

RESEARCH

Open Access



Development of a theory-informed questionnaire to assess the acceptability of healthcare interventions

Mandeep Sekhon^{1,2*}, Martin Cartwright¹ and Jill J. Francis^{1,3,4}

Abstract

Background: The theoretical framework of acceptability (TFA) was developed in response to recommendations that acceptability should be assessed in the design, evaluation and implementation phases of healthcare interventions. The TFA consists of seven component constructs (affective attitude, burden, ethicality, intervention coherence, opportunity costs, perceived effectiveness, and self-efficacy) that can help to identify characteristics of interventions that may be improved. The aim of this study was to develop a generic TFA questionnaire that can be adapted to assess acceptability of any healthcare intervention.

Methods: Two intervention-specific acceptability questionnaires based on the TFA were developed using a 5-step pre-validation method for developing patient-reported outcome instruments: 1) item generation; 2) item de-duplication; 3) item reduction and creation; 4) assessment of discriminant content validity against a pre-specified framework (TFA); 5) feedback from key stakeholders.

Next, a generic TFA-based questionnaire was developed and applied to assess prospective and retrospective acceptability of the COVID-19 vaccine. A think-aloud method was employed with two samples: 10 participants who self-reported intention to have the COVID-19 vaccine, and 10 participants who self-reported receiving a first dose of the vaccine.

Results: 1) The item pool contained 138 items, identified from primary papers included in an overview of reviews. 2) There were no duplicate items. 3) 107 items were discarded; 35 new items were created to maximise coverage of the seven TFA constructs. 4) 33 items met criteria for discriminant content validity and were reduced to two intervention-specific acceptability questionnaires, each with eight items. 5) Feedback from key stakeholders resulted in refinement of item wording, which was then adapted to develop a generic TFA-based questionnaire.

For prospective and retrospective versions of the questionnaire, no participants identified problems with understanding and answering items reflecting four TFA constructs: affective attitude, burden, perceived effectiveness, opportunity costs. Some participants encountered problems with items reflecting three constructs: ethicality, intervention coherence, self-efficacy.

Conclusions: A generic questionnaire for assessing intervention acceptability from the perspectives of intervention recipients was developed using methods for creating participant-reported outcome measures, informed by theory,

*Correspondence: Mandeep.sekhon@kcl.ac.uk

¹ School of Health Sciences, City University of London, Northampton Square, London EC1V 0JB, UK

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

previous research, and stakeholder input. The questionnaire provides researchers with an adaptable tool to measure acceptability across a range of healthcare interventions.

Keywords: Acceptability, Questionnaire development, Pre-validation methods, Healthcare intervention, Theoretical framework

Statement of contribution

What is known on this topic:

- It is important to assess the acceptability of healthcare interventions.
- A recently developed theoretical framework of acceptability (TFA) proposes seven component constructs (affective attitude, burden, ethicality, intervention coherence, opportunity costs, perceived effectiveness, and self-efficacy) to help identify characteristics of interventions that may be improved.

What this study adds:

- A generic, theoretically informed questionnaire for assessing acceptability of healthcare interventions has been developed.
- The comprehensibility, relevance, and answerability of the adapted items in the generic acceptability questionnaire were assessed using think-aloud methods.
- The generic TFA-based questionnaire is a brief and adaptable tool for researchers and clinicians to measure intervention/treatment acceptability across a range of healthcare settings.

Introduction

A health intervention that is not acceptable to the people who deliver or receive it is difficult to evaluate, as intervention facilitators are unlikely to deliver the key components faithfully and recipients are unlikely to engage with the intervention as required. Intervention acceptability from the perspective of patients and healthcare professionals has been proposed to have an impact on intervention implementation, uptake, adherence, intended outcomes and overall effectiveness [1–5].

The UK Medical Research Council (MRC) has published guidance for intervention developers and researchers, describing best practice methods for designing and evaluating complex interventions [4–6]. The guidance recommends that researchers assess intervention acceptability [5, 6] but provides few suggestions about how to do this.

