PEC Innovation 3 (2023) 100208

Contents lists available at ScienceDirect

PEC Innovation

journal homepage: www.elsevier.com/locate/pecinn

Is it fun or is it hard? Studying physician-related attributes of shared decision-making by ranking case vignettes

Laura Spinnewijn [a,b,*], Johanna Aarts [c], Didi Braat [a], Nikolaj Baranov [a], Klaas Sijtsma [d], Jules Ellis [e], Fedde Scheele [b,f]

[a] Radboud University Medical Center, Department of Obstetrics and Gynaecology, Nijmegen, the Netherlands
[b] VU University, Athena Institute for Trans-Disciplinary Research, Amsterdam, the Netherlands
[c] Amsterdam University Medical Centers, Department of Obstetrics and Gynaecology, Amsterdam, the Netherlands
[d] Tilburg University, Department of Methodology and Statistics, Tilburg School of Social and Behavioral Sciences, Tilburg, the Netherlands
[e] Radboud University, Behavioural Science Institute, Nijmegen, the Netherlands
[f] Amsterdam University Medical Centers, School of Medical Sciences, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Complexity
Job satisfaction
Gynaecologists
Non-technical problems
Shared decision-making
Culture change

ABSTRACT

Objective: This study investigated provider-related attributes of shared decision-making (SDM). It studied how physicians rank SDM cases compared to other cases, taking 'job satisfaction' and 'complexity' as ranking criteria.
Methods: Ten vignettes representing three cases of SDM, three cases dealing with patients' emotions and four with technical problems were designed to conduct a modified ordinal preference elicitation study. Gynaecologists and trainees ranked the vignettes for 'job satisfaction' or 'complexity'. Results were analysed by comparing the top three and down three ranked cases for each type of case using exact *p*-values obtained with custom-made randomisation tests.
Results: Participants experienced more satisfaction significantly from performing technical cases than cases dealing with emotions or SDM. Moreover, technical cases were perceived as less complex than those dealing with emotions. However, results were inconclusive about whether gynaecologists find SDM complex.
Conclusion: Findings suggest gynaecologists experience lower satisfaction with SDM tasks, possibly due to them falling outside their comfort zone. Integrating SDM into daily routines and promoting culture change favouring dealing with non-technical problems might help mitigate issues in SDM implementation.
Innovation: Our novel study assesses SDM in the context of task appraisal, illuminating the psychology of health professionals and providing valuable insights for implementation science.

1. Introduction

Over the past two decades, shared decision-making (SDM) has emerged as a prominent approach in healthcare, advocating for collaborative decision-making between healthcare providers and patients. It involves doctors and patients sharing the best available evidence, engaging in meaningful discussions about treatment options, and ultimately reaching a consensus based on the patient's values, preferences, and unique clinical circumstances [1]. SDM is considered an essential additional component of evidence-based medicine, enriching decision-making by actively involving patients in their healthcare choices [2,3]. Research suggests that SDM can lead to better patient understanding of available options, improved perception of potential benefits and risks, and ultimately, decisions that better align with patients' values and preferences [4,5].

Although promising as a concept and despite supportive measures like decision aids and training in decision-making, implementing SDM has been a laborious, challenging and time-consuming process [6-8]. Moreover, SDM utilization is perceived as incomplete or inconclusive [9-11], even though advice on overcoming obstacles for successful implementation seems readily available [8,12,13]. Therefore, whether previous studies on (surmounting) barriers to implementation have considered all crucial determinants for success remains to be debated.

A recent review on barriers to SDM implementation could be helpful, as it considers those barriers from a slightly different perspective. It stratifies obstacles according to the system, provider and patient levels,

* Corresponding author at: Radboud University Medical Center, Department of Obstetrics and Gynaecology, Post-box 9101, 6500 HB Nijmegen, the Netherlands.
E-mail address: laura.spinnewijn@radboudumc.nl (L. Spinnewijn).

https://doi.org/10.1016/j.pecinn.2023.100208
Received 17 March 2023; Received in revised form 29 August 2023; Accepted 30 August 2023
Available online 31 August 2023

with most challenges occurring at the provider level [14]. Therefore, it might be imperative to consider those provider-level barriers in particular. One relevant perspective for investigating provider-level barriers is Everet Rogers' Diffusion of Innovations (DOI) theory. As an influential author in the change management literature, Rogers extensively explores the spread of new ideas and technologies, also known as "innovations," and how attributes such as complexity can influence their successful adoption by individual users [15]. Complexity, in the context of DOI, refers to the perceived difficulty experienced by users in understanding or using an innovation [15]. Rogers contends that an innovation's complexity is negatively related to its adoption rate, making it a pertinent attribute to examine in the context of SDM.

