Making epistemic goods compatible: knowledge-making practices in a lifestyle intervention RCT on mindfulness and compassion meditation

BioSocieties (2023) 18:359–388
https://doi.org/10.1057/s41292-022-00272-w

ORIGINAL ARTICLE

Making epistemic goods compatible: knowledge-making practices in a lifestyle intervention RCT on mindfulness and compassion meditation

Mareike Smolka[1]

Accepted: 1 February 2022 / Published online: 8 March 2022

Abstract Mindfulness and compassion meditation is a popular lifestyle intervention in randomised controlled clinical trials (RCTs), which examine its efficacy to ameliorate health and well-being. Studying meditation in an RCT poses the challenge of standardising an intervention that relies on a mix of people, skills and activities. This article describes how, in meeting this challenge, researchers engage in diverging knowledge-making practices. It draws on praxiographic inquiry in an RCT on the effects of meditation compared to a foreign language training on healthy ageing. To analyse normative dimensions of knowledge-making practices, the concept of 'epistemic goods' is introduced. Researchers juggled partly incoherent epistemic goods—internal validity, social relevance, assessing efficacy, attending to qualitative effects, objectivity, trained judgment—and resolved tensions between them. Strategies to respond to unexpected events in the research process were: reinterpreting the study protocol, caring informally while playing by formal rules and adjusting the procedure of a study task. Analysing epistemic goods and strategies that make them coexist is relevant to problematise what counts as evidence in evidence-based medicine. Instead of evaluating knowledge by reference to a 'gold standard', evidence claims should be placed in the context of their production to evaluate them on their own terms.

Keywords Randomised controlled clinical trial · Sociology of standardisation · Empirical ethics · Knowledge-making practices · Evidence · Health promotion

✉ Mareike Smolka
m.smolka@maastrichtuniversity.nl; mareike.smolka@posteo.de

[1] Faculty of Arts and Social Sciences, Maastricht University, Grote Gracht 90-92, 6211 SZ Maastricht, The Netherlands

Introduction

Clinical research is currently shaped by the convergence of two turns (Winther and Hillersdal 2020). On the one hand, a turn towards standardisation and objectivity has placed the randomised controlled clinical trial (RCT) at the top of the knowledge hierarchy in evidence-based medicine (Knaapen 2014). On the other hand, a turn towards prevention has given rise to a growing number of studies on behavioural interventions that encourage healthy participants to adopt different lifestyles (Holman et al. 2018). These turns sit uncomfortably together because behavioural interventions are by default less likely to be tested and less likely to be proven effective in RCTs in comparison to pharmacological interventions (Lambert 2006). One of the reasons is that RCTs are designed to identify the efficacy of a specific intervention under highly standardised, controlled and monitored conditions that separate the intervention from external influences. Standardisation is particularly complicated in behavioural intervention trials, because these interventions rely on a mix of people, skills, devices, activities, processes and environments (Wells et al. 2012). Whether a behavioural intervention works, depends to a high degree on its adaptation to contexts of implementation (Cohn and Lynch 2017).

A solid body of literature in Science and Technology Studies (STS) on interventions involving medical drugs or devices has shown that the successful completion of a trial—one that is actually able to recruit participants and gather outcome data—depends on the alignment of standardised protocols with already existing practices in local settings (Berg 1998; Hauskeller et al. 2019; Hogle 1995; Jonvallen 2005; Keating and Cambrosio 2007; Timmermans and Berg 2003; Webster et al. 2011; Will and Moreira 2010). Behavioural lifestyle interventions have received less attention, but a few studies reveal that socio-material contexts (e.g., care practices, social relations, study equipment, infrastructures) and study participants (e.g., motivation, meaning-making, affect) shape clinical trial research processes and results (Cohn and Lynch 2017; Jespersen et al. 2014; Rogers et al. 2005; Wolters et al. 2018). Accordingly, mixed-methods approaches have been developed to improve study designs and evaluations through qualitative research (Mannell and Davis 2019). Although there is a growing trend towards more flexible, adaptive and ecological designs (Bonell et al. 2012; Montgomery 2016; Ong et al. 2014), the classic RCT remains the dominant approach to investigate the efficacy of lifestyle interventions (Green and Kolar 2015; Holman et al. 2018).

I contribute to STS research on clinical trials by shedding light on the knowledge-making practices in a lifestyle intervention RCT. My research builds on previous studies in the fields of disease prevention and public health that analyse how clinical trial researchers carefully negotiate between seemingly contradictory scientific norms—methodological purity on the one hand and social relevance on the other—to create what they consider as robust evidence (Rod et al. 2014; Will 2007; Winther and Hillersdal 2020). To expand on these analyses, I examine how knowledge-making practices in an RCT on mindfulness and compassion

meditation are related to multiple scientific norms and how researchers engage with tensions between these practices. An RCT on meditation is a relevant case because meditation has become an increasingly popular and contested lifestyle intervention over the last two decades (Gotink et al. 2015; van Dam et al. 2018). While some scientists and clinicians see RCT research as an opportunity to strengthen the therapeutic legitimacy of mindfulness (e.g., Kabat-Zinn 2011), others take the view that a rigid RCT design is too restricted to capture the processes underpinning the complex effects that meditation may have on the human body and mind (e.g., Lutz et al. 2015).

To study practices of doing good clinical trial research on meditation, I conducted praxiographic research from 2017 to 2020 in the Age-Well clinical trial, which is part of the Silver Santé Study. The Silver Santé Study, also known as Medit-Ageing among researchers, is a European Horizon 2020 project that combines two study protocols: SCD-Well and Age-Well (Fig. 1). Age-Well is a three-armed RCT with 137 study participants conducted in the city of Caen in France. It compares the effects of an 18-month meditation intervention with a foreign language (English) training intervention—hereafter abbreviated to 'English intervention'—and a passive control group on mental health and well-being in older adults (Poisnel et al. 2018). Moreover, the trial includes a group of 30 long-term meditators who undergo the same battery of study examinations as the older adults. Comparing "novice" with "expert" meditators (Lutz et al. 2018, p. 759) helps to understand the mechanisms underlying meditation that are assessed with cognitive, behavioural, biological, neuroimaging and sleep examinations.

My case study seeks to answer the following research questions: How are knowledge-making practices enacted in the Age-Well trial of the Silver Santé Study? How are these practices related to scientific norms of good clinical research on meditation? Which strategies do Silver Santé researchers deploy to resolve tensions between different knowledge-making practices? To answer these questions, I introduce the concept of 'epistemic goods' by combining Pol's (2015) approach to

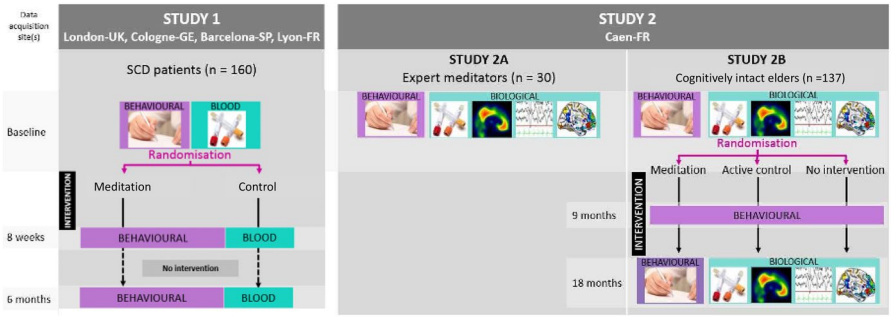

Fig. 1 Study scheme of the Silver Santé Study. The Silver Santé Study combines two clinical trials: Study 1 SCD-Well, Study 2 Age-Well. Age-Well consists of two studies: Study 2A is an observational study with expert meditators, while Study 2B is an RCT with healthy elderly adults who participate in a meditation intervention, an English intervention or a passive control group. The scheme was adapted from public communication materials published on www.chetelat-lab.fr/silver-sante-study/

empirical ethics with Daston and Galison's (2007) understanding of "epistemic virtues" (p. 39). In performing epistemic goods, research teams enact norms of doing good research. The verb 'to enact' highlights that what counts as good research is constituted in and through its practical accomplishment (Mol 2002). I identified multiple epistemic goods in the Age-Well clinical trial and traced seeming contradictions between them: internal validity and social relevance, assessing efficacy and attending to qualitative effects, objectivity and trained judgment. Although these paired epistemic goods are not necessarily incompatible, tensions occurred between them in research practice and occasionally gave rise to debates within the Silver Santé team about how to resolve these tensions.

For this purpose, Silver Santé researchers deployed different strategies: reinterpreting the study protocol, caring informally while playing by formal rules and adjusting the procedure of a study task. A strategy does not imply a pre-defined plan to reach a certain goal under conditions of uncertainty, but is here analysed as situated "tinkering" (Berg 1998, p. 237) with standardised procedures inscribed in a clinical trial protocol. Informed by the sociology of standardisation (Timmermans and Epstein 2010), I study how the Age-Well trial became "do-able" (Fujimura 1987) in light of multiple epistemic goods partly pulling practices into opposite directions—towards and away from strict adherence to the trial protocol. Studying epistemic goods and strategies that make them coexist in a clinical trial is relevant to rethink what counts as robust knowledge and to adopt an empirical ethics approach in knowledge evaluation.

Theoretical background

Empirical research on goods as practices is the programme of an empirical ethics (Mol et al. 2010; Pols 2013, 2015, 2018). Rather than reasoning about normativity in the abstract, empirical ethics studies how people attempt to accomplish something good with the help of devices, routines and concepts. What is good can be traced empirically in the activities that advance people's values, ideals or tastes. Studies have examined how ideals like dignity (Pols et al. 2018) and individualisation (Pols 2008) were enacted in multiple ways in healthcare practices.

Epistemic goods are practices of doing good research that enact scientific norms. The concept is informed by Daston and Galison's (2007) seminal work on *Objectivity*. The authors uncover how the norm of objectivity has infused practices of image-making for scientific atlases since the middle of the nineteenth century, thereby shaping how scientists view the world. To underline the intricate relation between scientific norms, practices and knowledge, they call objectivity an epistemic virtue. Epistemic virtues "are norms that are internalised and enforced by appeal to ethical values, as well as to pragmatic efficacy in securing knowledge" (pp. 40–41). Daston and Galison reconceive epistemology as ethics by considering it as a repository of multiple versions of the good that are products of distinct historical circumstances, but have persisted over time in knowledge-making practices. Looking at practices throws frictions between these goods into relief—for example, precision

and replicability can come at each other's expense—while acknowledging that what doing good research looks like is situated and context-dependent.

