

Ragin, C. C. (2004). Turning the Tables: How Case-Oriented Research Challenges. *Rethinking social inquiry: Diverse tools, shared standards, 123*.

Rip, A. (1985). Experts in Public Arenas. *Regulating industrial risks: Science, hazards and public protection*, 94-110.

Rip, A. (1986). Controversies as Informal Technology a Ssessment. *Knowledge, 8*(2), 349-371.

Rip, A. (1992). Expert Advice and Pragmatic Rationality. *The Culture and power of knowledge: inquiries into contemporary societies*, 363.

Russell, J. (2017). *The Rationality of Rationing: A Rhetorical Policy Analysis of Deliberations About Resource Allocation in the Nhs.* University of Oxford, UK.

Russell, J., & Greenhalgh, T. (2014). Being 'Rational'and Being 'Human': How National Health Service Rationing Decisions Are Constructed as Rational by Resource Allocation Panels. *Health:, 18*(5), 441-457.

Sayer, A. (2011). *Why Things Matter to People: Social Science, Values and Ethical Life*: Cambridge University Press.

Shah, K. K., Cookson, R., Culyer, A. J., & Littlejohns, P. (2013). Nice's Social Value Judgements About Equity in Health and Health Care. *Health Economics, Policy and Law, 8*(2), 145-165.

Singer, P. A., Martin, D., Giacomini Bhatia, M. V., & Purdy, L. (2000). Priority Setting for New Technologies in Medicine: Qualitative Case Study. *Bmj, 321*(7272), 1316-1318. doi: 10.1136/bmj.321.7272.1316

Stolk, E., Brouwer, W., & Busschbach, J. (2002). Rationalising Rationing: Economic and Other Considerations in the Debate About Funding of Viagra. *Health Policy, 59*(1), 53-63.

Stolk, E., & Poley, M. (2005). Criteria for Determining a Basic Health Services Package. *The European Journal of Health Economics, 6*(1), 2-7.

Wallenburg, I., Weggelaar, A. M., & Bal, R. (2019). Walking the Tightrope: How Rebels "Do" Quality of Care in Healthcare Organizations. *Journal of Health Organization and Management*.

Walley, T. (2007). Health Technology Assessment in England: Assessment and Appraisal. *Medical Journal of Australia, 187*(5), 283.

Zorginstituut Nederland. (2013). Pakketbeheer in De Praktijk Deel 3 (pp. 1-106). Diemen.

Zorginstituut Nederland. (2016). Reglement Adviescommissie Pakket. In Z. Nederland (Ed.). https://www.zorginstituutnederland.nl/over-ons/publicaties/besluit/2016/12/05/reglement-adviescommissie-pakket-zorginstituut-nederland (accessed 11 February 2020).

Zorginstituut Nederland. (2017). Pakketadvies in De Praktijk: Wikken En Wegen Voor Een Rechtvaardig Pakket. Diemen.