In a systematic overview of reviews, we found that 55% of included reviews used behavioural measures as indicators of acceptability [7], such as total trial dropout rate and reasons for discontinuation [8–10]. However, there are several reasons, other than acceptability problems, that may explain why participants discontinue treatment or withdraw from an intervention [11]. Reasons may include lack of motivation to take part, mistrust of the research process [12] or personal circumstances external to the trial or the intervention [13]. Furthermore, reliance on measures of observed behaviour does not provide information on which aspects of an intervention are (un) acceptable and hence does not inform how to enhance acceptability. Notably, in our overview of reviews, we found that there was no standardised or validated intervention acceptability questionnaire [7].

The Theoretical Framework of Acceptability (TFA) [7, 14] can guide assessments of intervention acceptability across three temporal perspectives (before, during and after participation in an intervention) and from the perspectives of intervention deliverers and recipients. In some contexts, a questionnaire approach may offer a time-efficient way to identify potential problems with intervention acceptability, within all four phases of intervention development and evaluation identified in the MRC guidance, i.e., development, feasibility and piloting, evaluation, implementation [4–6].

Questionnaires are often considered a practical and cost-effective method for assessing participant outcomes (e.g., quality of life, emotional health, experienced symptoms). They have the advantage of being able to be administered to large samples, and to enable quantitative analysis including longitudinal assessments and direct comparisons between different trial arms and across different studies [15, 16].

To assess whether other researchers had developed a TFA-based acceptability questionnaire, we completed a forward citation search in Google scholar (October 2021), and identified two articles that report using quantitative measures of the TFA constructs [17, 18]. Neither article reported using an established method to develop their questionnaire, and neither questionnaire included an item assessing overall acceptability. Keyworth et al. [18] describe useful methods for analysing data from Likert scale responses to TFA items but reported problems with the wording of the Opportunity Costs item and potential issues

with the response format of this item and the Self-Efficacy item. Renko et al's [17] questionnaire contained several examples where the item wording did not accurately reflect the TFA definitions or items that conflated multiple TFA constructs. This study included no assessment of the comprehensibility of the items to participants. These studies highlight the challenges of developing a TFA questionnaire and demonstrate the need for a generic TFA questionnaire that can be used as the starting point for questionnaire development across health interventions.

The development and validation of questionnaires is a complex process [19]. Common methods to develop questionnaires include both inductive "bottom-up" approaches and deductive "top-down" approaches. The bottom-up approach focuses on generating items from empirical data often generated by exploratory qualitative methods (e.g., semi-structured interviews, focus groups) to ensure items represent the perspectives of the target population [20]. The top-down approach focuses on reviewing the literature to identify existing items [20, 21] or generating items based on pre-existing theory.

To develop the questionnaire, we combined methods for the development and pre-validation of participant-reported outcome measures [20] and methods used to establish the content validity of items in theoretically informed questionnaires [22] (see [Method](#) section and supplementary files 1 and 2 for details).

Aims and objectives

The aim of the current study was to develop a generic TFA questionnaire that can be adapted to assess acceptability of any healthcare intervention. The objectives were to:

- (1) adopt the 5-step pre-validation Patient reported Outcomes (PRO) method [20] and the Discriminant Content Validation (DCV) method [22] to develop two preliminary acceptability questionnaires based on the TFA (one for healthcare professionals and one for patients);
- (2) optimise the two preliminary questionnaires using feedback from key stakeholders on comprehensibility and relevance of items;
- (3) assess the comprehensibility, relevance and answerability of items in the generic acceptability questionnaire using think-aloud methods.

Methods

Context of preliminary questionnaires

Healthcare professional acceptability questionnaire

The first questionnaire was developed to assess acceptability to intervention recipients (healthcare professionals) of two feedback interventions delivered as part of

the AFFINITIE Research Programme which aimed to optimise audit-and-feedback processes to improve blood transfusion practice [23, 24]. Intervention 1 consisted of feedback reports that were "enhanced", compared with usual feedback delivered to hospital staff. Intervention 2 consisted of "follow-on support" (a web-based toolkit and telephone support) provided to hospital transfusion teams to help them plan their response to the feedback reports [25].