Instead of calling these practices virtues, which links their mastery to character and skills development for becoming a good scientist, I use the concept of epistemic goods to analyse their relational nature. In line with Pols' (2015) approach to empirical ethics, I understand practices as "the interrelational achievement of people, as well as the technologies and concepts they use" (p. 82). Within these relations, practices (of attempting) to do good appear in different forms. As people are assumed to be relational entities, different versions of the good cannot be attributed to independently acting subjects, but instead emerge within networks of heterogeneous actors.

An empirical ethics approach also sheds light on the day-to-day labour and care practices that are necessary to do good research and produce sound knowledge (Swallow et al. 2020). It draws attention to private, embodied, emotional and messy aspects of science that are essential to its achievement but seem to conflict with epistemic goods like objectivity and precision (Friese 2013). STS researchers have invoked the concept of care to analyse a range of practices that remain invisible in the polished accounts of science (Latimer and Puig de la Bellacasa 2013; Puig de la Bellacasa 2011), such as the attentive and affective interactions with laboratory animals, study participants and databases (Lappé 2018; Pinel et al. 2020; Wadman and Hoeyer 2014). But researchers tend to consider such care practices as sources of 'local variation' and 'noise' that disturb standardised clinical trial procedures (Fisher 2006; Hallowell et al. 2009). These disturbances must be weeded out in the research process or should be corrected in the analysis to conduct methodologically rigorous science (Danziger 1990). Strict adherence to the procedures specified in the clinical trial protocol is thought to minimise personal biases and enhance comparability between different intervention groups (Dehue 2002, 2010).

As established by research on clinical trials in the sociology of standardisation (Bowker and Star 1999; Lampland and Star 2009; Timmermans 2014; Timmermans and Epstein 2010), however, protocols are not as inflexible in practice as they appear on paper (Berg 1998; Jonvallen 2005; Keating and Cambrosio 2007; Will 2010). Although protocols contain detailed sequenced prescriptions of how to act in a given situation, there are a number of roles, tasks and assumptions that are not written out in a protocol but become visible once it touches existing practices. Heterogeneous practices are always already in place in hospitals or research institutes where clinical trials are carried out. These practices entail, among other things, scheduled activities of researchers and physicians for other projects; availability and accessibility of measuring devices, laboratory tests and examination rooms; and specific skills of research team members to perform study examinations. What a protocol prescribes must be aligned with the practices given in particular organisational contexts (Hauskeller et al. 2019; Timmermans and Berg 2003).

Berg (1998) shows that the practices defined, coordinated and ordered by a clinical trial protocol are inextricably linked with the practices involved in constructing the protocol. As he argues: "[A] protocol is not simply imposed on the diverse practices. Rather, the construction (and implementation) of a protocol is a process of ongoing, continuing *negotiations*" (p. 235). Only in retrospect, after a clinical trial has been completed, the protocol loses the traces of "tinkering" (p. 237) that was

necessary to make it work. Tinkering involves situated judgments, local knowledge and creativity in using a standard, so that local practices become standardised and "the standard is localised" (Knaapen 2014, p. 830), especially in view of unexpected events and particularities.

In the empirical analysis below, I identify different forms of tinkering as strategies that researchers use to enact diverging epistemic goods together. These strategies are oft-informal "ordering process[es]" (Mesman 2008, p. 9) that emerge in the everyday organisation of work practices. Tracing ordering processes is relevant to better understand clinical research because they define who and what should be included and excluded from knowledge production. By analysing strategies for tinkering with the study protocol, I highlight how (social) order is co-produced with knowledge-making practices (Jasanoff 2004).

Methods

For an inquiry into practices, this study used the methodology of "praxiography" (Mol 2002, p. 31), which is akin to ethnography but differs in focus and emphasis. While *graphy* refers to the common task of recording, describing and writing about a phenomenon, praxiography is not so much interested in *ethno* (culture) but in *praxis* (practice). Knowledge-making practices have been a focus of attention in STS laboratory studies that trace the large amount of calibration work and various standardisation practices required to (re-)produce a scientific finding across sites and transform it into a 'universally' valid fact (e.g., Collins 1992; Latour 1983; Latour and Woolgar 1987). As noted by Garforth (2012), the methodology of laboratory studies centres on "seeing close up, in context, and in the middle of the action" (p. 269). Yet, observation may be perceived as intrusive, and some practices resist being witnessed because of their solitary or regulated nature (Garforth 2012; Star and Strauss 1999). STS researchers have therefore made use of alternative methods like interviews and document analysis to study (knowledge-making) practices (Beaulieu 2002, 2010).

My praxiographic research on the Silver Santé Study combined three means of data collection: participant observation, interviews and document analysis. I observed the Silver Santé team during two periods of fieldwork (Sep–Dec 2018, Oct–Dec 2019) and several short-term visits of the biomedical imaging platform Cyceron in Caen. I did observations of regular team meetings and study examinations (e.g., neuroimaging, behavioural and neuropsychological tests, polysomnography), as well as during public events for participant recruitment, science communication efforts and the annual two-day meeting of the European consortium. During fieldwork, I realised that social interaction, especially in conversations, helped me to learn more about practices, rather than simply being physically present. A reason is that a large part of scientific work was conducted silently behind computer screens, which is why it felt intrusive to shadow activities like data processing and analysis. Another reason is that I was not permitted to observe the study interventions because researchers were concerned that I could influence the dynamics in the English and meditation intervention groups. What is not accessible for observation, however,

generates its own data (Garforth 2012). I learned that the epistemic good of objectivity, or codified rule-following to minimise external influences on experimental conditions, was enacted in the Age-Well trial by limiting access to the interventions.

Interviews allowed me to circumvent the problem of direct access by asking questions about an interviewee's involvement in the Silver Santé Study, the tasks that had to be fulfilled to keep the study running, and the specifics of work practices. I conducted 74 semi-structured interviews with diverse actors involved in the Silver Santé Study. The large number of interviews resulted from the praxiographic approach to data collection. Mol's (2002) praxiography draws on actor-network theory, which situates practices in a flat network of actors. A key tenet is that one should not assume a priori which actors are included in a network and which of them are more important than others. Therefore, I did not only interview the local research team at Cyceron, but recruited interviewees from wider circles around my primary field site whenever I noticed them to be involved in the practices I was exploring. Table 1 categorises actors into social groups whose interviews turned out to be relevant for the empirical analysis.

To prepare for interviews, I studied a range of documents to gain a better understanding of the Silver Santé Study: scientific articles, conference presentations, participant recruitment materials and news items. Scientific articles and perspective pieces on meditation research more broadly helped me to become acquainted with this research field, making it possible for me to interpret practices in the Silver Santé Study in their wider context. Furthermore, the Age-Well clinical trial protocol was an important document to study practices. It provides detailed prescriptions of sequenced steps as to how to act in a given situation, criteria on whether and when the next step can be taken, as well as standards and classifications involved in recruitment, examinations and analysis. I got access to a confidential version of the protocol from 2018. For confidentiality reasons, I do not quote directly from the protocol; I only used it in the analysis process to triangulate information that I had gathered in interviews, participant observation and scientific publications.

Table 1 Social groups in the Age-Well clinical trial

Research team in Caen	Principal investigator, senior researchers, project managers, postdoctoral and PhD researchers, research assistants and technicians, neuropsychologists, physicians, communication officer, administrator, English teachers, meditation instructors
Participants	Healthy elderly adults who participated in the English intervention, the meditation intervention, the passive control group or the group of expert meditators
Sponsor	Representatives of a French public research organisation that assumes responsibility for the quality of scientific data and results, safety of participants, regulatory aspects and budget management
Methodologist	Independent experts supporting the sponsor and research team in data storage, research methodology and statistical analysis
European consortium	Researchers from European institutions (France, Belgium, Switzerland, London, Germany and Spain) involved in the data analysis, legal and administrative managers of the consortium, European communication officers

Interviewees were recruited from all social groups

Empirical analysis of the Age-Well clinical trial

To process the empirical data, this study draws on abductive analysis (Timmermans and Tavory 2012). Abductive analysis refers to an inferential process of producing theoretical insights based on surprising research evidence. While Timmermans and Tavory identify an empirical finding as surprising against a background of theoretical literature, I recognised a finding as surprising or unusual in relation to other observations. Three events stood out from interview transcripts and fieldwork materials because they broke with the ordinary. While most breaks with regular patterns of ordinary practices were repaired relatively swiftly in the Age-Well trial, the events selected for analysis created enough of a disruption to force actors to carefully rethink and renegotiate ways of doing good research. Moreover, as suggested by Pols (2013), studying events sheds light on routine ways of doing good research because they "may teach us something about the *conditions* that allow particular dramas to emerge" (p. 22). In analysing events, I examined how different ways of doing good research imbued day-to-day practices in Age-Well.

To gain an impression of everyday routines, extended periods of fieldwork and a substantial number of interviews were necessary. Only a part of the data corpus, however, helped me to describe the events selected for in-depth analysis. Limiting the amount of data presented here to a fraction of its corpus was a result of moving back and forth between empirical materials and theoretical literature. While I was searching for concepts that would help me interpret all the data, I was also looking for a definition of the research problem, which, without hiding anything, would make relevant only those data that fitted the concepts (cf. Katz 2001). In this way, the empirical focus on events and the analytical focus on tensions between epistemic goods co-emerged in the analysis process.

To avoid imposing theoretical concepts on the empirical material, I asked actors involved in the events for a written commentary on an earlier version of this article and presented preliminary analyses to the Age-Well research team in Caen and to the European consortium of the Silver Santé Study. I combined their feedback not only with my initial empirical findings but also with STS theory on scientific norms and knowledge production. Through this process, I reconstructed events, interpreted practices of doing clinical research as epistemic goods and analysed strategies to reconcile tensions between these practices. To further improve the validity of data analysis, I verified that the epistemic goods identified in relation to events could be applied across a multiplicity of scenes and actions captured in fieldnotes and interviews.