However, predicting the perceived difficulty of SDM in healthcare is challenging due to the multitude of factors influencing healthcare encounters. Lorenzetti et al. highlight that challenging encounters can be attributed to various factors associated with physicians, patients, situational context, or a combination of these [16]. Additionally, Self-Determination Theory (SDT), a renowned psychological theory, posits that perceived difficulty is interdependently related to an individual's sense of autonomy and competence [17]. If an individual perceives a task as overly challenging or beyond their perceived competence levels, they may experience greater difficulty and reduced motivation or satisfaction [17].

Moreover, physicians engaged in challenging patient-clinician encounters have been shown to experience higher stress levels, increased burnout, and reduced job satisfaction [18,19]. Nonetheless, encouragingly, recent research shows that doctors trained in SDM as part of a multilevel implementation program exhibit greater appreciation for SDM and self-report improved execution [20], which, in line with SDT theory, might be caused by increased feelings of competence. These findings suggest that aligning SDM with individual competence levels and needs, following SDT principles, may lead to higher intrinsic motivation and better execution of SDM.

Considering SDM implementation's intricacies, this study aims to investigate healthcare providers' perceived characteristics from SDM, explicitly focusing on job satisfaction and perceived complexity. By delving into these two attributes and their interplay, we seek to gain a deeper understanding of how physicians rank SDM tasks compared to other tasks in the clinical setting. To the best of our knowledge, no existing studies have specifically examined these attributes in relation to SDM within the current literature. To address this knowledge gap, we have designed an explorative study with a novel experimental approach where doctors will rate patient-related medical scenarios. Through this investigation, we aspire to illuminate the crucial determinants for successful SDM implementation and integration into routine clinical practice.

2. Methods

We conducted a ranking study using self-developed case scenarios, called case vignettes. We asked participants to rank ten case vignettes by placing each case in one of three ordinal preference categories. Ranking occurred in two groups. Depending on group allocation, the preference categories per participant related to either job satisfaction or the complexity of the case scenarios. We used an ordinal preference elicitation method, which has some advantages over other quantitative rating methods. Generally, ranking methods are easier to understand, produce more consistent responses, and therefore are expected to reflect participants' preferences better than quantitative ratings [21]. The ranking methodology will be discussed in more detail below.

2.1. Case vignettes

We described ten case vignettes in a general hospital's gynaecology and obstetrics department. We designed the vignettes purposefully as four 'technical' cases (TECH), which means cases dealing with medical-technical skills or knowledge, three 'SDM' cases (SDM), and three 'managing emotion' cases (EMO). All were purposefully designed to incorporate more or less difficult or potentially uncomfortable situations for doctors as we intentionally sought to address the perceived complexity of these more complex cases. Although informed about the study's aims, participants did not know beforehand about this division in task types. To reduce framing bias, we added the EMO cases that focused on emotional outbreaks amongst colleagues or patients. These cases depicted instances of intense emotional displays, like anger, frustration or sadness. This approach was taken, to prevent participants from recognising our distinct subjects of interest: doctors' ratings on technical and SDM-related scenarios [22]. Furthermore, cases in which doctors deal with psychosocial or emotional problems are rated as more difficult and lead to less job satisfaction than other practice scenarios, making them a useful comparison to the two case vignette types [19]. The TECH cases covered bleeding due to uterine wall tear during caesarean section, ultrasound cornual pregnancy diagnosis, haemostasis difficulties during loop excision of the cervical transformation zone, and vacuum extraction during vaginal delivery. The SDM cases concerned preferences surrounding delivery (e.g., no oxytocin and no scalp electrode) that are not in line with clinical practice guidelines, wish for a 'Woodruff' operation that is not the first preferred medical and doctor's option, and a patient that has a solid wish to continue oestrogen suppletion, while it is medically safer to discontinue after ten years of usage. The EMO cases involved a delirious patient asking for euthanasia, an emotional trainee, and an angry spouse after a complication during a caesarean section. One gynaecologist of the author team (FS) with experience in the design of vignettes prepared the vignettes. The other gynaecologists from the author team reviewed and improved the vignettes (LS, JA, DB). All judged the tasks related to the cases described as feasible for all participants without evident differences in difficulty. We asked the study participants whether they approved of the constructed vignettes to represent their work and whether these vignettes were suitable for ranking. Almost all responded positively, confirming representativeness and suitability for ranking.

