

University of Groningen

Re-assessing the validity of the Moral Sensitivity Questionnaire (MSQ)

Kuilman, Luppo; Jansen, Gerard J.; Mulder, Laetitia B.; Middel, Berrie; Roodbol, Petrie F.

Published in:
Journal of Evaluation in Clinical Practice

DOI:
[10.1111/jep.13353](https://doi.org/10.1111/jep.13353)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Kuilman, L., Jansen, G. J., Mulder, L. B., Middel, B., & Roodbol, P. F. (2020). Re-assessing the validity of the Moral Sensitivity Questionnaire (MSQ): Two new scales for moral deliberation and paternalism. *Journal of Evaluation in Clinical Practice*. <https://doi.org/10.1111/jep.13353>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



ORIGINAL ARTICLE

Re-assessing the validity of the Moral Sensitivity Questionnaire (MSQ): Two new scales for moral deliberation and paternalism

Luppo Kuilman BSN, MPA^{1,2} | Gerard J. Jansen PhD¹ |
Laetitia M. Mulder PhD³ | Berrie Middel PhD⁴ | Petrie F. Roodbol PhD¹

¹Nursing Research Section, Department of Health Sciences, University Medical Center Groningen, University of Groningen, Groningen, Netherlands

²Department of Physician Assistant Studies, College of Health and Human Service, Northern Arizona University, Arizona, USA

³Department of Human Resource Management & Organizational Behaviour, Faculty of Economics and Business, University of Groningen, Groningen, Netherlands

⁴Department of Health Sciences, Community & Occupational Medicine Division, University Medical Center Groningen, University of Groningen, Groningen, Netherlands

Correspondence

Luppo Kuilman, Nursing Research Section,
Department of Health Sciences, University
Medical Center Groningen, University of
Groningen, PO Box 30.001, 9700 RB
Groningen, Netherlands.
Email: l.kuilman@pl.hanze.nl

Abstract

Rationale, aims, and objectives: The current study and previous research have called the six-component model of Lützen's 30-item Moral Sensitivity Questionnaire (MSQ) into question. For this reason, we re-examined the construct validity of this instrument.

Methods: In this cross-sectional study, which was based on a convenience sample of Dutch nurse practitioners (NPs) and physician assistants (PAs), we tested the validity of MSQ items using exploratory and confirmatory factor analyses (EFA and CFA, respectively).

Results: The EFA revealed a two-component model, which was then tested as a target model with CFA and was found to have good model fit. Some items were correlated with two uncorrelated latent constructs, which we labelled as "paternalistic" and "deliberate" attitudes towards patients.

Conclusions: As in previous studies, the analyses in the current study, which was conducted among PAs and NPs, did not reveal six dimensions for the 30 items. Two new latent dimensions of moral sensitivity were psychometrically tested and confirmed. These two components relate to studies investigating ethical behaviour, and they can be used to describe the moral climate in healthcare organizations. The scales are indicators of the extent to which health professionals behave in a deliberate

List of abbreviations: AIC, Akaike information criterion; BCPH, Behavioral Control targeted at Preventing Harm; CFA, Confirmatory factor analysis; CFI, Comparative fit index; CI, Confidence interval; DIT-N2, Defining Issues Test-N2; EAS, Ethics Advocacy Scale; EFA, Exploratory factor analysis; FCM, Four-component model of moral behavior; GFI, Goodness-of-fit index; IRB, Institutional Review Board; MD, Medical doctor; MDS, Moral Disengagement Scale; MS, Moral sensitivity; MSQ, Moral Sensitivity Questionnaire; MSQ-DELIB, Deliberate Attitude Scale; MSQ-PATER, Paternalistic Attitude Scale; NP, Nurse practitioner; PA, Physician assistant; RMSEA, Root mean square error of approximation; SDM, Shared decision-making; SRMR, Standardized root mean square residual; STROBE, Strengthening the Reporting of Observational Studies in Epidemiology; Velicer's MAP, Velicer's Minimum Average Partial.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Journal of Evaluation in Clinical Practice* published by John Wiley & Sons Ltd

(sensitive) or paternalistic (insensitive) manner towards the opinions of patients within the context of medical decision-making.

KEYWORDS

deliberation, moral sensitivity, nurse practitioner, paternalism, physician assistant

1 | INTRODUCTION

In Western health systems, two interesting shifts with regard to professional and patient responsibility have taken place over the last few decades. First, the professional responsibility of making medico-ethical decisions that exclusively belonged to the realm of medical doctors (MDs) has been extended to other health professions, including nurse practitioners (NPs) and physician assistants (PAs).^{34,36} Second, in the past, MDs guided their patients through the medical treatment process according to a strong paternalistic attitude. In current practice, the perspective has shifted towards emphasizing the central role of patients in healthcare.⁵⁰ Within the models of shared decision-making (SDM) that are now prevalent, assigning a central role to the patient is regarded as an ethical imperative. Such models of SDM are consistent with the four principles of ethics in care: respecting autonomy, propagating beneficence, avoiding harm, and achieving justice.⁷ Medical decisions established through SDM have been shown to be associated with improved medication compliance, health-related quality of life, an increase in patients' perceived control over their choices with regard to treatment options, and a decrease in healthcare utilization.¹⁶ In the past, clinicians were accustomed to employing protocols and guidelines that were accepted as the gold standard for treatment. In contrast, computer-literate and empowered patients are adding a new dimension to the treatment relationship, thus potentially increasing the risk of tension and conflict.²⁵

In light of such changes in the treatment relationship, tension is likely to arise between what a clinician regards as the best treatment option (or even what rules and regulations dictate that they propose) and the treatment that is perceived as the best in the eyes of the patient. Such tension could create a moral dilemma, which could be described as a situation in which for example there are conflicting opinions (between health professional and patient) regarding what is the best treatment option.¹⁵

Health professionals may employ one of essentially two decision-making strategies or coping mechanisms to reduce dilemma-related stress: (i) a predominantly patient-centred, deliberate attitude focused on patient autonomy^{42,47} or (ii) a more dominant, clinical view, known as the "paternalistic approach."^{40,50} Health professionals adopting a paternalistic attitude are less likely to engage in dialogue regarding treatment options or the health beliefs of patients. They are more likely to decide what is best for the patient based on their own self-presumed professional knowledge and evidence-based practice. Health professionals who have adopted a deliberate attitude that takes the opinions and wishes of patients into account must reflect on their decisions in the light of the patient's views.¹