Reinterpreting the protocol in the English intervention: internal validity and social relevance

Event in the English intervention

In spring 2017, the first cohort of 43 healthy elderly adults started their participation in the three Age-Well study groups: English, meditation, or the passive control group. It is one of three cohorts in which 137 participants were recruited successively in the French city of Caen and its surroundings (Table 2). Participants in the English group followed a foreign language training programme to develop abilities in understanding, writing and speaking English. The language training programme was structurally matched to the meditation programme designed to cultivate mindfulness, kindness and compassion abilities. Apart from participating in weekly two-hour group sessions guided by experienced instructors, the English group and the meditation intervention group completed meditation and English exercises with a digital tablet at least 20 min every day at home, while also attending one day of intense group practice during the intervention.

Soon after the start of the intervention, the English group realised that participants had different levels of prior knowledge in English. While some could hold conversations in English, others had barely any command of the language. After trying to follow the English intervention for about three months, two participants decided to quit their study participation:

> We said: 'Stop! We quit.' I said: 'I quit. I cannot not keep up with this pace. This is not possible.' This was not the original rule . . . One could say that there were no exams and therefore no need to worry, but it is still dispiriting not to be able to follow [the English classes].

The "original rule" refers to the inclusion criteria of the study: age of 65 +, retired for at least one year, living autonomously, availability for the study for 24 months, overall health without chronic or acute diseases, normal performance on cognitive tests, level of education of more than seven years, no preference for any of the three study groups, neither practice meditation regularly nor speak English fluently (Poisnel et al. 2018).

To verify if participants met the inclusion criterion 'not speaking English fluently', they were asked whether they could hold a conversation in English and completed multiple choice comprehension exercises. The aim was to find participants with intermediate English competence to create a relatively homogeneous group in which all participants would manage to cope with the same learning material. But the screening may have failed to filter out beginners and advanced English speakers. Or participants with a high command of English may not have completed the screening test truthfully because of their eagerness to participate in the study. Another interpretation is that the intermediate level was stretched because it was difficult to recruit sufficient study participants who met all inclusion criteria.

Because the inclusion criterion left some leeway as to the exact meaning of 'not speaking English fluently' in practice, this may have eased the participant

Table 2 Schedule of the Age-Well clinical trial

	43 healthy elderly adults	50 healthy elderly adults	44 healthy elderly adults
Pre-selection questionnaire	Sep–End of Nov 2016	March–May 2017	March–Nov 2017
Individual interviews	Nov–Dec 2016	May–June 2017	Nov 2017–Jan 2018
Selection visit: medical and neuropsychological examinations	Nov–Dec 2016	June–July 2017	Jan–Feb 2018
Study examinations at inclusion	Dec 2016–Feb 2017	End of August–End of October 2017	Beginning of May 2018
Randomisation	Beginning of March 2017	Beginning of November 2017	Beginning of May 2018
Interventions (a) Meditation (b) English (c) Control	March 2017–Sep 2018	Nov 2017–May 2019	May 2018–Nov 2019
Study examinations after 18 months	Sep–Oct 2018	May–June 2019	Nov–Dec 2019

The schedule indicates the time periods allocated to the inclusion process, study examinations and interventions of the three cohorts of study participants in the Age-Well clinical trial. The table was created based on Poisnel et al. (2018) and on a presentation given during an Age-Well recruitment conference at the new faculty of medicine of the University of Caen in France, 25th of April 2017

recruitment process. At the same time, this leeway led to a situation in which two participants intended to quit the study because they felt uncomfortable in not being able to keep up with the weekly English classes. In trying to deal with this situation, the research team was juggling with two epistemic goods: internal validity and social relevance.

Internal validity and social relevance

Internal validity is a cardinal epistemic good in clinical trial research. The RCT is considered as the "gold standard" for intervention studies to assess a causal relation between an intervention and its effect (Timmermans and Berg 2003). A well-executed RCT has high internal validity because it effectively rules out other explanations for the observed effect. To rule out alternative explanations, RCTs have a control group, participants are randomly allocated to study groups, they follow strict compliance criteria, and both researchers and participants are blinded to the nature of the intervention.

The lack of internal validity in studies with a meditation intervention has become a key concern for meditation researchers. In light of the recent public and scientific 'hype' of mindfulness meditation as a panacea for a range of mental and physical ailments, researchers have called for caution regarding the robustness of evidence (Davidson and Kaszniak 2015; Rosenkranz et al. 2019; Vago et al. 2019; van Dam et al. 2018). In their critical evaluation of meditation studies, Davidson and Kaszniak (2015) state that studies on meditation have seldomly followed "double-blind placebo-controlled designs".

> This fact is partially responsible for the poor quality of clinical trials of meditation that have appeared in the scientific literature and is one important reason why recent meta-analyses of the clinical impact of meditation have reported so few rigorous studies that are judged to be methodologically sound. (p. 583)

To strengthen internal validity, the Silver Santé Study was designed in a "more rigorous" way (Klimecki et al. 2019, p. 223). The design of the English and the meditation intervention fulfils most of Davidson and Kaszniak's criteria for a "rigorous control condition" (p. 588). Participants were randomly assigned to the English group, the meditation group and the control group, so that results could be generalised to a random sample rather than to people who feel drawn to meditation or English. Moreover, the English intervention and the meditation intervention were structurally equivalent: the interventions were equal in length; they involved the same amount of group sessions and individual homework; and both the meditation instructors and the English teachers were comparably trained. This criterion was important to ensure that participants in both groups would be exposed to a similar amount of cognitive training. Similar levels of exposure in both study groups were necessary for a comparison of effect sizes regarding the changes in volume and cerebral blood flow in the anterior cingulate cortex and the insula—the main objective of the Age-Well RCT (Poisnel et al. 2018).

The pursuit of the hallmarks of internal validity in an RCT, however, creates frictions with another epistemic good: social relevance (Winther and Hillersdal

2020). An RCT has high social relevance if participants adhere to and engage with the intervention. Meditation researchers Rosenkranz et al. (2019) comment as follows on the importance of social relevance in their evaluation of mindfulness-based intervention research: "Choice is a strong predictor of adherence to and engagement with an intervention and effect sizes are typically higher when an intervention is individually initiated, rather than the consequence of random assignment" (p. 180). They point out that studies on meditation differ in some crucial respects from pharmaceutical trials for which the RCT design was championed. In pharmaceutical trials, one can be fairly confident that all participants in the treatment group receive the same dose of a drug. In behavioural interventions, the 'dose' depends on how an individual engages with the intervention. Therefore, they suggest to create study groups for whom an intervention is socially relevant. Studies must not necessarily aim to produce results that are generalisable to a random sample, but to individuals who would initiate an intervention such as a regular meditation practice or English language training.

The event in the English intervention highlights that internal validity was in tension with social relevance. The randomisation of Silver Santé participants neglected their social histories and the social dynamics that emerged in the study groups. It implies a model of social situations as a multitude of separate, identifiable elements with additive interconnections (Danziger 2000). In the English group, however, participants' social histories (discomforting childhood memories of studying at school) and social dynamics (experience of pressure to perform in front of a group) influenced their engagement with the intervention to the extent that two of them intended to leave the study. The reality of teaching English to a group of elderly people differed from what was assumed in the design of an RCT, whose internal validity depends on the artificial nature of study groups.

Reinterpreting the study protocol

In light of the difficulties in the English group of the first participant cohort, one particular fieldwork observation during the annual European consortium meeting of the Silver Santé Study in 2019 proved a surprise. A Silver Santé researcher congratulated his colleagues who had been in charge of the Age-Well data collection for a "big achievement" since he had rarely seen a clinical trial without drop-out. This implies that the participants who intended to leave the trial ultimately decided to continue following the English intervention until the end. What had motivated these participants to remain in the study?

One of them told me what happened when she met the Silver Santé project manager to announce that she and another participant would quit their study participation in summer 2017:

> [The project manager] proposed that we could stay in the study, keep the materials, meaning the tablet and the books, and do whatever we wanted, to not go to the [English] classes anymore . . . We accepted and this was great because we eventually played the game nevertheless. This means that we worked, but that we did so at our own pace. Every day we worked for half an hour, three

> quarters of an hour, or one hour, but at our pace, without the pressure of the group that was not at our level. I eventually progressed a lot.

The participant further emphasised that, from her point of view, she had engaged in a more effective and more comfortable way with the intervention without participating in the weekly English group sessions. She had spent more time practicing English on her own—completing English language exercises on a digital tablet and working with a language training book that was also used during the English group sessions—than the daily minimum of 20 min prescribed in the study protocol. Learning English had become relevant and manageable for her.

Making the English intervention socially relevant was a matter of doing good research in the Age-Well clinical trial, which the project manager described as follows:

> It seemed important to adapt to the field and to the people in front of us rather than following the rules foolishly. For example, we made English groups and noticed that the English level, although we had tried to be fairly homogeneous, was still pretty heterogeneous . . . this brings participants into very uncomfortable situations.

To create an intervention that was socially relevant—one that would hold in real life for people with social histories and for whom group dynamics matter—the research team reinterpreted what it meant to participate in the intervention. Instead of "following the rules foolishly", reinterpreting the study protocol implied adjusting the way the protocol was imposed so that the participants who had struggled with the weekly English sessions could complete the trial in a way that was suitable for them.

At the same time, keeping these participants in the study boosted its internal validity. Because statistical significance of study results partly depends on the number of participants, drop-outs introduce uncertainty in evaluating the efficacy of an intervention (Wadmann and Hoeyer 2014). If participants leave the study this may suggest that the results are not based on a random sample, but on a sample including people who did not leave the study, possibly because of a common characteristic. Enhancing social relevance thus fostered internal validity through reinterpreting the study protocol.

Caring informally while playing by formal rules in the meditation intervention: assessing efficacy and attending to effects

Event in the meditation intervention

In summer 2018, towards the end of the intervention of the first Age-Well cohort, a participant wrote a letter to the Silver Santé team on behalf of the meditation group. He explained that the group was motivated to go on meditating together in spite of the imminent end of the intervention. They intended to proceed with regular meditation group sessions supported by the principal meditation instructor who had facilitated most of the intervention. He also asked if they could continue

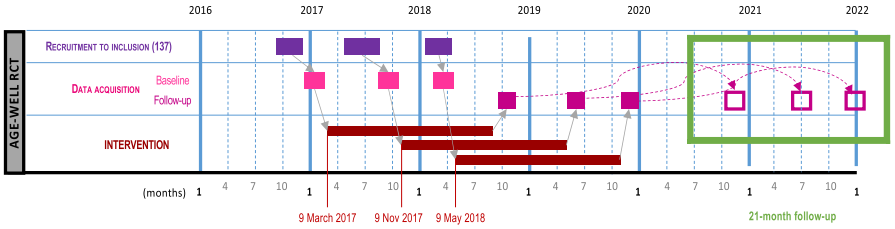

Fig. 2 Study scheme of the Age-Well clinical trial with 21-month follow-up examinations. The scheme visualises the periods of participant recruitment, data acquisition (baseline, follow-up after the 18-month intervention, and follow-up 21 month after baseline) and study interventions, which all started at different points in time in the three study cohorts. It is a modified version of a figure presented during an online European consortium meeting of the Silver Santé Study, 12th–13th of October 2020

using a room in the faculty of medicine in Caen where the weekly group sessions of the meditation intervention had taken place. The letter led to negotiations within the Silver Santé team about what should be done in response to the participants' request.