2.2. Participants and setting

For this study, we used a non-probability sample. Our study population consisted of gynaecologists and gynaecology trainees working within five non-academic training hospitals and one academic hospital in the Netherlands. We employed a convenience sample approach by selecting hospitals from two gynaecology training regions. This decision was influenced by the fact that the majority of researchers involved in the study were affiliated with one of the two regions. This affiliation allowed us to have easier access to participants and anticipate higher participation rates from these areas. Gynaecology training in the Netherlands takes six years. Trainees were included from year three of their training onwards and worked either in a non-academic or an academic hospital. We excluded trainees from junior years, as senior trainees were expected to be experienced beyond the novice level. Due to their rotation scheme, all trainees worked or had been working in a non-academic hospital and therefore had experience in the full breadth of a gynaecologist's work. Gynaecologists working in academic hospitals were not invited to participate in the study. This decision was based on the fact that in the Netherlands, these gynaecologists are often highly subspecialised, and as a result, they only encounter some, but not all, of the predesigned case scenarios in their clinical practice.

2.3. Data collection and ranking procedure

Contact data for gynaecologists invited to participate were extracted from available lists from the national society in Obstetrics and Gynaecology (NVOG). They were all contacted through e-mail or in person. Two researchers approached all eligible trainees, as both were involved in gynaecology training (LS, FS). The case vignette ranking exercise was

sent out by e-mail or printed on paper.

Participants were assigned to either the 'complexity' ranking group or the 'job satisfaction' ranking group before they were invited to participate. Group assignment was not blinded.

Participants had to rank cases into three preference categories. The 'job satisfaction' group ranked the three most rewarding (H), four neutral (N) and the three least rewarding (L) cases. The 'complexity' group ranked the three most complex (H), four neutral (N) and three least complex (L) case vignettes. We only asked for a partial ranking of all ten vignettes in three categories. Full ordinal ranking becomes more complicated when ranking alternatives increase, and results may get biased [21]. Decreasing ordering complexity from ten options to categorising in three ordinal classes reduces this risk of ranking bias.

2.4. Research questions

The concrete research questions for our statistical analyses were: Do the frequencies and distributions of SDM, EMO and TECH vignette counts differ within the lower (L), and higher (H) ranked categories of 'job satisfaction' and 'complexity', respectively?

2.5. Statistical analyses

Due to the nature of our data, it was impossible to use standard statistical tests, like the Wilcoxon-signed rank test, as our data did not meet the assumptions for these tests (e.g., measures are not continuous). Therefore, an expert team of statisticians (JE, KS) developed an algorithm for analyses based on randomisation tests [23]. The algorithm made it possible to distinguish whether ranking results were based on chance.

We tested the null hypothesis that cases were randomly assigned to Category L, each with equal probability within each participant and with a multivariate hypergeometric distribution of the variables within Category L. [24] First, we determined how often each case vignette type SDM, EMO or TECH was assigned to the lowest Category L per participant. Then, we compared case vignette observed sum scores (S) with chance level, reporting both sum scores and the statistic's expected value (E) in each test. Then, using a convolution algorithm, we computed the sum scores' probability distribution from the hypergeometric distributions [25]. According to the doubling formula, the two-sided p-values were calculated from S based on these probability distributions [26]. Next, we repeated the previous steps and compared all observed vignette counts pairwise.

Subsequently, all analyses mentioned above were repeated for the highest Category, H. Because we conducted six significance tests in each block, the Bonferroni correction was also applied, meaning all p-values were multiplied by 6 to correct for false-positive results. A p-value <0.05 was considered significant. Analyses were executed separately for the 'job satisfaction' and 'complexity' ranking groups. In reporting, we use subscript L and H letters to reference the scoring category; for example, SDM_L for SDM case vignette scores in the lowest Category, L. The supplemental file, the Appendix, provides a more comprehensive description of the statistical methods used.

3. Results

Overall, 114 potential participants were contacted from five different hospitals. Thirty-eight (33%) participated. Participants consisted of nine men and twenty-nine women. Fifteen were trainees. Twenty-three participants ranked vignettes for job satisfaction, and fifteen ranked for complexity.

Table 1 provides an overview of case vignette-type contributions per preference category. Results from comparisons and statistical analysis are represented in the following two paragraphs. Most observed results are compared to expected results based on our null hypothesis.

3.1. Job satisfaction

When comparing observed vignette counts with expected sums, as displayed in Table 1, it was clear that sum scores S deviated significantly from the chance level. SDM was assigned to Category L more often than expected and less often than expected to Category H. The same conclusion holds for EMO_L and EMO_H. The opposite conclusion holds for TECH, which was assigned significantly less often than expected to Category L and more often to Category H. After the Bonferroni correction was applied, all tests except for EMO_H were still significant.