2 | BACKGROUND

Regardless of whether health professionals cope with moral dilemmas through either a deliberate or paternalistic attitude, moral dilemmas arising within interactions must necessarily be resolved through an ethical decision-making process. For example, James Rest captures this ethical decision-making process in the "four-component model of moral behavior" (FCM). The FCM states that moral decision-making is influenced by moral sensitivity, moral reasoning, moral motivation, and moral character. In this model, Rest conceptualizes moral sensitivity as the first and essential precursor in ethical decision-making, defining it as "a combination of one's recognition of moral issues, and how one reacts and processes these issues from an affective perspective within a social context."⁴⁵

Lütznén and colleagues³² defined the concept of moral sensitivity (MS) in theoretical terms as "a personal attribute involving the ability to recognize a moral conflict, a contextual and intuitive understanding of a person's vulnerable situation and insight into the ethical consequences of decisions made on behalf of another person." They operationalized this concept of moral sensitivity using the Moral Sensitivity Questionnaire (MSQ) in study populations consisting of psychiatrists³⁰ and psychiatric nurses.^{29,31} Based on their results, they reported six dimensions (ie, latent variables). With reference to exploratory analysis, Lütznén and colleagues label these dimensions as follows: 1) interpersonal orientation, 2) structuring moral meaning, 3) expressing benevolence, 4) modifying autonomy, 5) experiencing moral conflict, and 6) having confidence in medical knowledge.²⁹ In a methodological and statistical appraisal of the results as published, however, a weak structure emerges as a result of three observations. First, factor loadings (correlations between items and the underlying construct) were too low, as items should be sufficiently correlated (factor loading $\geq .40$) with the target dimension in the data. Second, some correlations were biased such that the target construct could not be interpreted, as the full matrix of factor loadings was not presented. And third, several items were correlated with more than one construct of moral sensitivity, thereby violating the necessary condition that each item should exclusively tap an aspect of only one underlying construct or dimension. As a consequence of these problems, the indices of reliability or internal consistency (Cronbach's alpha) for these six scales ranged from 0.36 to 0.61, thus indicating poor intercorrelations between the items.

Other MSQ studies conducted in many different countries^{10,14,23,54} have also evaluated the content and psychometric quality of a 30-item MSQ. These combinations of items proposed in these studies deviate from the latent constructs proposed by Lütznén.



As was the case with the instrument-testing performed by Lützn and colleagues, the aforementioned studies consisted exclusively of exploratory factor analysis (EFA). This method is not the most suitable for arriving at conclusive results about the factor structure of a scale, given that EFA based solely on the Kaiser criterion could potentially generate an excessively inclusive result.¹⁸

Thus, as the mixed results of the above studies suggest, there are still some unclaritys about what the MSQ measures and how it should be used. On top of this, the 30-item MSQ has so far only been validated among psychiatrists and among nurses.^{29,30,33} However, it will be particularly interesting and relevant to develop and validate the MSQ among PAs and NPs. This is because these healthcare professionals have a special role that distinguishes them from nurses and doctors. As their responsibility lies in between that of MDs and nurses, their role is largely characterized by having medical-decisional responsibilities. In this role, both deliberate and paternalistic attitudes may take a prominent place in their professional identities. Indeed, in a feasibility study that we first performed, we found some indication that the MSQ administered among this specific group particularly distinguishes paternalistic and deliberative attitudes. We conducted this feasibility study among Master's-level PA students (N = 32). By employing the method developed by Ruscio and Roche,⁴⁸ we found a two-factor structure. One of these factors apparently reflects a paternalistic attitude in decision-making, while the other reflects a deliberate attitude.

2.1 | Study objective

To test the reproducibility of this two-component factor structure, we performed another study based on a larger sample consisting of the particular group of Dutch PAs and NPs. The objectives of this study are as follows: a) evaluate how many factors of the MSQ should be retained for further factor analysis, and b) apply both exploratory and confirmatory factor analysis (CFA) to test the dimensionality, scalability, and construct validity of the items remaining from the MSQ.

2.2 | Research Questions

To meet the study objectives, the following research questions will be addressed:

Which items of the MSQ are unequivocally correlated with latent constructs when using the criteria of EFA after having determined how many factors should be retained, and how strong is the model fit, based on CFA?

Are the items of these components scalable? And do these scales have sufficient internal consistency (reliability)?

Do the retained scales confirm discriminant or convergent validity as hypothesized when correlated with the following scales, which are known to tap moral aspects of the decision-making process: (i) the Behavioral Control targeted at Preventing Harm (BCPH) scale,

(ii) the Ethics Advocacy Scale (EAS), (iii) the Moral Disengagement Scale (MDS), and (iv) the Defining Issues Test (DIT-N2)?

3 | METHODS

3.1 | Study design, participants, and data collection

In this cross-sectional study, five PA and one NP degree programs were selected as sources for approaching alumni. Based on the Dutch Personal Data Protection Act, the researchers were not granted permission to use the databases of the programs in order to retrieve alumni email addresses. The information letter concerning the current study was therefore sent to 470 NP alumni and to 426 PA alumni by the programs' administrators. By activating a hyperlink to a private web-based system included in this letter, individual alumni were free to reveal their contact details to the researchers. When respondents returned permission to use their email addresses, this was regarded as informed consent. Upon receipt of their permission, these alumni were sent the access key to the web-based set of questionnaires.

In all, 294 subjects were willing to participate: 176 PAs and 118 NPs, meaning a response rate of 52.7% (ie, 155/294). Upon closure of the online survey (between January and March 2015), 155 respondents had completed all of the questionnaires. In all, 139 alumni, who initially consented to participate, eventually did not reply to the survey. Therefore, no information about this group was available that could be used to test for selection bias. All questions in the Qualtrics online survey environment were forced choice, so there were no missing data. The "Strengthening the Reporting of Observational Studies in Epidemiology" (STROBE) checklist was employed (see Supporting Information S1).

3.2 | Ethical approval and consent to participate

According to the statement by the Central Committee on Research Involving Human Subjects (www.ccmo.nl), no Institutional Review Board (IRB) approval was warranted for this type of survey study among volunteer professionals. This study was performed in accordance with the tenets of the Declaration of Helsinki.²¹ Only the first author (LK) had access to the online survey data.