On the one hand, researchers were reluctant to meet the request because they had recently amended the clinical trial protocol to include a third battery of study examinations 21 months after the end of the intervention (Fig. 2). The scientific objective of this amendment was to validate whether the efficacy of the intervention would be maintained in the long run. The amended protocol specified that participants could proceed with practicing meditation or learning English on their own after the end of the intervention but not with the support of the Silver Santé team. Supporting participants in organising post-intervention meditation sessions would violate the protocol. Moreover, it would curtail the comparability between the 21-month follow-up study examinations of the English group and the meditation group. The reason is that no arrangements were made to prolong the English intervention beyond its official end because the English group had not put in such a request.

On the other hand, the research team acknowledged that meditation differed from learning English. While learning English was considered pre-eminently a cognitive training, meditation was also an affective training. Although both interventions were thought to give rise to affective group dynamics and might therefore affect participants' emotional well-being, meditation was assumed to have a specific effect on emotion regulation (Poisnel et al. 2018). A participant reported how meditation had helped him "channel" his emotions, so that he managed to stay calm in situations that used to make him angry. Meditation had "modified something within" which made him "live differently". As meditation had impacted how participants related to themselves and their lives, the principal meditation instructor stressed the importance of supporting the meditation group after the end of the intervention:

> Meditation is a process, it's not like learning about mathematics or a language, it's learning about yourself. It's transforming something inside you and when you are in the process, something organically grows . . . [A]fter 18 months they [the meditation group] want to continue to do something and we have to show them some tools or give them some information or support them.

Accordingly, meditation had different effects on study participants than the English intervention. Attending to these effects, however, seemed to be at odds with measuring the efficacy of the meditation intervention in the framework of the AgeWell RCT.

Assessing efficacy and attending to effects

RCTs assess the efficacy of an intervention, or, in other words, whether it meets the expected outcome in the study sample under controlled conditions (Streiner 2002). They promise to prove the instantaneous efficacy of an ameliorative intervention unambiguously (Dehue 2002). For this reason, evidence derived from RCTs is the dominant type of knowledge used for the development of guidelines that inform clinical decision-making and help governments as well as health insurers to allocate scarce resources (Wieringa et al. 2018). The growing number of RCTs which demonstrate the benefits of mindfulness-based interventions has been a major impetus to integrate mindfulness in medicine, healthcare, education and other institutions (Kabat-Zinn 2019; Stanley 2015; Wilson 2015). Along these lines, a Silver Santé researcher elucidated:

> The research we do in the Silver Santé Study, which is protocolled and a randomised controlled trial, has the advantage that it allows us to verify a hypothesis and to have legitimacy in comparison to pharmaceutical industries that will propose medicinal drugs. We can tell them: 'We realised a protocol that is as demanding as yours and we show positive results.'

Examination of the efficacy of meditation featured as an epistemic good in the Silver Santé Study that researchers were striving for by testing a hypothesis and adhering to a demanding study protocol. Doing good research meant producing results that could inform guideline development for the governance of healthcare.

The assessment of an intervention's efficacy is based on the assumption that the research hypothesis includes variables that are relevant for a particular study group and that one intervention works best for everyone (Dehue 2010). In this respect, however, one of the Siler Santé meditation instructors pointed out: "Meditation does not work for everybody all the time and at the same time." A meditation intervention has different effects on study participants depending on various factors, such as their world view, life circumstances, physical disposition and relationship to the meditation instructors. Meditation researchers Lutz et al. (2015) have cautioned against the negligence of context in clinical trials:

> [There] has been the need to frame mindfulness-based interventions in ways that are maximally compatible with clinical medicine and psychology, such that these practices are seen through the lens of current scientific thinking and are articulated in ways that can be readily communicated to potential patients, healthcare providers, and researchers. Although clearly crucial to basic and clinical research, this restricted perspective increases the risk of misrepresenting (or missing altogether) the active ingredients underlying the potentially

> transformative effects of these practices whose techniques emerge in a context broader than clinical medicine, psychology or neuroscience. (p. 633)

This critique of mindfulness-based clinical trial research underlines the significance of attention for the qualitative effects of meditation, instead of maintaining a narrow focus on efficacy (Farias and Wikholm 2015; Goleman and Davidson 2019).

Attending to an intervention's variable effects is not only a way to produce good knowledge (knowledge of which cognitive, affective and social processes might be altered throughout an intervention), but it may also give rise to a form of "care work" (Federici 2012, p. 368; Wadmann and Hoeyer 2014). Care work usually refers to the relational labour involved in holding communities together and generating conditions for "living as well as possible" (Puig de la Bellacasa 2017, p. 4). In acknowledging the affective effects of the meditation intervention on participants' well-being in the Age-Well RCT, the relevance of this kind of labour for doing good research came to the fore. The Silver Santé team wondered whether good research implied caring for participants after the end of the intervention—the phase in which researchers commonly "erase care" (Jespersen et al. 2014, p. 664) to direct their attention towards data analysis. Yet, performing extra care work to attend to the affective effects of meditation was in tension with assessing the intervention's efficacy. If the Silver Santé team had facilitated post-intervention meditation sessions, they would have biased the inquiry into efficacy in the 21-month follow-up examinations.

Caring informally while playing by formal rules

To take care of the affective effects of the meditation intervention without compromising the assessment of the intervention's efficacy, the Silver Santé team ultimately decided to play by the formal rules inscribed in the Age-Well protocol. Following a rule with fidelity does not contradict situational adjustments of its imposition; it also leaves room for courses of action not specified by the rule (Garfinkel 1967; Lynch 1993). Playing by formal rules meant that the meditation intervention of the first cohort of participants ended as planned in September 2018 (Table 2). The protocol prescribed that the weekly group sessions facilitated by Silver Santé meditation instructors should end after 18 months. It was not specified in the protocol, however, whether and in which way participants could continue meditating on their own.

The protocol left room for study participants to create an autonomous meditation group. Right after the end of the intervention, one participant managed to arrange a space, through the municipality of Caen, where the meditation group could meet on a biweekly basis for meditation sessions. This group invited the principal Silver Santé meditation instructor to attend a few sessions and teach members how to guide meditation practices for each other. The meditation instructor accepted this invitation in her personal capacity independent from the Silver Santé Study. By helping participants transition from a study group to a self-organised community of practice, she cared for participants' well-being after the end of the intervention while respecting the study protocol.

In this way, meditation was treated as an affective training whose effects were different from those of the English intervention and thus required the provision of "extra things" (Jespersen et al. 2014, p. 9) that were not included in the protocol. These extra things were part of an informal trial economy existing side by side with its formal procedures (Daston 1995). Playing by formal rules ensured that the official procedures specified in the study protocol were followed. At the same time, the research team, especially the meditation instructor, engaged in informal care practices that the protocol did not address. Doing informal care work while following official procedures made different epistemic goods compatible. The affective effects of the meditation intervention were taken care of without jeopardising the assessment of the intervention's efficacy.

Adjusting the protocol for a study task with expert meditators: objectivity and trained judgment

Event in a study task with expert meditators

The Age-Well protocol combines a three-armed RCT with an observational study on expert meditators (Lutz et al. 2018). Researchers sought to recruit 30 healthy participants aged 65 years or older who clocked at least 10.000 h of meditation in Vipassana, Dzogchen (Tibetan Buddhism) or Zen traditions to undergo a battery of study examinations on behavioural, neuroimaging, sleep and biological measures. The results of their examinations were supposed to be compared to those from participants of the Age-Well RCT. One aim of this cross-sectional study was to address a limitation of the RCT. Participants in the RCT learned two meditation practices over the course of the meditation intervention: mindfulness meditation (MM) and loving-kindness and compassion meditation (LKCM). As the RCT only assessed the combined effect of MM and LKCM, the observational study was supposed to help distinguish between the effects of MM and LKCM on brain functions, emotional reactivity and emotional regulation.

For this purpose, the battery of study examinations included an adapted version of the Socio-affective Video Task (SoVT) developed by Klimecki et al. (2013). The video task involved short silent video clips with high emotional content (human suffering in distressing situations) and low emotional content (people performing everyday activities). Novice meditators from the three-armed Age-Well RCT watched these videos while resting in the brain scanner. Expert meditators watched one set of videos in MM (relaxed openness to and awareness of any thought or feeling that arises) and another set in LKCM (generating feelings of loving kindness and compassion). Their brain activity was measured in response to the video clips. After the scanning session, participants partook in a debriefing in which they re-watched the videos and provided self-reports for the experience of each video in MM and LKCM. These self-reports involved ratings on a 0-to-10 scale of (1) empathy with the characters in the video, (2) positive emotions and (3) negative emotions (see Table 3: Standard SoVT debriefing).

Table 3 Standard and adjusted SoVT debriefing

Standard SoVT debriefing Three-armed Age-Well RCT with novice meditators	Adjusted SoVT debriefing Age-Well observational study with expert meditators
(1) At which intensity did you feel the emotions of the characters? (2) Indicate the intensity of your positive emotions (3) Indicate the intensity of your negative emotions	(1) At which intensity did you feel the emotions of the characters? (2) Indicate the intensity of your positive emotions (3) Indicate the intensity of your negative emotions (4) To which degree of openness were you available to experience the emotional content of the video? (5) To which degree of intensity were you upset and distressed while watching the emotional content of the video? (6) To which degree did you experience loving kindness and compassion towards the protagonists of the video?

The standard SoVT debriefing by Klimecki et al. (2013) was employed in the three-armed Age-Well RCT. The adjusted SoVT debriefing was developed by the Silver Santé team in response to expert meditators' feedback and was subsequently used in the Age-Well observational study.