Patient acceptability questionnaire

The second questionnaire was developed to assess acceptability to intervention recipients of a new appointment-booking system. Recipients were patients who regularly attended an eye clinic for management of Benign Essential Blepharospasm (BEB) and Hemifacial Spasm (HFS) [26]. There were two conditions: standard care (i.e. appointments booked for patients by the clinic at approximately regular intervals) and the new service model (patient-initiated appointment booking service where patients called a nurse helpline for an appointment when their symptoms necessitated it). Trial participants were randomly allocated to either the standard service or patient-initiated appointment services for 9 months.

Step 1: item generation

Supplementary file 2 presents further details of the 5-step pre validation PRO method. To generate a pool of potentially relevant items, we adapted an empirical approach to selecting systematic review papers and their included primary papers, based on the findings from our overview of systematic reviews [7] to determine how others have approached the issue of assessing acceptability. In our overview of reviews, the findings indicated that review authors assessed acceptability either via measures of observed behaviour, self-reported measures, or a combination of both. Next, an inclusive approach was applied to the review papers, and their included primary papers that reported using self-report measures to assess acceptability (or related constructs) if authors stated acceptability had been assessed via:

- A measure of satisfaction
- Reasons for discontinuation
- Qualitative open-ended interviews
- User perspectives and evaluations of the intervention

Reviews that identified assessing acceptability via 'observed measures of behaviour' were not reviewed or included in our item pool as they would not have been relevant to our primary aim of developing a generic TFA questionnaire.

Primary papers that reported assessing acceptability, as described above, were eligible if they met any of following inclusion criteria:

- a) Exact item wording and response format is presented in the text or in the appendix or supplementary file of the paper
- b) Wording of interview questions used to assess acceptability are reported in the paper
- c) Descriptions of reported reasons for discontinuation reflected assessments of acceptability reported in the paper, e.g., reasons for discontinuation provided by participants included side effects of treatment [27]; preference for treatment choice [28]
- d) Descriptions of user perspectives and evaluations applied to assess intervention acceptability are reported in the text of the paper (e.g., evaluations of programme acceptability among programme planners, policy makers and members of the community) [29].

Information relating to the origin of the item, response format, content and wording of interview questions and descriptions of reasons for dropout, user perspectives and evaluations was entered into an Excel database (the item pool).

Step 2: Deduplication

Extracted items were reviewed for the three types of duplication by one author as suggested by Prior et al. (2011) [20] (Supplementary file 2).

Step 3: item reduction and item creation

Two researchers independently reviewed items extracted in Step 1 and removed items:

- a) If items were specific to an intervention and non-generalisable (e.g., do you follow a special diabetes diet?);
- b) If reasons for discontinuation and descriptions of user perspectives and evaluations of an intervention could not be reworded as a question (e.g., “loss to follow up, other reasons”).

To maximise coverage of the TFA constructs, one author drafted new items based on the definitions of the seven TFA constructs (Table 1) for the healthcare professional questionnaire and the patient questionnaire. The new items were specific to each intervention, and the temporal perspective was also represented in item wording. For example, in the BEB/HFS questionnaire, not all TFA constructs were appropriate for assessing the acceptability of the standard service (control condition).

Participants receiving standard care did not perform a behaviour (i.e., book their own appointment) because the next appointment was scheduled by their treating healthcare professional [26]. Thus, the constructs of burden and self-efficacy were not relevant. The response options of the new items also reflected the TFA constructs (Table 1).

Step 4: assessment of content coverage against a pre-existing theoretical framework

To test the discriminant content validity of items against the seven TFA constructs, the DCV method was applied [22, 30]. Previous research on the number of judges required for judgement tasks suggests between 2 and 20 [31, 32]. All members of the Psychology Group within the Centre of Health Services Research at City University London were invited to take part. Eight members agreed to participate as expert judges, including four PhD students, three postdoctoral research fellows, one research assistant and one senior lecturer. Participants were provided with the construct definitions and an Excel table of 65 items to be classified. Five judges received a paper copy of the Excel table to complete in a face-to-face session; three elected to receive electronic versions of these materials and instructions on how to complete the DCV task.