Previous results were consistent with pairwise comparisons. Results from these comparisons are partially displayed in Table 2. For Category L, the differences between $TECH_L$ and SDM_L ($S = -32$, $E = 6.9$, $p < 0.001$) and between $TECH_L$ and EMO_L ($S = -28$, $E = 6.9$, $p < 0.001$) were significantly smaller than expected, with TECH vignettes being picked far less often in Category L. The difference between SDM_L and

Table 1
Case vignette type distributions per preference category.

	Case vignette type per category	Observed sum (S) (1)	Expected sum (E) (2)	Difference (S-E)	Cumulative probability (3)	Point probability (4)	Two-sided p-value (5)	Bonferroni-corrected p-value (6)
Job satisfaction (N = 23 participants)	SDM_L	35	20.7	14.3	1.000	<0.001	**<0.001**	**<0.001**
	EMO_L	31	20.7	10.3	0.999	0.001	**0.004**	**0.024**
	$TECH_L$	3	27.6	−24.6	<0.001	<0.001	**<0.001**	**<0.001**
	SDM_H	4	20.7	−16.7	<0.001	<0.001	**<0.001**	**<0.001**
	EMO_H	13	20.7	−7.7	0.014	0.008	**0.028**	0.169
	$TECH_H$	52	27.6	24.4	1.000	<0.001	**<0.001**	**<0.001**
Complexity (N = 15 participants)	SDM_L	15	13.5	1.5	0.772	0.124	0.704	1.000
	EMO_L	5	13.5	−8.5	0.001	0.001	**0.002**	**0.011**
	$TECH_L$	25	18	7	0.995	0.008	**0.025**	0.152
	SDM_H	17	13.5	3.5	0.929	0.063	0.268	1.000
	EMO_H	16	13.5	2.5	0.866	0.094	0.456	1.000
	$TECH_H$	12	18	−6	0.027	0.016	0.054	0.326

SDM is shared decision making, EMO is emotional, and TECH is technical. L or H in the subscript refers to the low or high ranking category, respectively.
All significance tests are performed according to the algorithm based on randomisation tests, as described under 2.5 Statistical analyses and in the Appendix.
(1) number of times that the task (SDM, EMO, or TECH) is assigned to the position (L or H, as indicated in subscript), added across all persons in the samples.
(2) the expected value of 'sum' if all persons assign the tasks randomly to positions.
(3) probability that the sum is less than or equal to the observed value if all persons assign the tasks randomly to positions.
(4) probability that the sum is equal to the observed value if all persons assign the tasks randomly to positions.
(5) computed from the previous two probabilities.
(6) 6 times the original two-sided probability, as six comparisons are made; $p < 0.05$ is here considered statistically significant.
All significance tests are performed according to the algorithm based on randomisation tests, as described under 2.5 Statistical analyses and in the Appendix.

Table 2
Case vignette type pairwise comparisons per preference category for 'job satisfaction'.

Case vignettes per type and preference category for 'job satisfaction' ranking		SDM_L	EMO_L	$TECH_L$	SDM_H	EMO_H	$TECH_H$
	Observed sum of case vignettes per category	35	31	3	4	13	52
SDM_L	35	0	4 (*NS*)	32 ($p < 0.001$)			
EMO_L	31	−4 (*NS*)	0	28 ($p < 0.001$)			
$TECH_L$	3	−32 ($p < 0.001$)	−28 ($p < 0.001$)	0			
SDM_H	4				0	−9 (*NS*)	−48 ($p < 0.001$)
EMO_H	13				9 (*NS*)	0	−39 ($p < 0.001$)
$TECH_H$	52				48 ($p < 0.001$)	39 ($p < 0.001$)	0

SDM is shared decision making, EMO is emotional, and TECH is technical. L or H in the subscript refers to the low or high ranking category, respectively.
All significance tests are performed according to the algorithm based on randomisation tests, as described under 2.5 Statistical analyses and in the Appendix.
For each comparison in the cross-tabulation, the value is calculated by subtracting the vignette count in the horizontal row from the count in the vertical row; between brackets, the *p*-value is provided from the 2-sided *p*-test; only significant *p*-values ($p < 0.05$) are displayed; *NS* means non-significant.
Irrelevant cells are left empty.

EMO_L was not significant ($S = 4$, $E = 0$, $p = 0.428$). For Category H, the difference between $TECH_H$ and SDM_H was significantly larger than expected ($S = 48$, $E = 6.9$, $p < 0.001$), with TECH being picked far more often. The difference between $TECH_H$ and EMO_H was more extensive than expected as well ($S = -39$, $E = 6.9$, $p < 0.001$), but the difference between SDM_H and EMO_H was not significant ($S = -9$, $E = 0$, $p = 0.094$). These patterns remained intact after the Bonferroni correction.

These results are reasons to reject the null hypothesis of indifference in the low (L) and high (H) rating categories of 'job satisfaction', as handling TECH cases seems to lead to higher job satisfaction and dealing with SDM and EMO cases leads to lower job satisfaction.