I interviewed several expert meditators about their experience of the video task; one of them remembered it vividly:

> Expert meditator: The last session in the brain scanner was mind-blowing. This was my first experience of non-duality. Everything revealed itself to me although the videos shown were actually pretty sensory and, in part, pretty painful. Explicit suffering was depicted and, nevertheless, I had suddenly the feeling that it was not like that. Duality was completely dissolved.
> . . .
> Smolka: After your experience in the brain scanner, you talked to a researcher who asked you about positive and negative emotions. What did you answer?
> Expert meditator: I said that there was no positive or negative anymore.

The expert meditator's experience of non-duality escaped the protocol of the video task. She did not perceive a distinction between positive and negative emotions and, therefore, felt unable to rate her emotions in the debriefing. The protocol failed to standardise her subjective experience.

This case and similar incidences with other expert meditators gave rise to discussions among Silver Santé researchers about how to capture the subjective experience of the video task. One of the researchers suggested that expert meditators' trained judgment regarding their first-person experience of meditation could help improve the study task and data analysis. He considered long-term meditators as experts in observing and reporting their inner experience of meditation, which is why he was interested in collaborating with them. Other Silver Santé researchers, however, insisted that performing rigorous clinical trial research required objectivity, that is, adherence to standardised procedures. Changing the study task in response to expert meditators' trained judgment in the midst of data collection would impair objectivity

and could render SoVT data from Age-Well incomparable with SoVT data sets from other research groups.

Objectivity and trained judgment

The RCT has become the privileged form of knowledge production in biomedical research because it is deemed to produce claims about the efficacy of a drug or intervention on the basis of objective testing (Marks 1997). According to Cambrosio et al. (2006), the emergence of biomedicine in the 1950s has been accompanied by "regulatory objectivity", which ensures compatibility of measurements across laboratories and hospitals through norms and systems of collective production of evidence. Enacting regulatory objectivity in the Age-Well RCT meant weighing any change of the SoVT protocol against the loss of comparability of results with other studies that make use of this video task. The results of a single study have little meaning in isolation, but turn into solid evidence if compared with other findings. Regulatory objectivity seeks standards that allow for such comparisons and for cancelling out individual variation in subjective experiences. Its target is not so much the individual but a population of potential participants on whom a study task or an intervention could be used.

Regulatory objectivity incorporates elements of earlier forms of objectivity, in particular what historians of science call "mechanical objectivity" (Daston and Galison 2007, p. 18; Porter 1995, p. 4). Mechanical objectivity is synonymous with the exclusion of personal judgment through adherence to standardised methodological procedures. To understand why mechanical objectivity has become dominant not only in clinical research (Dehue 2001) but in the sciences and public life more broadly, Porter (1995) suggests to analyse the authority status of expert communities involved in knowledge production. Drawing on case studies of engineering, accountants and actuaries, and the rise of cost–benefit analysis, Porter argues that expert communities endorse mechanical objectivity whenever their authority becomes vulnerable. As long as their authority is considered legitimate by other scientists and the wider public, their knowledge claims rely on expert consensus. Accordingly, the rise of mechanical objectivity in clinical research is related to mistrust in subjectivity: scientific analyses could be biased by interests and clinical decision-making could be impaired by convictions (Dehue 1999).

Porter's explanation of the rise of mechanical objectivity in expert communities applies to meditation research. The first wave of meditation research in the 1970s was spurred by the Transcendental Meditation (TM) movement, which sought to validate the benefits of its meditation technique with scientific evidence. Meditation researchers like many of their colleagues in biomedicine, neuroscience and psychology consider TM research as "sloppy" pseudoscience (Farias and Wikholm 2015, p. 132; Harrington and Dunne 2015; Tøllefsen 2014). As of the early 2000s, second-wave meditation researchers have been trying to avoid such accusations by distancing themselves from spirituality, publishing in well-respected academic journals and carefully adhering to scientific standards, first and foremost the RCT (Kucinskas 2019).

Although objectivity is an important pursuit in meditation research, partly to consolidate the social basis of its authority, it is not coherent with another epistemic good that is at the heart of its scientific programme: "trained judgment". This concept was introduced by Daston and Galison (2007, p. 18) as an alternative practice of good science that relies on tacit expert knowledge (Collins 2010). While mechanical objectivity has banned both researchers' and participants' subjectivity (Danziger 1990), trained judgment makes knowledge claims based on subjective criteria. Bringing subjective experience back into science is a key objective of second-wave meditation researchers (Komjathy 2018). As one of their challenges is to conceptualise and operationalise meditation, they seek to collaborate with long-term meditators who have expertise in observing their inner experience and in reporting the activity of their minds (Wiles 2018). The mission of the Mind and Life Institute, one of the main drivers and sponsors of meditation research, is to foster "the dialogue between Western science and Buddhism" (Hasenkamp and White 2017, p. 7). Buddhist meditators participate in this dialogue as study participants and active scientific collaborators who help to refine research protocols and contribute to scientific analyses and publications (e.g., Singer 2017).

Still, ongoing collaboration with expert meditators in which researchers remain responsive to feedback and flexibly adapt their work throughout the research process challenges strict adherence to a protocol, thus interfering with objectivity in a clinical trial. As a Silver Santé researcher reflected on this interference in interactions with an expert meditator:

> Often there were situations—and I would totally agree with him there – where he, the [expert] meditator, would make a really fine point about the lack of definition in our task or in the way we present the test, but instead of really engaging in a collaborative investigation of where our conceptual frameworks might meet, we just say: 'Yeah, we get it, we understand, but could we just move on now?' And I agree. When I put my 'scientist head' on, I also just pushed him through.

As observed by the researcher, performing the scientific norm of objectivity did not allow him to be responsive to the expert meditator's trained judgment regarding distinctions between meditative states. Although attending to trained judgment could help to improve the task design so as to better capture experts' experiences of meditation, it would blur the lines between subjectivity and objectivity (e.g., Dor-Ziderman et al. 2013; Lutz et al. 2002; Winter et al. 2020). As the striving for objectivity fused with the "taboo against subjectivity" (Wallace 2007, p. 67) have a long-lasting history in Western science, meditation researchers consider the incorporation of trained judgment of expert meditators in their work a radical departure from traditional scientific inquiry (Lutz and Thompson 2003; Varela et al. 1997).

Adjusting the study protocol

Silver Santé researchers brought objectivity and trained judgment together by adjusting the study protocol of the aforementioned video task (SoVT) in the observational study with expert meditators. The task is based on the assumption

that study participants have an emotional response to videos with content of suffering that can be rated as positive and/or negative on a numerical scale (Engen and Singer 2015, 2016). Yet, this assumption did not hold for a number of expert meditators, as one of the Silver Santé researchers observed:

> One thing we realised, I think in the middle of the study, is that there are a couple of experts who have difficulties rating some of our scales in terms of valence [the extent to which an emotion is positive or negative] because they have more nuanced and complex ways to perceive things or images that do not fit our categories . . . Through mental training they have changed their worldview at such a level that it changed perception.

Some expert meditators commented on the videos that they experienced a lot of compassion without feeling any strong positive or negative emotions, especially when performing LKCM in the scanner. Nor did they perceive any distress in response to the suffering of the characters displayed in the videos. They might have changed the way they interpret and perceive suffering through long-term meditation practice (Dahl et al. 2015, 2016).

To find out whether expert meditators indeed perceive suffering in a way that escapes emotional valence, the Silver Santé team added three scale ratings to the SoVT debriefing (Table 3: Adjusted SoVT debriefing). Adjusting the SoVT protocol allowed them to investigate whether expert meditators experienced compassion (question 6) without strong positive/negative emotions (question 2 and 3) and/or distress (question 5), and whether these ratings were mediated by a particular interpretation of the video content (question 4). The protocol accommodated expert meditators' trained judgment in the research process, which opened up the possibility for new knowledge to emerge. While the protocol of the observational study with expert meditators was adjusted, the SoVT debriefing with novice meditators in the three-armed Age-Well RCT remained unchanged.

By limiting the adjustment of the study protocol to the observational study, researchers combined trained judgment with objectivity. They incorporated an emerging scientific insight into their work that resulted from trained judgment about how long-term meditation practice may change the experience of other people's suffering. At the same time, objectivity was warranted in the Age-Well RCT in which researchers and participants continued adhering to pre-defined procedures. The RCT is less flexible than the observational study because the former is supposed to provide conclusive results about the truth or falsity of a hypothesis, whereas the latter seeks to obtain novel biological and behavioural markers of meditation (Lutz et al. 2018). Despite its more flexible nature, researchers also paid attention to objectivity in the observational study. They developed a new standardised procedure for the SoVT debriefing that was followed in all ensuing study examinations with expert meditators. In this way, they collected objective data on how expert meditators in the state of rest, MM or LKCM subjectively experienced the videos.

Discussion

The adoption of the RCT standard in research on lifestyle interventions means that researchers continuously engage in balancing a variety of epistemic goods. My analysis of these balancing acts in the Age-Well trial of the Silver Santé Study underlines that a lifestyle intervention RCT is a labour-intensive, careful and situated achievement, which often remains hidden behind a standardisation machinery. To make diverging epistemic goods compatible, Silver Santé researchers tinkered with the study protocol by drawing on different practical strategies. I have not presented an exhaustive list of all the strategies mobilised by researchers to enact epistemic goods together. Rather, I have discussed three events that foreground dramatic situations in which researchers needed to find ways to pacify tensions so as to keep their day-to-day research going. These events disrupted ordinary research practices so much that the disruptions, the strategies for repair and the epistemic goods at stake became apparent and analysable. The strategies that researchers deployed to respond to these events were not exceptional, but could be identified in multiple scenes where practices diverged from standardised procedures.

Adjusting the protocol, reinterpreting its procedures and caring informally while playing by formal rules were some of the strategies that researchers took up on a regular basis to respond to more or less dramatic emergencies and surprises. What participation in the study interventions entailed was reinterpreted multiple times in both the English and the meditation intervention: at the occasion of participants' vacation requests, in cases of extended durations of illness, and when life took a toll on either participants or intervention instructors. There were also a few 'loopholes' in the protocol which allowed researchers to engage in informal care practices while playing by formal rules. For instance, the protocol did not specify whether participants could undergo psychotherapeutic treatment while being enrolled in the study. Meditation instructors could thus provide occasional psychological support for participants in times of trouble without breaching their professional task descriptions. In response to the emergence of new technologies and scientific advances, several official adjustments were made to the protocol, including the introduction of a new sleep monitoring device and an extra study task on memory consolidation.