3.3 | Academic integrity statement

The dataset in the current study was the same as the one in Kuilman et al (2019).²⁶ However, from that pool different variables were used, focusing on different research questions. Only the MDS and the indicator for moral reasoning (DIT-N2) were used in both studies, albeit with different hypotheses and functionality (independent vs dependent variable).

3.4 | Statistical analyses

All data were analysed using SPSS Statistics for Windows, Version 25.0. CFA was performed using SPSS AMOS, Version 23.0.

3.5 | Bivariate analysis

For categorical data, we used the chi-square test (Fisher's exact tests for 2×2 contingency tables) and the difference-between-proportions test.³⁸ For continuous variables, we used the Student's *t*-test for independent samples.

3.6 | Multivariate analysis

To assess the structural validity of the MSQ, we performed factor analyses and calculations of reliability estimates, as explained below.

3.7 | Model fit through Confirmatory Factor Analysis

We used the following goodness-of-fit indices to determine model fit using CFA: a) chi-square/degrees of freedom (χ^2/df), b) root mean square error of approximation (RMSEA), c) standardized root mean square residual (SRMR), d) comparative fit index (CFI), and e) a goodness-of-fit index (GFI). The χ^2/df with a ratio between 0 and 2 is indicative of a good fit.^{49,52} For the RMSEA, a cut-off value less than or close to 0.06 was assumed to be appropriate.²⁴ The lower limit of the confidence interval (CI) should be close to 0, and the upper limit should not exceed 0.08.³⁵ We also report the SRMR, as its standardized nature makes it easier to interpret. Values for the SRMR ranged from zero to 1.0, with good-fitting models having an acceptable threshold of less than 0.08.²⁴ For the CFI, values equal to or greater than 0.95 are deemed indicative of a good model fit.²⁴ For the GFI, cut-off values greater than or equal to 0.95 are recommended for relatively low factor loadings and sample sizes.³⁷ The Akaike information criterion (AIC) was used to compare different models. This criterion is a descriptive measurement, in which the preferred model is the one with the lowest value.³

3.8 | Internal consistency

Cronbach's alpha values were calculated to examine the reliability of all scales. In general, values equal to or greater than ≥ 0.70 are considered sufficient.⁹

3.9 | Convergent and divergent validity

Convergent validity refers to the extent to which a construct measures what it is purported to measure.^{41,51} It is assessed according to

data showing that different measurements of conceptually related dimensions of moral behaviour are conceptually associated in the hypothesized direction.

In this study, convergent validity was imputed according to statistically significant associations (linear associations between measurements of moral behaviour), while divergent validity was assumed when there was no correlation (ie, $P > 0.05$). The degree of overlap between constructs was estimated by calculating the nonparametric effect size of Rho (given the sample size). The statistically significant small effect size, with Rho in the range of ≥ 0.10 to < 0.30 and Rho ≥ 0.30 to < 0.50 , indicates a medium effect that is comparable to relevant effect sizes in terms of differences between two means.¹¹ Divergent validity was analysed according to correlations between measurements of moral behaviour that were expected to be unrelated (ie, no statistically significant correlation).

3.10 | Measurements

3.10.1 | Sociodemographic characteristics

In this study, the following sociodemographic characteristics were self-reported: age, gender, working environment, and religion.

3.10.2 | Moral Sensitivity Questionnaire

In order to adjust the psychiatry oriented MSQ²⁹ for use in research populations of NPs and PAs, it was necessary to rephrase nine items. For example, references to "psychiatrist" were replaced with references to either "NP" or "PA" in two items, and the terms "psychiatric care" and "psychiatric practice" were rephrased as "care" or "practice," respectively, in five items. Furthermore, two items referring to "treatment under the Mental Health Act" were rephrased to refer to "care provided to incapacitated patients." Respondents were asked to use a 7-point Likert scale (1 = fully disagree to 7 = fully agree) to indicate how they perceived their own manner of decision-making in moral dilemmas. Each of the items reflected either a paternalistic or deliberate attitude, as assumed in a previous feasibility study conducted among PA students.

For each scale, item scores were coded, summed, and transformed into a scale ranging from 0 to 100 (with higher scores reflecting greater sensitivity or insensitivity) and calculated by subtracting the lowest possible scale score from the raw summed scale score, divided by the range of scores on the scale and multiplied by 100.

The instruments used for testing the convergent and divergent validity of the hypothesized latent MSQ constructs (as found in the feasibility study), as described in Appendix S1, include the following: a) the BCPH scale; b) the EAS; c) the MDS, and d) the DIT-N2. Appendix S1 contains the following references: 2, 4, 5, 43, 44, 46. All of the scales used in the current study were transformed towards normality through a two-step transformation process, conducted prior to the analyses.⁵³

3.11 | Translation of measurement instruments

Questionnaires were translated into Dutch following the procedure proposed by Guillemin and colleagues.²² First, two certified translators working independently of each other translated the original English version of the questionnaires into Dutch. Second, two other certified translators each back-translated the Dutch translation into English. The resulting English versions were compared with the originals and discrepancies were discussed and resolved by consensus between the researchers LK, GJ, and BM.

3.12 | Hypotheses regarding convergent and divergent validity

We examined the strength of the correlation coefficients as indicators of conceptual overlap between paternalistic and deliberate attitudes according to four concurrent self-report measurements. The following hypotheses were formulated:

3.12.1 | Divergent validity

Although paternalistic and moral deliberate attitudes are usually pictured as two opposites, the traits are nevertheless expected to be independent of each other. This is because the features of both traits are not incompatible with each other. For example, an important feature of a deliberate attitude is valuing to have a relationship with patients. This is not necessarily in contradiction with one's inclination to follow rules and regulations and base one's decision on medical practice (which is a feature of a paternalistic attitude). So, even though a healthcare professional may be aimed at having a relationship with a patient and treat the patient with respect (ie, deliberate attitude), still the healthcare professional can decide to base his/her decision on medical knowledge or regulations, even if that is against the will of a patient (ie, paternalistic attitude), if he/she really thinks this is in the best interest of the patient. We therefore hypothesize that

H1 :There is no correlation between the two scales measuring a deliberate attitude and paternalistic attitude respectively.