3.2. Complexity

The significance tests showed that EMO was assigned to L less frequently than expected from the chance level, and TECH was assigned to L more frequently than expected. However, only the first deviation was significant after the Bonferroni correction. The other frequencies, displayed in Table 1, did not deviate significantly from the chance level.

The significance tests of the differences showed that for Category L, the difference between $TECH_L$ and EMO_L was significantly larger than expected ($S = 20$, $E = 6.9$, $p = 0.001$), rating fewer TECH cases as complex. There was a considerable difference in case count between $TECH_L$ and SDM_L; however, this difference did not reach significance ($S = 10$, $E = 4{,}5$, $p = 0.157$). Bonferroni correction yielded the same results. The other two differences in Category L and all three differences in

Table 3
Case vignette type pairwise comparisons per preference category for 'complexity' ranking.

Case vignettes per type and preference category for 'complexity' ranking		SDM_L	EMO_L	$TECH_L$	SDM_H	EMO_H	$TECH_H$
	Observed sum of case vignettes per category	15	5	25	17	16	12
SDM_L	15	0	10 ($p = 0.037$)*	−10 (*NS*)			
EMO_L	5	−10 ($p = 0.037$)*	0	−20 ($p = 0.001$)			
$TECH_L$	25	10 (*NS*)	20 ($p = 0.001$)	0			
SDM_H	17				0	1 (*NS*)	5 (*NS*)
EMO_H	16				−1 (*NS*)	0	4 (*NS*)
$TECH_H$	12				−5 (*NS*)	−4 (*NS*)	0

*) non-significant after Bonferroni correction.
SDM is shared decision making, EMO is emotional, TECH is technical. L or H in the subscript refers to the low or high ranking category, respectively.
All significance tests are performed according to the algorithm based on randomisation tests, as described under 2.5 Statistical analyses and in the Appendix.
For each comparison in the cross-tabulation, the value is calculated by subtracting the vignette count in the horizontal row from the count in the vertical row; between brackets, the *p*-value is provided from the 2-sided *p*-test; only significant *p*-values ($p < 0.05$) are displayed; *NS* means non-significant
Irrelevant cells are left empty

Category H were insignificant. More results are shown in Table 3.

We determined that the three task types (SDM, EMO, TECH) have different L ratings because the TECH and EMO difference is significant within Category L. TECH cases were perceived as less complex than EMO cases. However, we could not establish that this is also true within Category H. SDM did not differ significantly from EMO or TECH cases in all comparisons.

4. Discussion and conclusion

4.1. Discussion

Curious about the reasons behind the internationally reported delay in implementation and inspired by SDT and DOI theory, we set out to study two aspects of SDM: the job satisfaction it would offer physicians and the perceived complexity of the concept. Results are inconclusive about whether doctors find SDM complex, yet they provide some clues for explaining the lack of SDM implementation. Compared with technical cases, our study suggests that participants get less job satisfaction from SDM. When comparing technical cases to all others, we observed a trend in our data that gynaecologists rate these technical cases as less complex yet provide higher satisfaction levels. These results hint that gynaecologists might look more for predictable outcomes from technical activities within their comfort zone. Subsequently, SDM cases might not be within the gynaecologists' comfort zone.

These propositions raise several questions. The first question is what causes SDM not to be within a physician's comfort zone. One explanation might be found in the psychology of human decision-making itself. Research within this domain shows that people prefer familiar alternatives over unknown alternatives [27]. In general, selecting a familiar option is preferred [27,28]. As SDM is associated with uncertainty, with less control over physician-patient interaction and potentially different results than medical guidelines advise, it is understandable why it is not within the physician's comfort zone [29]. These findings align with SDT, which underscores that increased autonomy and competence are more comfortable within the zone of familiarity [17].

The second question is whether bringing SDM within the physicians' comfort zone would still be possible. Providing doctors with a sense of control by familiarising them with SDM concepts might help. For example, by teaching doctors how to master SDM skills better and educating them in dealing with uncertainties in doctor-patient communication and subsequent out-of-protocol healthcare choices [29,30]. Or even more practical, by providing them with more decision aids to help them structure the consultation according to SDM principles [31,32]. Training initiatives and decision aids have increasingly been developed over the past few decades [32,33], so this seems to be an easy fix to the previously presented problem. Nevertheless, extensive training is still minimally implemented, and decision aids are only relatively little available in the medical workspace. Furthermore, results from training interventions described in research do not yield encouraging results, as many studies do not show positive effects like increased or better use of SDM by doctors [32,33].