When comparing these strategies, it becomes apparent that tinkering with the study protocol occurred formally and informally. Informal tinkering refers to the strategies that the Silver Santé team used to tinker with the imposition of the study protocol through reinterpreting its procedures and engaging in care practices that left written rules untouched. Formal tinkering, by contrast, means adjusting the official procedure inscribed in the protocol. Whereas care and social dynamics remain in the informal "back stage" of a clinical trial, tinkering is performed "front stage" (Goffman 1959, p. 115) if science is at stake. This means that responding flexibly to unexpected events in the research process is not necessarily "invisible work" (Wolters et al. 2018; see also Shapin 1989; Star and Strauss 1999). It becomes visible in the study protocol and related scientific publications if researchers recognise its relevance for epistemology.

The event in a study task with expert meditators put science at stake because it revealed that meditation as defined in the study protocol differed from meditation as experienced by expert meditators in the brain scanner. In consequence, Silver Santé researchers adjusted the protocol for expert meditators, so that it became redefined as a contemplative practice deserving further investigation because it could give rise to extraordinary states, such as the experience of non-duality. By leaving such states unexamined for participants in the Age-Well RCT, meditation was researched as a behavioural lifestyle intervention that could enhance cognition and emotional regulation. Through formal tinkering, Silver Santé researchers came to study different versions or "multiple objects" (Mol 2002, p. 5) of meditation. While this finding is not surprising for STS researchers who have long insisted that science is not simply representational, but generates new entities or modifies existing ones (Beaulieu 2001; Latour and Woolgar 1987; Law 2002; Pickering 1992), it could be relevant for the Silver Santé Study. After all, it calls into question whether the effects of meditation on brain function as captured in the Age-Well RCT versus those recorded in the observational study with expert meditators are actually comparable.

Conclusion

This article has explored the relevance of the concept of epistemic goods for STS literature on clinical trial research. Specifically, I have introduced epistemic goods as a theoretical lens to capture knowledge-making practices and their normative dimensions. Prior social studies of clinical trials scrutinised frictions between fidelity to the study protocol and adaptation to the local context of implementation (e.g., Cohn and Lynch 2017; Winther and Hillersdal 2020; Rod et al. 2014; Timmermans and Berg 2003; Timmermans 2010; Will 2007). By looking at these frictions through the lens of epistemic goods, it becomes visible that they emerge from the enactment and negotiation of multiple, partly incoherent norms of doing good research.

The contextual, situated accomplishment of scientific norms does not imply that they are arbitrary. Instead, their co-existence stems from larger historical developments: the turn towards cultural specificity in the international clinical trial industry (Brives et al. 2016; Rosemann 2019), the shift from generalisable to personalised knowledge in precision medicine (Au 2020) and in an emerging "precision science of meditation" (Schlosser et al. 2022, p. 11), as well as the revived recognition of expert judgment in biomedicine (Cambrosio et al. 2006). These developments have not led to radical transformations, but have allowed seemingly incoherent epistemic goods to be present together (cf. Daston and Galison 2007). The historical origin of epistemic goods and how they become internalised through socialisation processes are topics for further investigation that go beyond the scope of this article. What is important to recognise here is that the definition of good research emerges in the relational interactions between scientists, other members of a research team, study participants, trial protocols, scientific methods, technologies and concepts *in the process of knowledge production*. This insight has implications for the evaluation of scientific evidence in guideline development.

The concept of epistemic goods helps problematise the notion of evidence in evidence-based medicine and healthcare (cf. Knaapen 2013). Critics have pointed out that evidence-based medicine prioritises knowledge that appeals to internal validity, the assessment of efficacy and objectivity (Knaapen 2014; Lambert 2006; Wahlberg and McGoey 2007). As RCTs and their systematic reviews embody these epistemic goods, they constitute the "gold standard" (Timmermans and Berg 2003) of robust evidence production. Other kinds of knowledge and the epistemic goods enacted in their construction tend to be undervalued in guideline development (Wieringa et al. 2018; Zuiderent-Jerak 2012). Following up on these critiques, the concept of epistemic goods reveals that what counts as robust evidence cannot be tied to an a priori defined standard.

In line with Gomart and Hajer's (2003) reading of science study's philosophy of good experiments (in particular Stengers 1993), I argue that it is impossible to define universal criteria of robust evidence if we consider research settings as sites of emergence where "no one knows beforehand what are the essences, and therefore the vulnerabilities or the resistances of the entities [e.g., research participants] that pass through a setting" (Gomart and Hajer 2003, p. 39). Neither can one fix once and for all the essence of a good clinical study. This does not mean that there is no 'good' but that each study proposes a new definition of what good might be. This definition is contingent on the interplay between the research methodology (e.g., case study, quasi-experimental research, RCT), study object (e.g., pharmacological product, technological device, behavioural intervention), research participants (e.g., college students, healthy elderly adults, patients) and situated knowledge-making practices (e.g., adhering to protocolled procedures, attending to effects, incorporating trained judgment). Therefore, evidence claims must be placed in the context of their production to be evaluated on their own terms.

For this purpose, I seek to introduce an empirical ethics approach (Pols 2015) to the evaluation of clinical trial research. In evaluating knowledge claims for the development of guidelines and policies, professionals have been shown to be reflexive about the narrow definition of evidence in evidence-based medicine and to pay attention to alternative (non-epistemic) considerations (Boswell 2017; Moreira 2005; Stewart and Smith 2015; Verkerk et al. 2006; Zuiderent-Jerak 2021). As observed by Lagerlöf et al. (2021), however, the tenets of evidence-based medicine prevail in the development process of national guidelines on lifestyle habits in Sweden. They further propose: "If considerations pertaining to public health and healthcare are to be integrated more firmly into the National Guidelines, methodological rigour needs to be complemented with a wider latitude for epistemological deliberation" (p. 16). I propose that an empirical ethics approach could widen the latitude for epistemological deliberation through empirically informed descriptions of locally configured knowledge-making practices and their normative dimensions.

An empirical ethics approach does not take local normativities to carry prescriptive force, but helps questioning what counts as good evidence by comparing the epistemic goods identified somewhere with those found elsewhere. Empirically described goods could further be compared with oft-tacit knowledge evaluation criteria in healthcare policy and guideline development processes so as to scrutinise and revise them. Along these lines, the results of this study suggest that the

epistemological deliberation on internal validity, the assessment of efficacy and objectivity in knowledge evaluations of RCTs could be widened by addressing whether a trial is socially relevant, attends to effects and incorporates trained judgment. The epistemic goods identified here could play a role in the appraisal of knowledge for developing health promotion guidelines, disease prevention programmes and public health measures.

Acknowledgements This research benefitted from funding from the German Academic Scholarship Foundation. I wish to thank the Medit-Ageing research group and the participants in the Silver Santé Study for contributing to this research. I am also grateful for the comments by Cornelius Borck, Gili Yaron, Cyrus Mody, Geert Somsen and the anonymous reviewers on earlier versions of this article.

Declarations

Conflict of interest The author states that there is no conflict of interest.

Ethical approval The research received ethical approval from the Ethical Review Committee Inner City Faculties of Maastricht University.

References

Au, L. 2020. Imagining the public: Anticipatory discourses in China's push for precision medicine. *BioSocieties*. https://doi.org/10.1057/s41292-020-00205-5.

Beaulieu, A. 2001. Voxels in the brain: Neuroscience, informatics and changing notions of objectivity. *Social Studies of Science* 31 (5): 635–680.

Beaulieu, A. 2002. Mediating ethnography: Objectivity and the making of ethnographies of the internet. *Social Epistemology* 18 (2–3): 139–163.

Beaulieu, A. 2010. From co-location to co-presence: Shifts in the use of ethnography for the study of knowledge. *Social Studies of Science* 40 (3): 453–470.

Berg, M. 1998. Order(s) and disorder(s): Of protocols and medical practice. In *Differences in medicine. Unraveling practices, techniques, and bodies*, ed. M. Berg and A. Mol, 226–246. Durham: Duke University Press.

Bonell, C., A. Fletcher, M. Morton, T. Lorenc, and L. Moore. 2012. Realist randomised controlled trials: A new approach to evaluating complex public health interventions. *Social Science & Medicine* 75: 2299–2306.

Boswell, J. 2017. What makes evidence-based policy making such a useful myth? The case of NICE guidance on bariatric surgery in the United Kingdom. *Governance* 31 (2): 199–214.

Bowker, G.C., and S.L. Star. 1999. *Sorting things out: Classification and its consequences*. Cambridge: MIT Press.

Brives, C., F. Le Marcis, and E. Sanabria. 2016. What's in a context? Tenses and tensions in evidence-based medicine. *Medical Anthropology* 35 (5): 369–376.

Cambrosio, A., P. Keating, T. Schlich, and G. Weisz. 2006. Regulatory objectivity and the generation and management of evidence in medicine. *Social Science & Medicine* 63: 189–199.

Cohn, S., and R. Lynch. 2017. Falling into a routine: From habits to situated practices. *Sociology of Health & Illness* 39 (8): 1398–1411.

Collins, H.M. 1992. *Changing order: Replication and induction in scientific practice*. Chicago: University of Chicago Press.

Collins, H.M. 2010. *Tacit and Explicit Knowledge*. Chicago: The University of Chicago Press.

Dahl, C.J., A. Lutz, and R.J. Davidson. 2015. Reconstructing and deconstructing the self: Cognitive mechanisms in meditation practice. *Trends in Cognitive Sciences* 19 (9): 515–523.

Dahl, C.J., A. Lutz, and R.J. Davidson. 2016. Cognitive processes are central in compassion meditation. *Trends in Cognitive Sciences* 20 (3): 161–162.