We further assume that paternalistic and the deliberate attitude are different from moral reasoning. After all, moral reasoning reflects a cognitive, intra-personal process, in which a person engages in a deliberation on what is the moral thing to do. The paternalistic and deliberative attitudes refer more to a person's general preferences for how they relate to patients. This is more an inter-personal issue and reflect one's tendencies of how to behave in a patient-professional relationship. We therefore hypothesize that

H2 :Neither the paternalistic nor the deliberate attitude scale is expected to have any significant overlap with the level of moral reasoning (DIT-N2).

3.12.2 | Convergent validity

We assume that deliberative attitude is related to several ethical tendencies of people. After all, the main characteristic of a deliberative attitude to value a respectful relationship with patients. Hence, deliberate patient-centred decision-making is expected to more strongly possess preferences and traits that are ethics-related. For example, they may be more likely to adhere to the fundamental principle of "First, do not harm"; they may be more likely to value ethical considerations in care and are less inclined to use mechanisms of moral disengagement. We therefore hypothesize that

H3.1 :There is a positive correlation between a moral deliberate attitude and the BCPH scale.

H3.2 :There is a positive correlation between a moral deliberate attitude and the EAS.

H3.3 :There is a negative correlation between the moral deliberate attitude and the MDS.

On the other hand, a paternalistic attitude may differently relate to ethical tendencies. Because people with a paternalistic attitude are more likely to avoid empathizing with the patient's dilemmas and prefer rules and regulations, they may exhibit little inclination to advocate the need for ethics in patient care. Hence, they may also have a weaker tendency to control impulses of morally disengaged behaviour. We therefore hypothesize that

H4.1 :The NPs' and PAs' paternalistic attitudes are expected to have no correlation with the EAS.

H4.2 :The NPs' and PAs' paternalistic attitudes are expected to have a positive correlation with the MDS.

Despite the expectation that a paternalistic attitude is negatively related to the need to advocate for ethics and positively to moral disengagement, this does not necessarily mean that they do not care about the "First, do not harm" principle. In fact, people who score high on paternalism may also adhere to this principle, although they try to achieve this in different ways than people who score high of moral deliberation (namely, through the adherence to regulations and expert information rather than through building a relationship with the patient). Therefore, we hypothesize

H4.3 :There is a positive correlation between a paternalistic attitude and the BCPH scale.

4 | RESULTS

4.1 | Sociodemographic characteristics

An overview of the sociodemographic characteristics broken down by professional group (ie, NPs and PAs) is provided in Table 1. The mean age of the PAs (42.5 years) was lower than that of the NPs (48.8 years). The two groups did not differ in terms of gender, religion, or work setting. Moreover, no significant differences were found between the two professions with regard to the prevalence of politically conservative or liberal orientations. Based on these results, we considered it acceptable to merge the samples for analyses.

4.2 | Confirmation of the latent MSQ dimensions

The Velicer's Minimum Average Partial (MAP) test yielded three factors to be retained (MAP squared: 0.017) for consecutive analysis.¹³ This test was followed by EFA, which was also based on a polychoric correlation matrix³⁹ using principal axis factoring and oblimin-quartimin-Q rotation.⁶ A three-factor model converged well, detecting 16 items with loadings exceeding 0.40 and having no cross-loadings exceeding 0.20 on any other factors.¹² CFA was used to assess the model fit of the three-factor solution retained from the MSQ item pool. At first glance, the model fit parameters for the CFA appeared to indicate an acceptable model fit. Upon closer examination, however, five items (all of which were clustered in one factor) had standardized loadings well below 0.40. These items were therefore eliminated from the model. Finally, a two-factor solution demonstrated a very good model fit, with $\chi^2/df = 1.168$ and a significance of $P = 0.223$, RMSEA = 0.033 (CI 90% lower bound = 0.000 and CI 90% upper bound = 0.069), SRMR = 0.0622, CFI = 0.965, and

GFI = 0.951. The lower AIC value (101.218) that was found for the two-factor solution, as compared with the AIC of the initially anticipated three-factor solution model (203.371), provided evidence that eliminating the weak items was necessary in order to establish a good model with two latent constructs. Despite this good model fit, three items in Factor 1 continued to exhibit standardized regression weights less than 0.40: Item 4 ("When I need to make a decision contrary to the will of a patient, I do so according to my opinion about what is good care"), Item 12 ("If I am unacquainted with the case history of a patient, I follow the rules that are available"), and Item 16 ("I think that good care often includes making decisions for the patient"). Additional CFA, in which the three items with factor loadings less than 0.40 were separately excluded from the analysis, indicated deterioration in the model fit parameters. Once all of the items with factor loadings less than 0.40 were excluded from the analysis, the model became unidentifiable. For that reason, these three items were not included in Factor 1.

As in the feasibility study, the results of this study indicate that the MSQ item pool represents two dimensions. In light of these findings, we conducted a meticulous recheck of the content of the factors retained from the MSQ items. This led to the conclusion that the findings were congruent with the content of the moral deliberation and paternalism of the physician-patient relationship models hypothesized by Emanuel and Emanuel¹⁷ and empirically tested by Falkum and Førde.¹⁹ Factor 1 thus represents a construct that we have labelled the "Paternalistic Attitude Scale" (MSQ-PATER), as the items reflect the thinking that one is acting in the patient's best interest while disregarding the patient's will in the matter. In this study, we defined the concept of paternalism as "a tendency to avoid empathizing with the patient's dilemmas and taking decisions with a strong emphasis on rules and regulations, as well as on medical knowledge and practice, and based on professional opinions about the best treatment options". Factor 2 represents a construct that

TABLE 1 Sociodemographic characteristics of participants, stratified by profession (NPs and PAs)

Sociodemographic characteristics		Physician assistant N = 88	Nurse practitioner N = 67	Total N = 155	(P-value)
Age mean (SD)		42.5 (8.4)	48.8 (8.7)	45.2 (9.1)	<.001 ^a
Gender	Female N (%)	56 (63.6)	53 (79.1)	109 (70.3%)	.05 ^b
	Male N (%)	32 (36.4)	14 (20.9)	46 (29.7%)	
Religion	Not religious	48 (54.5)	35 (52.3)	83 (53.5%)	.54 ^b
	No denomination, but spiritual	3 (3.4)	4 (4.5)	7 (4.5%)	
	Christian	35 (39.8)	25 (37.3)	60 (38.7%)	
	Islam	1 (1.1)	0	1 (0.7%)	
	Other religions	1	3 (4.5)	4 (2.6%)	
Working environment	Hospital, N (%)	64 (72.7%)	49 (73.1%)	113 (72.9%)	.58 ^b
	General practice, N (%)	13 (14.8%)	7 (10.5%)	20 (12.9%)	
	Mental health, N (%)	3 (3.4%)	6 (9%)	9 (5.8%)	
	Disability care, N (%)	1 (1.1%)	1 (1.5%)	2 (1.3%)	
	Other, N (%)	7 (8%)	4 (5.9%)	11 (7.1%)	
Political orientation	Conservative, N (%)	15 (17%)	6 (9%)	21 (13.5%)	.14 ^b
	Liberal, N (%)	73 (83%)	61 (91%)	134 (86.5%)	

^aIndependent-sample t-test.