Normalising SDM use by integrating it into physicians' daily working routines might better reduce physicians' levels of uncertainty or unfamiliarity. For example, by standard addressing SDM in any (multidisciplinary) team meeting and at all patient case discussions and incorporating SDM principles in every clinical guideline [29]. Unfortunately, SDM is still far too often considered ‘different’. Even in our current study, we started from the position that we need to think about change: changing the culture, engaging and enabling the organisation, and implementing and sustaining the change [34]. It is time for us to leave the we-still-need-to-make-a-change standpoint and begin normalising SDM as an element of ‘the air we breathe’ as healthcare professionals.

However, a further question is whether attempts to put SDM within the physician's comfort zone would mitigate its backlog on technical tasks. Previous research shows that the most often mentioned reasons to pursue a medical career are helping people and their interest in medicine [35]. Against this background, would physicians learn to appreciate being confronted with potentially more demanding patients or not? A culture change will be needed for clinicians to appreciate these non-technical aspects of their profession [29]. Previous study results provide hope for the future of this culture change. For example, an extensive international survey amongst general practitioners confirmed that performing technical tasks is related to higher job satisfaction; however, patient satisfaction is even more connected to physicians' job satisfaction [36]. If SDM delivers its promise of more satisfied patients, it could compete with technical tasks by yielding equal or even better job satisfaction scores.

Transitioning to the practice implications of our research findings, a notable observation arises: despite the conviction that SDM is important and the morally just thing to do [37], many physicians display signs of the so-called Dunning-Kruger effect, a phenomenon in which people with little understanding of and competence in a particular domain seem to overestimate their performance levels [38]. It is essential to understand more about these phenomena and the psychology of health professionals in general, as studies tapping into these subjects are primarily neglected in medical research. Furthermore, the previously mentioned culture change is required amongst health professionals, which cannot be promoted by simply addressing barriers to SDM implementation [29]. Therefore, for future research, we advocate for studies of SDM that make the psychological and cultural issues of healthcare professionals and patients transparent. In addition, we must know the ‘dark side’ of the concept to understand the implementation delay and prevent naive approaches. Additionally, another future research question might be which kind of patient care SDM can best show quick wins, as visible quick wins can be significant determinants for normalising SDM in daily healthcare [15].

Our study has several limitations. It quantifies sentiments concerning SDM in practising physicians as essential stakeholders in the implementation process yet includes only a small set of physicians from within one specialty. Ideally, a bigger sample size with a better spread over several types of hospitals and medical specialties would have been used, leading to improved validity and generalizability of study results. In addition, our vignette study lacks an in-depth consideration of patient variations, including differences in skills, competencies, or opportunities to engage in SDM with their healthcare providers effectively. These variations significantly contribute to the complexity of SDM [39]. Furthermore, we did not comment on the differences between TECH and SDM cases with the appreciation of EMO cases in our discussion section, as these comparisons were outside the scope of our study. Despite our efforts to prevent it, the vignette study could have been subject to framing bias, which occurs when people choose based on how the information is presented instead of the facts themselves [22]. Moreover, the design of the vignettes could have biased the ranking results as well, as we purposely involved more complicated cases, and cases are non-equivalent. Next, we did not register whether participants were trained or skilled in SDM or not; differences in SDM skills present could have influenced results. Sampling was not blinded, and we obtained a higher number of responses in the job satisfaction sample compared to the complexity sample. Finally, the finding that physicians perceive SDM as less rewarding and complex does not imply that they will refuse to do it. The question remains how to seduce them to enjoy ‘the art of SDM’.

4.2. Innovation

An important strength of our study is the innovative design and the statistical methodology purposefully designed to analyse our study results. Moreover, this is the first time that the concept of SDM has been compared to the appraisal of technical tasks regarding their perceived complexity and job satisfaction ratings amongst physicians. These

comparisons, tapping into the psychology of health professionals, are still a largely neglected field in medical research. Nevertheless, those types of studies could provide valuable insights for implementation science.

4.3. Conclusion

SDM is compared to performing technical tasks perceived as complex and less of a source of job satisfaction. In solving the questions behind the delay in implementing SDM, psychological issues and culture deserve to be at the centre of attention. Only when we understand these human aspects might we invent better ways to implement SDM.

Submission declaration and verification

The authors declare that this work has not been previously published.

Funding source declaration

No funding source.

Informed consent

All participating gynaecologists and gynaecology trainees provided their consent for use of their ranking results. No personal patient details are described in the manuscript.