Danziger, K. 1990. *Constructing the subject. Historical origins of psychological research*. Cambridge: Cambridge University Press.
Danziger, K. 2000. Making social psychology experimental: A conceptual history, 1920–1970. *Journal of the History of Behavioural Sciences* 36 (4): 329–347.
Daston, L. 1995. The moral economy of science. *Osiris* 10: 2–24.
Daston, L., and P. Galison. 2007. *Objectivity*. Brooklyn: Zone Books.
Davidson, R.J., and A.W. Kaszniak. 2015. Conceptual and methodological issues in research on mindfulness and meditation. *American Psychologist* 70 (7): 581–592.
Dehue, T. 1999. Testing treatments, managing life: On the history of randomized clinical trials. *History of the Human Sciences* 12 (1): 115–124.
Dehue, T. 2001. Establishing the experimenting society: The historical origin of social experimentation according to the randomized controlled design. *American Journal of Psychology* 2 (114): 283–302.
Dehue, T. 2002. A Dutch treat: Randomized controlled experimentation and the case of heroin-maintenance in the Netherlands. *History of the Human Sciences* 15 (2): 75–98.
Dehue, T. 2010. Comparing artificial groups: On the history and assumptions of the randomised controlled trial. In *Medical proofs, social experiments: clinical trials in shifting contexts*, ed. C. Will and T. Moreira, 103–119. Farnham: Ashgate.
Dor-Ziderman, Y., A. Berkovich-Ohana, J. Glicksohn, and A. Goldstein. 2013. Mindfulness-induced selflessness: A MEG neurophenomenological study. *Frontiers in Human Neuroscience* 7: 582.
Engen, H.G., and T. Singer. 2015. Compassion-based emotion regulation up-regulates experienced positive affect and associated neural networks. *Social Cognitive and Affective Neuroscience* 10 (9): 1291–1301.
Engen, H.G., and T. Singer. 2016. Affect and motivation are critical in constructive meditation. *Trends in Cognitive Sciences* 20 (3): 159–160.
Farias, M., and C. Wikholm. 2015. *The Buddha Pill. Can meditation change you?* London: Watkins.
Federici, S. 2012. *Revolution at point zero: Housework, reproduction, and feminist struggle*. Oakland: PM Press.
Fisher, J.A. 2006. Co-ordinating 'ethical' clinical trials: The role of research coordinators in the contract research industry. *Sociology of Health & Illness* 28 (6): 678–694.
Friese, C. 2013. Realizing potential in translational medicine. The uncanny emergence of care as science. *Current Anthropology* 54 (7): 129–138.
Fujimura, J.H. 1987. Constructing 'Do-Able' Problems in cancer research: Articulating alignment. *Social Studies of Science* 17 (2): 257–293.
Garfinkel, H. 1967. *Studies in ethnomethodology*. Englewood Cliffs: Prentice Hall.
Garforth, L. 2012. In/visibilities of research: Seeing and knowing in STS. *Science, Technology, & Human Values* 37 (2): 264–285.
Goffman, E. 1959. *The presentation of self in everyday life*. New York: Anchor.
Goleman, D., and R.J. Davidson. 2019. *Altered traits Science reveals how meditation changes your mind, brain, and body*. New York: Penguin Random House.
Gomart, E., and M. Hajer. 2003. Is *That* politics? For an inquiry into forms in contemporary politics. In *Social studies of science and technology: Looking back ahead*, ed. B. Joerges and H. Nowotny, 33–61. Amsterdam: Kluwer Academic Publishers.
Gotink, R.A., P. Chu, J.J.V. Busschbach, H. Benson, G.L. Fricchione, and M.G. Hunink. 2015. Standardised mindfulness-based interventions in healthcare: An overview of systematic reviews and meta-analyses of RCTs. *PLoS ONE* 10: 1–17.
Green, A., and K. Kolar. 2015. Engineering behaviour change in an epidemic: The epistemology of NIH-funded HIV prevention science. *Sociology of Health & Illness* 37 (4): 561–577.
Hallowell, N., S. Cooke, G. Crawford, A. Lucassen, and M. Parker. 2009. Distinguishing research from clinical care in cancer genetics: Theoretical justifications and practical strategies. *Social Science & Medicine* 68 (11): 2010–2017.
Harrington, A., and J.D. Dunne. 2015. When mindfulness is therapy: Ethical qualms, historical perspectives. *American Psychologist* 70 (7): 621–631.
Hasenkamp, W., and J.R. White. 2017. Introduction. In *The monastery and the microscope*, ed. W. Hasenkamp and J.R. White, 1–18. London: Yale University Press.
Hauskeller, C., N. Baur, and J. Harrington. 2019. Standards, harmonization and cultural differences: Examining the implementation of a European stem cell clinical trial. *Science as Culture* 28 (2): 174–199.

Hogle, L.F. 1995. Standardization across non-standard domains: The case of organ procurement. *Science, Technology, & Human Values* 20 (4): 482–500.
Holman, D., R. Lynch, and A. Reeves. 2018. How do health behaviour interventions take account of social context? A literature trend and co-citation analysis. *Health* 22 (4): 389–410.
Jasanoff, S., ed. 2004. *States of knowledge: The co-production of science and social order*. Oxford: Routledge.
Jespersen, A.P., J. Bønnelycke, and H.H. Eriksen. 2014. Careful science? Bodywork and care practices in randomised clinical trials. *Sociology of Health & Illness* 36 (5): 655–669.
Jonvallen, P. 2005. Testing pills, enacting objectivity: The work of localizing tools in a clinical trial. PhD dissertation, Linköping University, Sweden.
Kabat-Zinn, J. 2011. Some reflections on the origins of MBSR, skillful means, and the trouble with maps. *Contemporary Buddhism* 12: 281–306.
Kabat-Zinn, J. 2019. Foreword: Seeds of a necessary global renaissance in the making: the refining of psychology's understanding of the nature of mind, self, and embodiment through the lens of mindfulness and its origins at a key inflection point for the species. *Current Opinion in Psychology* 28: xi–xvii.
Katz, J. 2001. From how to why: On luminous description and causal inference in ethnography (part 1). *Ethnography* 2 (4): 443–473.
Keating, P., and A. Cambrosio. 2007. Cancer Clinical Trials: The Emergence and Development of a New Style of Practice. *Bulletin of the History of Medicine* 81 (1): 197–223.
Klimecki, O.M., S. Leiberg, M. Ricard, and T. Singer. 2013. Functional neural plasticity and associated changes in positive affect after compassion training. *Cerebral Cortex* 23: 1552–1561.
Klimecki, O.M., N.L. Marchant, A. Lutz, G. Poisnel, G. Chételat, and F. Collette. 2019. The impact of meditation on healthy ageing—The current state of knowledge and a roadmap to future directions. *Current Opinion in Psychology* 28: 223–228.
Knaapen, L. 2013. Being 'evidence-based' in the absence of evidence: The management of non-evidence in guideline development. *Social Studies of Science* 43 (5): 681–706.
Knaapen, L. 2014. Evidence-based medicine or cookbook medicine? Addressing concerns over the standardization of care. *Sociology Compass* 8 (6): 823–836.
Komjathy, L. 2018. *Introducing contemplative studies*. Hoboken: Wiley.
Kucinskas, J. 2019. *The mindful elite. Mobilizing from the inside out*. Oxford: Oxford University Press.
Lagerlöf, H., T. Zuiderent-Jerak, and M. Sager. 2021. Epistemological deliberation: The challenges of producing evidence-based guidelines on lifestyle habits. *Evidence & Policy*. https://doi.org/10.1332/174426421X16149619907286.
Lambert, H. 2006. Accounting for EBM: Notions of evidence in medicine. *Social Science & Medicine* 62: 2633–2645.
Lampland, M., and S.L. Star. 2009. *Standards and their stories: How quantifying, classifying, and formalizing practices shape everyday life*. Ithaca: Cornell University Press.
Lappé, M. 2018. The paradox of care in behavioral epigenetics: Constructing early-life adversity in the lab. *BioSocieties* 13 (4): 698–714.
Latimer, J., and M. Puig de la Bellacasa. 2013. Re-thinking the ethical: Everyday shifts of care in biogerontology. In *Ethics, law and society*, ed. N. Priaulx and A. Wrigley, 153–174. London: Routledge.
Latour, B. 1983. Give me a laboratory and I will raise the world. In *Science observed: Perspectives on the social study of science*, ed. K. Knorr-Cetina and M. Mulkay, 141–170. London: Sage.
Latour, B., and S. Woolgar. 1987. *Laboratory life*. Princeton: Princeton University Press.
Law, J. 2002. *Aircraft stories: Decentering the object in technoscience*. Durham: Duke University Press.
Lutz, A., A.P. Jha, J.D. Dunne, and C.D. Saron. 2015. Investigating the phenomenological matrix of mindfulness-related practices from a neurocognitive perspective. *American Psychologist* 70 (7): 632–658.
Lutz, A., O.M. Klimecki, F. Collette, G. Poisnel, E. Arenaza-Urquijo, N.L. Marchant, V. de La Sayette, et al. 2018. The age-well observational study on expert meditators in the Medit-Ageing European project. *Alzheimer's & Dementia* 4: 756–764.
Lutz, A., J.-P. Lachaux, J. Martinerie, and F.J. Varela. 2002. Guiding the study of brain dynamics by using first-person data: Synchrony patterns correlate with ongoing conscious states during a simple visual task. *Proceedings of the National Academy of Sciences* 99 (3): 1586–1591.
Lutz, A., and E. Thompson. 2003. Neurophenomenology integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies* 10: 31–52.