^bDifference between proportions test.

we describe as the "Deliberate Attitude Scale" (MSQ-DELIB). All of the items in this scale centre on the dimension of a professional relationship between the clinician and the patient, as indicated by such socio-cognitive, affective themes as "autonomy," "relationship," "giving respect," and "providing patients with insight." The concept of moral deliberation thus implies that NPs and PAs engage in careful and serious deliberation before making any important medical decisions. This finding is based on the independent content analysis of MSQ items. We defined the concept of moral deliberation, measured by the MSQ-DELIB as "medical decision-making aimed at helping patients to determine the best health-related values that can be realized in the clinical situation after considerable deliberation."

The new scales and their assigned items are presented in Figure 1, along with (a) their respective standardized regression weights (ie, factor loadings) from latent constructs to the variables measured and (b) their standard errors. All beta weights were statistically significant ($P < .001$). Both the "MSQ-PATER" and the "MSQ-DELIB" scales had internal consistency of 0.70.

4.3 | Construct validity of the MSQ-DELIB and MSQ-PATER scales

4.3.1 | Divergent validity

As demonstrated by the results of CFA, the constructs of MSQ-PATER and MSQ-DELIB had no conceptual overlap (see Figure 1) and

were not correlated ($r = .03$). The hypothesis concerning the divergent validity of the MSQ-DELIB and MSQ-PATER scales (H1) was confirmed. The results further provide evidence of divergent validity for both scales, given the absence of any correlation between either scale and the DIT-N2 (H2).

4.3.2 | Convergent validity

Our analyses revealed several statistically significant correlations, which could be used to establish convergent validity, as hypothesized. First, (H3.1), the MSQ-DELIB scale is positively correlated with a) the "Behavioral Control targeted at Preventing Harm (BCPH)" scale ($r = .34$) and b) (H3.2) the "Ethics Advocacy Attitude Scale (EAS)" ($r = .42$), and it is thus negatively correlated (H3.3) with c) "Moral Disengagement Total (MDS)" ($r = -.17$). Second, there is a significant correlation between the MSQ-PATER scale and a) the BCPH scale ($r = .17$) and b) MDS ($r = .20$), with no inclination towards ethics advocacy ($-.06$, ns), as hypothesized (H4.1, H4.2, H4.3).

Given that the correlation between paternalism (MSQ-PATER) and BCPH was weaker than the correlation between moral deliberation (MSQ-DELIB) and BCPH, it could be that care providers who tend to follow a model of negotiation in their interactions with patients are likely to attach greater importance to the prevention of harm ($r = .34$) than are care providers who are more inclined towards "command management" ($r = .17$).

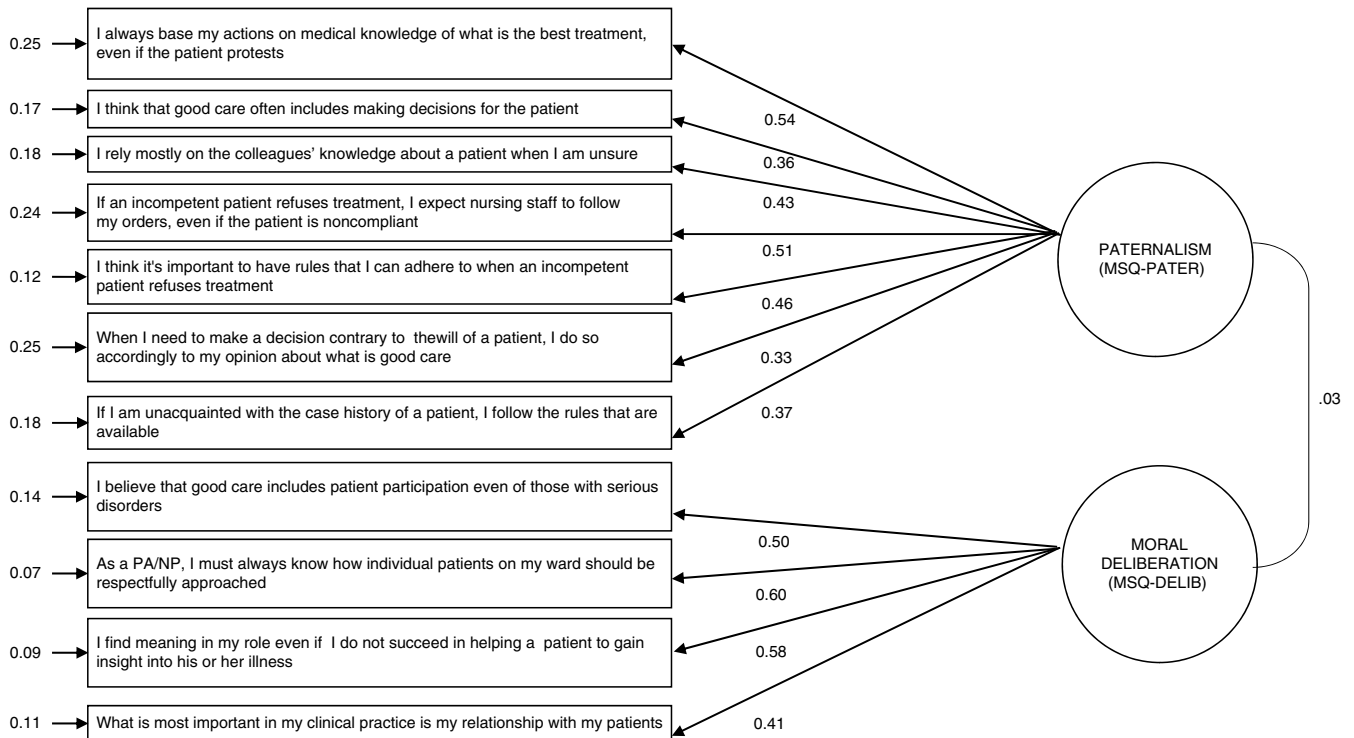


FIGURE 1 Confirmatory factor analysis (CFA) two-factor solution with standardized estimates on the Moral Sensitivity Questionnaire (MSQ) items. MSQ-DELIB; Deliberate Attitude Scale; MSQ-PATER, Paternalistic Attitude Scale; NP, nurse practitioner; PA, physician assistant