Step 5: feedback on preliminary version of acceptability questionnaire from key stakeholders

Prior et al. (2011) [20] recommend conducting a think-aloud study with members of the target population to provide feedback on a newly developed questionnaire. Due to time constraints of both trials, this was not possible. Instead, two principal investigators (with clinical backgrounds) working on the AFFINITIE programme, and two patient representatives on the BEB/HFS study steering group, were asked to provide feedback on the draft questionnaire, which was emailed to them together with an invitation to read and comment on each item for comprehensibility and relevance.

Development of a generic TFA questionnaire

By reviewing both TFA questionnaires, a generic (not intervention-specific) acceptability questionnaire was developed. Item wording that was common to both the intervention-specific questionnaires was reviewed to assess whether the specific intervention descriptions and behavioural descriptions could be replaced with equivalent generic terms and phrases, “[intervention]” and “[engage with the intervention]”, so that all items could be applicable to a range of healthcare interventions (examples presented in Results section below).

The generic version of the questionnaire was adapted to assess perceptions of both prospective

Table 1 Generic form of TFA acceptability questionnaire

TFA Construct	Questionnaire item																				
Affective attitude <i>How an individual feels about the intervention</i>	Affective attitude: Did you like or dislike [intervention]? <table border="1"> <tr> <td>Strongly dislike</td> <td>Dislike</td> <td>No opinion</td> <td>Like</td> <td>Strongly like</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> </table> OR How comfortable did you feel [behaviour e.g. to engage with] [intervention]? <table border="1"> <tr> <td>Very uncomfortable</td> <td>Uncomfortable</td> <td>No opinion</td> <td>Comfortable</td> <td>Very comfortable</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> </table>	Strongly dislike	Dislike	No opinion	Like	Strongly like	1	2	3	4	5	Very uncomfortable	Uncomfortable	No opinion	Comfortable	Very comfortable	1	2	3	4	5
Strongly dislike	Dislike	No opinion	Like	Strongly like																	
1	2	3	4	5																	
Very uncomfortable	Uncomfortable	No opinion	Comfortable	Very comfortable																	
1	2	3	4	5																	
Burden <i>The amount of effort required to participate in the intervention</i>	How much effort did it take [behaviour e.g. to engage with] [intervention]? <table border="1"> <tr> <td>No effort at all</td> <td>A little effort</td> <td>No opinion</td> <td>A lot of effort</td> <td>Huge effort</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> </table>	No effort at all	A little effort	No opinion	A lot of effort	Huge effort	1	2	3	4	5										
No effort at all	A little effort	No opinion	A lot of effort	Huge effort																	
1	2	3	4	5																	
Ethicality <i>The extent to which the intervention has good fit with an individual's value system</i>	Ethical Consequences: There are moral or ethical consequences [behaviour e.g. to engage with] [intervention] <table border="1"> <tr> <td>Strongly disagree</td> <td>Disagree</td> <td>No opinion</td> <td>Agree</td> <td>Strongly agree</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> </table> OR How fair is [Intervention] for [people/ participants/ recipients] with [condition]? <table border="1"> <tr> <td>Very unfair</td> <td>Unfair</td> <td>No opinion</td> <td>Fair</td> <td>Very fair</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> </table>	Strongly disagree	Disagree	No opinion	Agree	Strongly agree	1	2	3	4	5	Very unfair	Unfair	No opinion	Fair	Very fair	1	2	3	4	5
Strongly disagree	Disagree	No opinion	Agree	Strongly agree																	
1	2	3	4	5																	
Very unfair	Unfair	No opinion	Fair	Very fair																	
1	2	3	4	5																	
Perceived effectiveness <i>The extent to which the intervention is have achieved its intended purpose</i>	The [intervention] has improve [behaviour/ condition/ clinical outcome]: <table border="1"> <tr> <td>Strongly disagree</td> <td>Disagree</td> <td>No opinion</td> <td>Agree</td> <td>Strongly agree</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> </table>	Strongly disagree	Disagree	No opinion	Agree	Strongly agree	1	2	3	4	5										
Strongly disagree	Disagree	No opinion	Agree	Strongly agree																	
1	2	3	4	5																	
Intervention coherence <i>The extent to which the participant understands how the intervention works</i>	It is clear to me how [intervention] will help [manage/ improve] my [behaviour/ condition/clinical outcome] <table border="1"> <tr> <td>Strongly disagree</td> <td>Disagree</td> <td>No opinion</td> <td>Agree</td> <td>Strongly agree</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> </table> *Please tell us more about your views	Strongly disagree	Disagree	No opinion	Agree	Strongly agree	1	2	3	4	5										
Strongly disagree	Disagree	No opinion	Agree	Strongly agree																	
1	2	3	4	5																	
Self-efficacy <i>The participants confidence that they can perform behaviour(s) required to participate in the intervention</i>	Self-efficacy: How confident did you feel about [behaviour e.g. engaging with] [intervention]? <table border="1"> <tr> <td>Very unconfident</td> <td>Unconfident</td> <td>No opinion</td> <td>Confident</td> <td>Very confident</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> </table>	Very unconfident	Unconfident	No opinion	Confident	Very confident	1	2	3	4	5										
Very unconfident	Unconfident	No opinion	Confident	Very confident																	
1	2	3	4	5																	
Opportunity costs <i>The benefits, profits or values that were given up to engage in the intervention</i>	Opportunity Costs: [Behaviour e.g. engaging in] [intervention] interfered with my other priorities <table border="1"> <tr> <td>Strongly disagree</td> <td>Disagree</td> <td>No opinion</td> <td>Agree</td> <td>Strongly agree</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> </table>	Strongly disagree	Disagree	No opinion	Agree	Strongly agree	1	2	3	4	5										
Strongly disagree	Disagree	No opinion	Agree	Strongly agree																	
1	2	3	4	5																	
General Acceptability	How acceptable was the [intervention] to you? <table border="1"> <tr> <td>Completely unacceptable</td> <td>Unacceptable</td> <td>No opinion</td> <td>Acceptable</td> <td>Completely acceptable</td> </tr> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> </table>	Completely unacceptable	Unacceptable	No opinion	Acceptable	Completely acceptable	1	2	3	4	5										
Completely unacceptable	Unacceptable	No opinion	Acceptable	Completely acceptable																	
1	2	3	4	5																	