CRediT authorship contribution statement

Laura Spinnewijn: Conceptualization, Methodology, Formal analysis, Visualization, Investigation, Project administration, Writing – original draft. **Johanna Aarts:** Supervision, Writing – review & editing, Validation. **Didi Braat:** Supervision, Writing – review & editing, Validation. **Nikolaj Baranov:** Investigation, Project administration. **Klaas Sijtsma:** Methodology, Data curation, Writing – review & editing. **Jules Ellis:** Methodology, Data curation, Formal analysis, Writing – review & editing. **Fedde Scheele:** Conceptualization, Methodology, Investigation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pecinn.2023.100208.

References

[1] Charles C, Gafni A, Whelan T. Decision-making in the physician-patient encounter: revisiting the shared treatment decision-making model. Soc Sci Med 1999;49(5):651–61. https://doi.org/10.1016/s0277-9536(99)00145-8.
[2] Hoffmann TC, Montori VM, Del Mar C. The connection between evidence-based medicine and shared decision making. JAMA. 2014;312(13):1295–6. https://doi.org/10.1001/jama.2014.10186.
[3] Coulter A. Partnerships with patients: the pros and cons of shared clinical decision-making. J Health Serv Res Policy 1997;2(2):112–21. https://doi.org/10.1177/135581969700200209.
[4] Stacey D, Legare F, Lewis K, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev Apr 12 2017;4:CD001431. https://doi.org/10.1002/14651858.CD001431.pub5.
[5] Hughes TM, Merath K, Chen Q, et al. Association of shared decision-making on patient-reported health outcomes and healthcare utilization. Am J Surg 2018;216(1):7–12. https://doi.org/10.1016/j.amjsurg.2018.01.011.
[6] Couet N, Desroches S, Robitaille H, et al. Assessments of the extent to which health-care providers involve patients in decision making: a systematic review of studies using the OPTION instrument. Health Expect 2015;18(4):542–61. https://doi.org/10.1111/hex.12054.
[7] Stiggelbout AM, Van der Weijden T, De Wit MP, et al. Shared decision making: really putting patients at the Centre of healthcare. BMJ. 2012;344:e256. https://doi.org/10.1136/bmj.e256.
[8] Guirguis M, Thompson E, Miller J, Sommer R, Curran-Cook D, Kaba A. Qualitative examination of shared decision-making in Canada's largest health system: more work to be done : shared decision-making-more work to be done. J Patient Exp 2021;8. https://doi.org/10.1177/23743735211064141. 23743735211064141.
[9] Scholl I, LaRussa A, Hahlweg P, Kobrin S, Elwyn G. Organizational- and system-level characteristics that influence implementation of shared decision-making and strategies to address them - a scoping review. Implement Sci 2018;13(1):40. https://doi.org/10.1186/s13012-018-0731-z.
[10] van Veenendaal H, van der Weijden T, Ubbink DT, Stiggelbout AM, van Mierlo LA, Hilders C. Accelerating implementation of shared decision-making in the Netherlands: An exploratory investigation. Patient Educ Couns 2018;101(12):2097–104. https://doi.org/10.1016/j.pec.2018.06.021.
[11] Joseph-Williams N, Elwyn G, Edwards A. Knowledge is not power for patients: a systematic review and thematic synthesis of patient-reported barriers and facilitators to shared decision making. Patient Educ Couns 2014;94(3):291–309. https://doi.org/10.1016/j.pec.2013.10.031.
[12] Elwyn G, Laitner S, Coulter A, Walker E, Watson P, Thomson R. Implementing shared decision making in the NHS. BMJ. Oct 14 2010;341:c5146. https://doi.org/10.1136/bmj.c5146.
[13] Moore L, Britten N, Lydahl D, Naldemirci O, Elam M, Wolf A. Barriers and facilitators to the implementation of person-centred care in different healthcare contexts. Scand J Caring Sci 2017;31(4):662–73. https://doi.org/10.1111/scs.12376.
[14] Grover S, Fitzpatrick A, Azim FT, et al. Defining and implementing patient-centered care: An umbrella review. Patient Educ Couns 2022;105(7):1679–88. https://doi.org/10.1016/j.pec.2021.11.004.
[15] Rogers EM. Diffusion of innovations. 5th ed. Free Press; 2003.
[16] Lorenzetti CR, Jacques CH, Donovan C, Cottrell S, Buck J. Managing difficult encounters: understanding physician, patient, and situational factors. Am Fam Physician 2013;87(6):419–25.
[17] Deci EL, Ryan RM. Self-determination theory. In: Lange PAMV, Kruglanski AW, Higgins ET, editors. Handbook of theories of social psychology. SAGE; 2012.
[18] An PG, Rabatin JS, Manwell LB, et al. Burden of difficult encounters in primary care: data from the minimizing error, maximizing outcomes study. Arch Intern Med 2009;169(4):410–4. https://doi.org/10.1001/archinternmed.2008.549.
[19] Krebs EE, Garrett JM, Konrad TR. The difficult doctor? Characteristics of physicians who report frustration with patients: an analysis of survey data. BMC Health Serv Res 2006;6:128. https://doi.org/10.1186/1472-6963-6-128.
[20] van Veenendaal H, Voogdt-Pruis HR, Ubbink DT, et al. Evaluation of a multilevel implementation program for timeout and shared decision making in breast cancer care: a mixed methods study among 11 hospital teams. Patient Educ Couns 2022;105(1):114–27. https://doi.org/10.1016/j.pec.2021.05.005.
[21] Ali S, Ronaldson S. Ordinal preference elicitation methods in health economics and health services research: using discrete choice experiments and ranking methods. Br Med Bull 2012;103(1):21–44. https://doi.org/10.1093/bmb/lds020.
[22] Popovich I, Szecket N, Nahill A. Framing of clinical information affects physicians' diagnostic accuracy. Emerg Med J 2019;36(10):589–94. https://doi.org/10.1136/emermed-2019-208409.
[23] Edgington ES. Randomisation tests. In: Lovric M, editor. International encyclopedia of statistical science. Springer; 2011.
[24] Mehta CR, Patel NR. A network algorithm for performing fisher exact test in R X C contingency-tables. J Am Stat Assoc 1983;78(382):427–34. https://doi.org/10.2307/2288652.
[25] Johannssen A, Chukhrova N, Castagliola P. Efficient algorithms for calculating the probability distribution of the sum of hypergeometric-distributed random variables. Methodsx. 2021;8:101507. https://doi.org/10.1016/j.mex.2021.101507.
[26] Rivals I, Personnaz L, Taing L, Potier MC. Enrichment or depletion of a GO category within a class of genes: which test? Bioinformatics. 2007;23(4):401–7. https://doi.org/10.1093/bioinformatics/btl633.
[27] Samuelson W, Zeckhauser R. Status quo bias in decision making. J Risk Uncertain 1988;1:7–59.
[28] Tversky A, Kahneman D. Loss aversion in riskless choice - a reference-dependent model. Q J Econ 1991;106(4):1039–61. https://doi.org/10.2307/2937956.
[29] Spinnewijn L, Aarts J, Verschuur S, Braat D, Gerrits T, Scheele F. Knowing what the patient wants: a hospital ethnography studying physician culture in shared decision making in the Netherlands. BMJ Open 2020;10(3):e032921. https://doi.org/10.1136/bmjopen-2019-032921.
[30] Simpkin AL, Schwartzstein RM. Tolerating uncertainty - the next medical revolution? N Engl J Med 2016;375(18):1713–5. https://doi.org/10.1056/NEJMp1606402.
[31] Elwyn G, Scholl I, Tietbohl C, et al. "Many miles to go ...": a systematic review of the implementation of patient decision support interventions into routine clinical practice. BMC Med Inform Decis Mak 2013;13(Suppl. 2):S14. https://doi.org/10.1186/1472-6947-13-S2-S14.
[32] Legare F, Adekpedjou R, Stacey D, et al. Interventions for increasing the use of shared decision making by healthcare professionals. Cochrane Database Syst Rev 2018;7:CD006732. https://doi.org/10.1002/14651858.CD006732.pub4.
[33] Siyam T, Shahid A, Perram M, et al. A scoping review of interventions to promote the adoption of shared decision-making (SDM) among health care professionals in