	1	2	Alpha
1 Deliberate Attitude Scale (MSQ-DELIB)			0.70
2 Paternalistic Attitude scale (MSQ-PATER)	.03		0.70
3 Behavioral Control targeted at Preventing Harm (BCPH) scale	.34***	.17*	0.72
4 Ethics Advocacy Scale (EAS)	.42***	-.06	0.72
5 Moral Disengagement Total (MDS)	-.17*	.20*	0.85
6 Defining issues test (N2-index)	-.04	-.00	0.80

* $P < .05$.; ** $P < .01$.

The Cronbach's alpha values, which serve as indicators of internal consistency for all of the scales used, are included in the right-hand column of Table 2.

5 | DISCUSSION

Prior to the current study, a feasibility study was conducted among students in a master's program for PAs, in order to test the psychometric properties of several instruments, some of which have also been included in this study. The objectives of the feasibility study also included modifying and validating a revised version of the MSQ, as developed by Lützné and colleagues,²⁹ for use among NPs and PAs. In that study, however, a simple EFA using Varimax rotation revealed 10 latent components, instead of the six that were theoretically assumed by Lützné and colleagues. A subsequent narrative review of the literature revealed that international scholars building on the work of Lützné and colleagues^{10,14,23,54} had also been unable to reproduce the six factors proposed for the original instrument.

In light of these developments, we decided not to re-evaluate the six-component structure, but instead to modify and validate a revised version of the instrument. The outcomes of the current study support the validity and reliability of two new scales: MSQ-DELIB and MSQ-PATER. These findings are obviously preliminary, given that this is the first time that the validity of these new MSQ dimensions have been evaluated among Dutch NPs and PAs. The solid methodology of this study nevertheless contributes to these two new scales, which were established through CFA to produce a two-factor solution with good model fit and satisfactory internal consistency (reliability estimates). Our findings are in line with work by Emanuel and Emanuel, who identify deliberative and paternalistic attitudes as two of the four parts of the clinician-patient relationship (the other two being informative and interpretive attitudes).¹⁷ The first factor identified in the current study was labelled the MSQ-PATER. Examination of the seven items of these scale reveals that they centre on such themes as "following the rules," "personal opinion about good care," and "best treatment." All of these themes are consistent with the general perception of paternalism, that is, one is acting in the patient's best interest, while disregarding the patient's will in the matter. In this light, we formulated the following operational definition of paternalism was formulated to capture the meaning of the MSQ-PATER: "The tendency toward paternalism in medical decision-making is activated by a clinician's

TABLE 2 Convergent and divergent validity of the Deliberate Attitude Scale (MSQ-DELIB) and the Paternalistic Attitude Scale (MSQ-PATER)

preference for arguments based on rules and regulations. Decisions are established through the interplay between the clinician's own opinion, medical knowledge, and experience, as well as the opinions of others, while ignoring the will of the patient."

The second factor identified in the current study relates to the dimension of the professional relationship between the clinician and the patient, as indicated by affective, socio-cognitive considerations (eg, consideration for "autonomy," "relationship," "giving respect," and "providing patients with insight"). We combined these four items to form the MSQ-DELIB, which reflects "the clinician's aim of helping patients to determine the best health-related values that can be realized in the clinical situation".¹⁹ Such an aim requires morally sensitive reflection on the ethical consequences of decisions in treatment. It encompasses the desire to treat patients with the proper respect and to find meaningfulness in working with patients. A such, it is broadly consistent with the definition proposed by Lützné and colleagues²⁹: "the contextual and intuitive understanding of the vulnerability of a person's situation and insight into the ethical consequences of decisions made on behalf of the person" (p. 474). Be that as it may, based on our content analysis, the items identified in Factor 2 provide no basis for adhering to Lützné's concept of moral sensitivity. The items in Factor 2 do not reflect sensitivity for moral issues of the patient but rather represents an attitude towards moral dialogue.

After the two new measurement scales were developed, they were tested for construct validity. These tests yielded favourable convergent and divergent outcomes thus indicating good construct validity. As hypothesized, the indicator of moral reasoning (DIT-N2) showed no significant correlation with the two new scales. This is in line with a review by Muriel Bebeau⁸ positing that one could question whether the four components of the FCM should necessarily be correlated. Such questions are particularly justified in light of Bebeau's view on this assumption: "Conclusions to date suggest that measures of the components are assessing abilities that are distinct from one another." In addition, and more importantly, the two new scales do not measure a cognitive process but a morally deliberate and paternalistic attitude. Convergent validity is thus not something that should actually be expected.

Once the MSQ-DELIB and MSQ-PATER were confirmed as valid scales—measuring moral deliberation and paternalism, respectively—we considered the question of why only 11 of the 30 items in the original MSQ²⁹ addressed in this study were retained. Given that Lützné and colleagues propose a theoretical construct that includes



30 operationalized aspects that are presumed to measure the six domains of moral sensitivity, it is remarkable to note that, in a more recent study,²⁸ only 9 of those 30 items emerge as valid operationalizations for measuring the construct of moral sensitivity. The current study used the same pool of 30 items from the original physician's version of the MSQ, and factor analysis was used in order to assess whether the items correlated with the underlying construct. Our results indicate a comparable reduction in the number of items. This suggests that the 9-item MSQ of Lützné and colleagues may not actually measure moral sensitivity.²⁸ Moreover, our results provide evidence that the two new scales, which are based on the original items of the MSQ, measure levels of moral deliberation (MSQ-DELIB), and paternalism (MSQ-PATER) that are broadly in line with the findings of Falkum and Førde (2001). Our results also adds to Falkum and Førde as the scales in the current study are presented in the first person (eg, "When I need to make a decision contrary to the will of a patient, I do so accordingly to my opinion about what is good care"), whereas Falkum and Førde (2001) present statements in the third person (eg, "The physician expert should decide"). As such, the scales may be more likely to reflect a deliberate predisposition towards a paternalistic and moral stance, rather than any broader, general values concerning moral deliberation and paternalism. We do not wish to make any value judgement concerning whether a deliberate or a paternalistic attitude is better. Even though it may seem that we now regard a paternalistic approach as inappropriate within the clinician-patient relationship, it might be the case that this is a reflection of our contemporary culture. However, societies change, and it is possible that, in a future era, a paternalistic or a deliberate attitude is differently valued than it is now.