Lynch, M. 1993. *Scientific practice and ordinary action. Ethnomethodology and social studies of science*. Cambridge: Cambridge University Press.
Mannell, J., and K. Davis. 2019. Evaluating complex health interventions with randomized controlled trials: How do we improve the use of qualitative methods. *Qualitative Health Research* 29 (5): 623–631.
Marks, H.M. 1997. *The progress of experiment: Science and therapeutic reform in the United States, 1900–1990*. New York: Cambridge University Press.
Mesman, J. 2008. *Uncertainty in medical innovation. Experienced pioneers in neonatal care*. New York: Palgrave Macmillan.
Mol, A. 2002. *The body multiple: Ontology in medical practice*. Durham: Duke University Press.
Mol, A., I. Moser, and J. Pols, eds. 2010. *Care in practice: On tinkering in clinics, homes and farms*. Bielefeld: Transcript Verlag.
Montgomery, C.M. 2016. From standardization to adaptation: Clinical trials and the moral economy of anticipation. *Science as Culture* 26 (1): 232–254.
Moreira, T. 2005. Diversity in clinical guidelines: The role of repertoires of evaluation. *Social Science & Medicine* 60: 1975–1985.
Ong, B.N., A. Rogers, A. Kennedy, P. Bower, T. Sanders, A. Morden, S. Cheraghi-Sohi, J.C. Richardson, and F. Stevenson. 2014. Behavior change and social blinkers? The role of sociology in trails of self-management behaviour in chronic conditions. *Sociology of Health & Illness* 2: 226–238.
Pickering, A.R., ed. 1992. *Science as practice and culture*. Chicago: Chicago University Press.
Pinel, C., B. Prainsack, and C. McKevitt. 2020. Caring for data: Value creation in a data-intensive research laboratory. *Social Studies of Science* 50 (2): 175–197.
Poisnel, G., E. Arenaza-Urquijo, F. Collette, O.M. Klimecki, N.L. Marchant, M. Wirth, V. de La Sayette, et al. 2018. The Age-Well randomized controlled trial of the Medit-Ageing European project: Effect of meditation or foreign language training on brain and mental health in older adults. *Alzheimer's & Dementia* 4: 714–723.
Pols, J. 2008. Which empirical research, whose ethics? Articulating ideals in long-term mental health. In *Empirical ethics in psychiatry*, ed. G. Widdershoven, J. McMillan, T. Hope, and L. van der Scheer, 51–67. Oxford: Oxford University Press.
Pols, J. 2013. The chronification of illness. Empirical ethics in care. Inaugural lecture at University of Amsterdam. Amsterdam, 27 September.
Pols, J. 2015. Towards an empirical ethics in care: Relations with technologies in health care. *Medicine, Health Care and Philosophy* 18: 81–90.
Pols, J. 2018. A reader's guide to the anthropology of ethics and morality—Part III. Somatosphere, 1 November, http://somatosphere.net/2018/a-readers-guide-to-the-anthropology-of-ethics-and-morality-part-iii.html/. Accessed 23 July 2021.
Pols, J., B. Pasveer, and D. Willems. 2018. The particularity of dignity: Relational engagement in care at the end of life. *Medicine, Health Care and Philosophy* 21 (1): 89–100.
Porter, T. 1995. *Trust in numbers. The pursuit of objectivity in science and public life*. Princeton: Princeton University Press.
Puig de la Bellacasa, M. 2011. Matters of care in technoscience: Assembling neglected things. *Social Studies of Science* 41 (1): 85–106.
Puig de la Bellacasa, M. 2017. *Matters of care. Speculative ethics in more than human worlds*. London: University of Minnesota Press.
Rod, M.H., L. Ingholt, B.B. Sørensen, and T. Tjørnhøj-Thomsen. 2014. The spirit of the intervention: Reflections on social effectiveness in public health intervention research. *Critical Public Health* 24 (3): 296–307.
Rogers, A., A. Kennedy, E. Nelson, and A. Robinson. 2005. Uncovering the limits of patient-centeredness: Implementing a self-management trial for chronic illness. *Qualitative Health Research* 15 (2): 224–239.
Rosemann, A. 2019. Alter-standardizing clinical trials: The gold standard in the crossfire. *Science as Culture* 28 (2): 125–148.
Rosenkranz, M.A., J.D. Dunne, and R.J. Davidson. 2019. The next generation of mindfulness-based intervention research: What have we learned and where are we headed. *Current Opinion in Psychology* 28: 179–183.
Schlosser, M., T. Bahnhofer, F. Requier, Y.I. Deza-Araujo, O. Abdoun, N.L. Marchant, G.C. Chételat, et al. 2022. Measuring psychological mechanisms in meditation practice: Using a

phenomenologically grounded classification system to develop theory-based composite scores. *Mindfulness*. https://doi.org/10.1007/s12671-021-01816-0.
Shapin, S. 1989. The invisible technician. *American Scientist* 77 (6): 554–563.
Singer, T. 2017. I feel your pain: The social neuroscience of empathy. In *The monastery and the microscope*, ed. W. Hasenkamp and J.R. White, 234–269. London: Yale University Press.
Stanley, S. 2015. *Sila* and *Sati:* An exploration of ethics and Mindfulness in Pāli Buddhism and their implications for secular mindfulness-based applications. In *Buddhist foundations of mindfulness*, ed. E. Shonin, W. van Gordon, and N.N. Sing, 89–113. New York: Springer.
Star, S.L., and A. Strauss. 1999. Layers of silence, arenas of voice: The ecology of visible and invisible work. *Computer Supported Cooperative Work* 8: 9–30.
Stengers, I. 1993. *L'invention des sciences modernes*. Paris: La Decouverte.
Stewart, E., and K. Smith. 2015. 'Black magic' and 'gold dust': The epistemic and political uses of evidence tools in public health policy making. *Evidence and Policy* 11 (3): 415–437.
Streiner, D.L. 2002. The 2 'Es' of research: Efficacy and effectiveness trials. *The Canadian Journal of Psychiatry* 47 (6): 552–556.
Swallow, J., A. Kerr, C.K. Chekar, and S. Cunninham-Burley. 2020. Accomplishing an adaptive clinical trial for cancer: Valuation practices and care work across the laboratory and the clinic. *Social Science & Medicine* 252 (112949): 1–8.
Timmermans, S. 2010. Reconciling research with medical care in RCTs. In *Medical proofs, social experiments: Clinical trials in shifting contexts*, ed. C. Will and T. Moreira, 17–32. Farnham: Ashgate.
Timmermans, S. 2014. Trust in standards: Transitioning clinical exome sequencing from bench to bedside. *Social Studies of Science* 45 (1): 77–99.
Timmermans, S., and M. Berg. 2003. *The gold standard. The challenge of evidence-based medicine and standardization in health care*. Philadelphia: Temple University Press.
Timmermans, S., and S. Epstein. 2010. A world of standards but not a standard world: Toward a sociology of standards in standardization. *Annual Review of Sociology* 36: 69–89.
Timmermans, S., and I. Tavory. 2012. Theory construction in qualitative research: From grounded theory to abductive analysis. *Sociological Theory* 30 (3): 167–186.
Tøllefsen, I.B. 2014. Transcendental meditation, the art of living foundation, and public relations. In *Controversial new religions*, ed. J. Lewis and J. Petersen, 159–175. New York: Oxford University Press.
Vago, D.R., R.S. Gupta, and S.W. Lazar. 2019. Measuring cognitive outcomes in mindfulness-based intervention research: A reflection on confounding factors and methodological limitations. *Current Opinion in Psychology* 28: 143–150.
van Dam, N.T., M.K. van Vugt, D.R. Vago, L. Schmalzl, C.D. Saron, A. Olendzki, T. Meissner, et al. 2018. Mind the hype: A critical evaluation and prescriptive agenda for research on mindfulness and meditation. *Perspectives on Psychological Sciences* 13 (1): 36–61.
Varela, F.J., E. Thompson, and E. Rosch. 1997. *The embodied mind. Cognitive science and human experience*. Cambridge: The MIT Press.
Verkerk, K., H. van Veenendaal, J.L. Severens, E.J.M. Hendriks, and J.S. Burgers. 2006. Considered judgment in evidence-based guideline development. *International Journal for Quality in Health Care* 18 (5): 365–369.
Wadmann, S., and K. Hoeyer. 2014. Beyond the 'therapeutic misconception': Research, care and moral friction. *BioSocieties* 9 (1): 3–23.
Wahlberg, A., and L. McGoey. 2007. An elusive evidence base: The construction and governance of randomized controlled trials. *BioSocieties* 2: 1–10.
Wallace, A. 2007. *Contemplative science. Where Buddhism and neuroscience converge*. New York: Columbia University Press.
Webster, A., C. Haddad, and C. Waldby. 2011. Experimental heterogeneity and standardisation: Stem cell products and the clinical trial process. *BioSocieties* 6 (4): 401–419.
Wells, M., B. Williams, S. Treweek, J. Coyle, and J. Taylor. 2012. Intervention description is not enough: Evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. *Trials* 13: 95.
Wieringa, S., D. Dreesens, F. Forland, C. Hulshof, S. Lukersmith, F. Macbeth, B. Shaw, A. van Vliet, and T. Zuiderent-Jerak. 2018. Different knowledge, different styles of reasoning: A challenge for guideline development. *BMJ Evidence-Based Medicine* 23: 87–91.

Winter, U., P. LeVan, T.L. Borghardt, B. Akin, M. Wittmann, Y. Leyens, and S. Schmidt. 2020. Content-free awareness: EEG-fcMRI correlates of consciousness *as Such* in an expert meditator. *Frontiers in Psychology* 10: 3064.

Wiles, K. 2018. Embodied practitioner knowledge: Reconfiguring the boundary between science and meditation. *Science as Culture* 28 (4): 449–469.

Will, C.M. 2007. The alchemy of clinical trials. *BioSocieties* 2: 85–99.

Will, C.M. 2010. Addressing the commercial context: The collective critique of clinical trials. In *Medical proofs, social experiments: Clinical trials in shifting contexts*, ed. C. Will and T. Moreira, 67–84. Farnham, Surrey: Ashgate.

Will, C.M., and T. Moreira, eds. 2010. *Medical proofs, social experiments. Clinical trials in shifting contexts*. Farnham: Ashgate.

Wilson, J. 2015. *Mindful America. The mutual transformation of Buddhist Meditation and American culture*. Oxford: Oxford University Press.

Winther, J., and L. Hillersdal. 2020. Balancing methodological purity and social relevance: Monitoring participant compliance in a behavioural RCT. *BioSocieties* 15: 555–579.

Wolters, A., G. de Weert, O.C.P. van Schayck, and K. Horstman. 2018. Invisible work, actors, and knowledge: An analysis of a clinical trial for a vaccine to stop smoking. *BioSocieties* 15: 1–27.

Zuiderent-Jerak, T. 2012. Guidelines should reflect all knowledge, not just clinical trials. *BMJ Evidence-Based Medicine* 36 (4): 392–393.

Zuiderent-Jerak, T. 2021. STS as a third space between evidence-based medicine and the human sciences. In *Making & doing. Activating STS through knowledge expression and travel*, ed. G.L. Downey and T. Zuiderent-Jerak, 197–217. Cambridge: The MIT Press.

Mareike Smolka is a PhD candidate at the Faculty of Arts and Social Sciences of Maastricht University in the Netherlands and an affiliated researcher at the Institute of the History of Medicine and Science Studies of University Lübeck in Germany. She is interested in STS, empirical ethics, and responsible innovation.