clinical practice. Patient Educ Couns 2019;102(6):1057–66. https://doi.org/10.1016/j.pec.2019.01.001.
[34] Kotter JP. Leading change. Harvard Business Review Press; 2012.
[35] Goel S, Angeli F, Dhirar N, Singla N, Ruwaard D. What motivates medical students to select medical studies: a systematic literature review. BMC Med Educ 2018;18(1):16. https://doi.org/10.1186/s12909-018-1123-4.
[36] Stobbe EJ, Groenewegen PP, Schafer W. Job satisfaction of general practitioners: a cross-sectional survey in 34 countries. Hum Resour Health 2021;19(1):57. https://doi.org/10.1186/s12960-021-00604-0.
[37] Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. J Pers Soc Psychol 1999;77(6):1121–34. https://doi.org/10.1037//0022-3514.77.6.1121.
[38] Tan NQP, Volk RJ. Addressing disparities in patients' opportunities for and competencies in shared decision making. BMJ Qual Saf Feb 2022;31(2):75–8. https://doi.org/10.1136/bmjqs-2021-013533.
[39] Duggan PS, Geller G, Cooper LA, Beach MC. The moral nature of patient-centeredness: is it "just the right thing to do"? Patient Educ Couns 2006;62(2):271–6. https://doi.org/10.1016/j.pec.2005.08.001.