5.1 | Strengths and limitations

One of the major strengths of this research lies in its robust study design (employing CFA), including the assessment of the convergent and divergent validity of the scales. Another important strength is that the results are based on a representative sample that reflects the characteristics of the PA and NP workforces in the Netherlands with regard to gender and age.⁸ For this reason, the results can be generalized to a certain extent to both the NP and PA workforces at large. We also expect that the MSQ-DELIB and MSQ-PATER are applicable to other healthcare professionals who share a comparable framework regarding knowledge, skills, and legal boundaries (eg, MDs). One weakness of the study, however, is that the stability of the instrument (ie, its test-retest reliability) was not assessed. The study design did not allow for testing the two scales for longitudinal validity. By definition, cross-sectional studies cannot examine the stability of the attitudes or traits of subjects over time. A follow-up study will investigate longitudinal psychometric research questions focusing on the test-retest stability of the instruments.

Another limitation of the current study is that no a priori calculations of sample size were performed. Given the lack of studies assessing moral sensitivity among PAs and NPs, however, the field

was open to exploration. Given the actual sample size addressed in the study (155 records), it may not be necessary to assume that the results of our CFA were compromised by the sample size. Although we are aware of the various rules and opinions used to determine the sample size needed for CFA, this study was based on a convenience sample with an $N (=155)$ to P (number of items = 11) ratio of 14.1. We therefore felt confident that the assumption underlying CFA was not violated.²⁰ Finally, even though our analyses revealed statistically significant correlations for both the MSQ-DELIB and the MSQ-PATER scales based on convergent and divergent instruments, the explained variances were relatively low.

6 | CONCLUSION

The results of this study provide evidence of two new latent dimensions derived from the items of the original MSQ. Because the scales MSQ-DELIB and MSQ-PATER have been validated only for NPs and PAs, further exploration and validation may be needed before the three items with loadings less than 0.40 from Factor 1 (MSQ-PATER) can be eliminated. To this end, these three items should be rephrased to be more closely aligned with the target construct of paternalism. The most important contribution of this study is the introduction of the two new MSQ scales, both of which have good structural and construct validity. They therefore have the potential to serve as an impetus for structural equation modelling in relation to analysing paths within the four-component model of moral behaviour. Given the increase in the number of PAs and NPs throughout the world, such efforts will require validation in a number of countries.

7 | IMPLICATIONS

Healthcare professionals are quite likely to perceive working with patients as a natural calling, prompted by an intrinsic motivation to do good. Such inherent sympathy and empathy, however, which is perceived as beneficent, may become blurred by blind spots with regard to the personal attitudes held by individual clinicians and, consequently, their behaviour. With the exception of prejudice, the majority of complaints and disciplinary cases are based on either miscommunication or a feeling on the part of patients that they have been treated discourteously. With this in mind, both the MSQ-DELIB and MSQ-PATER could be used and applied as self-report tools for clinicians who would like to become more aware of their own underlying attitudes (eg, moral deliberation and paternalism) when communicating with patients. The two scales could also function as a type of "thermometer" with which to assess the moral climate and the work-related moral stress experienced by health employees.²⁷

AUTHOR CONTRIBUTIONS

Luppo Kuilman: study conception/design; acquisition (and storage) of data; analysis and interpretation of data; writing and critical revision of the manuscript.

Gerard Jansen: study conception/design; drafting manuscript; critical revision.

Laetitia Mulder: drafting of the manuscript; critical revision.

Berrie Middel: study conception/design; supervision on analysis and interpretation of data; drafting of the manuscript; critical revision.

Petrie Roodbol: study conception/design; critical revision.

ORCID

Luppo Kuilman  <https://orcid.org/0000-0003-3483-3265>

Gerard J. Jansen  <https://orcid.org/0000-0003-2464-4188>

Laetitia M. Mulder  <https://orcid.org/0000-0003-3203-7601>

REFERENCES

- Abma TA, Molewijk B, Widdershoven GA. Good care in ongoing dialogue. Improving the quality of care through moral deliberation and responsive evaluation. *Health Care Anal.* 2009;17(3):217-235.
- Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process.* 1991;50(2):179-211.
- Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control.* 1974;19(6):716-723.
- Bandura A. Moral disengagement in the perpetration of inhumanities. *Pers Soc Psychol Rev.* 1999;3(3):193-209.
- Bandura A, Barbaranelli C, Caprara G. Mechanisms of moral disengagement in the exercise of moral agency. *J Pers Soc Psycho.* 1996;71(2):364-374.
- Basto M, Pereira JM. An SPSS R-menu for ordinal factor analysis. *J Stat Softw.* 2012;46(4):1-29.
- Beauchamp TL, Childress JF. *Principles of Biomedical Ethics.* USA: Oxford University Press; 2001.
- Bebeau MJ. The defining issues test and the four-component model: contributions to professional education. *Journal of Moral Education.* 2002;31(3):271-295.
- Bernstein IH, Nunnally J. *Psychometric theory.* New York: McGraw-Hill/Oliva, TA, Oliver, RL, & MacMillan, IC (1992). *A catastrophe model for developing service satisfaction strategies.* *Journal of Marketing.* 1994; 56:83-95.
- Borhani F, Abbaszadeh A, Mohamadi E, Ghasemi E, Hoseinabad-Farahani MJ. Moral sensitivity and moral distress in Iranian critical care nurses. *Nurs Ethics.* 2017;24(4):474-482.
- Cohen J. *Statistical power analysis for the behavioral sciences.* Hillsdale NJ: Lawrence Earlbaum Associates. 1988;2.
- Costello AB. Getting the most from your analysis. *Pan.* 2009;12(2): 131-146.
- Courtney MGR, Gordon M. Determining the number of factors to retain in EFA: using the SPSS R-menu v2. 0 to make more judicious estimations. *Pract Assess Res Eval.* 2013;18(8):1-14.
- Dalla Nora CR, Zoboli EL, Vieira MM. Validation of a Brazilian version of the moral sensitivity questionnaire. *Nurs Ethics.* 2017; 0969733017720849.
- De Haan J. The definition of moral dilemmas: a logical problem. *Ethical Theory and Moral Practice.* 2001;4(3):267-284.
- Driever EM, Stiggelbout AM, Brand PL. Shared decision making: physicians' preferred role, usual role and their perception of its key components. *Patient Educ Couns.* 2019.
- Emanuel EJ, Emanuel LL. Four models of the physician-patient relationship. *JAMA.* 1992;267(16):2221-2226.
- Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods.* 1999;4(3):272-299.
- Falkum E, Førde R. Paternalism, patient autonomy, and moral deliberation in the physician-patient relationship: attitudes among norwegian physicians. *Soc Sci Med.* 2001;52(2):239-248.
- Gagne P, Hancock GR. Measurement model quality, sample size, and solution propriety in confirmatory factor models. *Multivar Behav Res.* 2006;41(1):65-83.
- General Assembly of the World Medical Association. World Medical Association declaration of Helsinki: ethical principles for medical research involving human subjects. *J Am Coll Dent.* 2014;81(3):14-18.
- Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol.* 1993;46(12):1417-1432.
- Han S, Kim J, Kim Y, Ahn S. Validation of a Korean version of the moral sensitivity questionnaire. *Nurs Ethics.* 2010;17(1):99-105.
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model Multidiscip J.* 1999;6(1):1-55.
- Jacobson P. Empowering the physician-patient relationship: the effect of the internet. *Partnership: The Canadian Journal of Library and Information Practice and Research.* 2007;2(1).
- Kuilman L, Jansen GJ, Middel B, Mulder LB, Roodbol PF. Moral reasoning explained by personality traits and moral disengagement: a study among Dutch nurse practitioners and physician assistants. *J Adv Nurs.* 2019;75(6):1252-1262.
- Lützn K, Blom T, Ewalds-Kvist B, Winch S. Moral stress, moral climate and moral sensitivity among psychiatric professionals. *Nurs Ethics.* 2010;17(2):213-224.
- Lützn K, Dahlqvist V, Eriksson S, Norberg A. Developing the concept of moral sensitivity in health care practice. *Nurs Ethics.* 2006;13(2): 187-196.
- Lützn K, Evertzon M, Nordin C. Moral sensitivity in psychiatric practice. *Nurs Ethics.* 1997;4(6):472-482.
- Lützn K, Johansson A, Nordstrom G. Moral sensitivity: some differences between nurses and physicians. *Nurs Ethics.* 2000;7(6): 520-530.
- Lützn K, Nordin C, Brolin G. Conceptualization and instrumentation of nurses moral sensitivity in psychiatric practice. *Int J Methods Psychiatr Res.* 1994;4(4):241-248.
- Lützn K, Nordström G, Evertzon M. Moral sensitivity in nursing practice. *Scand J Caring Sci.* 1995;9(3):131-138.
- Lützn K, Nordstrom G, Evertzon M. Moral sensitivity in nursing practice. *Scand J Caring Sci.* 1995;9(3):131-138.
- Maier CB. The role of governance in implementing task-shifting from physicians to nurses in advanced roles in Europe, US, Canada, New Zealand and Australia. *Health Policy.* 2015;119(12):1627-1635.
- McQuitty S. Statistical power and structural equation models in business research. *J Bus Res.* 2004;57(2):175-183.
- Merkle F, Ritsema TS, Bauer S, Kuilman L. The physician assistant: shifting the paradigm of European medical practice? *HSR Proc Intensive Care Cardiovasc Anesth.* 2011;3(4):255-262.
- Miles J, Shevlin M. A time and a place for incremental fit indices. *Personal Individ Differ.* 2007;42(5):869-874.
- Newcombe RG, Altman DG. Proportions and their differences. *Statistics with Confidence.* 2000;2:45-56.
- Olsson U. Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika.* 1979;44(4):443-460.
- Pellegrino ED. Toward a reconstruction of medical morality. *Am J Bioeth.* 2006;6(2):65-71.
- Polit DF, Beck CT. *Nursing Research: Principles and Methods.* Lippincott Williams & Wilkins; 2004.
- Quill TE, Brody H. Physician recommendations and patient autonomy: finding a balance between physician power and patient choice. *Ann Intern Med.* 1996;125(9):763-769.
- Raaijmakers Q, Engels R, Van Hoof A. Delinquency and moral reasoning in adolescence and young adulthood. *Int J Behav Dev.* 2005;29(3): 247-258.
- Rest J, Thoma SJ, Narvaez D, Bebeau MJ. Alchemy and beyond: indexing the defining issues test. *J Educ Psychol.* 1997;89(3):498-507.

45. Rest JR. *Moral Development: Advances in Research and Theory*. New York: Praeger; 1986.
46. Rest JR. *DIT manual*. Minnesota, MN: Center for the Study of Ethical Development; 1990.
47. Robinson JH, Callister LC, Berry JA, Dearing KA. Patient-centered care and adherence: definitions and applications to improve outcomes. *J Am Assoc Nurse Pract*. 2008;20(12):600-607.
48. Ruscio J, Roche B. Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychol Assess*. 2012;24(2):282-292.
49. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol Res Online*. 2003;8(2):23-74.
50. Siegler M. The progression of medicine: from physician paternalism to patient autonomy to bureaucratic parsimony. *Arch Intern Med*. 1985; 145(4):713-715.
51. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to their Development and Use*. USA: Oxford University Press; 2015.
52. Tabachnick B, Fidell L. Multivariate analysis of variance and covariance. *Using Multivariate Statistics*. 2007;3:402-407.
53. Templeton GF, Burney LL. Using a two-step transformation to address non-normality from a business value of information technology perspective. *J Inf Syst*. 2016.
54. Yilmaz Sahin S, Iyigun E, Acikel C. Validity and reliability of a Turkish version of the modified moral sensitivity questionnaire for student nurses. *Ethics Behav*. 2015;25(4):351-359.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Kuilman L, Jansen GJ, Mulder LM, Middel B, Roodbol PF. Re-assessing the validity of the Moral Sensitivity Questionnaire (MSQ): Two new scales for moral deliberation and paternalism. *J Eval Clin Pract*. 2020;1-11. <https://doi.org/10.1111/jep.13353>