Note: [intervention] should be replaced with the name of the healthcare intervention (e.g. Did you like or dislike the [feedback materials]?)

[Engaging with/ engaging with intervention] should be replaced with the specific behaviour participants are required to complete to engage with the intervention (e.g. how much effort did it to take you to [book your own appointment]? [clinical condition] should be replaced with the name of the clinical condition associated with the intervention (e.g. it is clear to me how [feedback reports] will result in improvements in [blood transfusion practice].

and retrospective acceptability of the COVID-19 vaccine. The questionnaire consisted of 9 items: one item reflecting each of the TFA constructs of affective attitude, burden, perceived effectiveness, intervention coherence, self-efficacy, and opportunity costs, two items reflecting ethicality, and one general acceptability item. Both ethicality items were included in the questionnaire, as we were not certain which of the two items would be most comprehensible and answerable.

Whilst the TFA identifies seven component constructs of acceptability, we propose a generic TFA questionnaire should include an item to assess overall acceptability, for two main reasons. First, by including a general acceptability item, researchers are able to explore which of the seven TFA constructs influences or drives participants' general acceptability judgment and secondly, it allows researchers to determine evaluations of overall acceptability of an intervention, which cannot be inferred from the sub-constructs alone (owing to uncertainties about the relative weightings of the items).

Think-aloud study

A 'think-aloud' study was conducted between March–April 2021 to explore people's views on the comprehensibility, relevance, and answerability of the items in two versions of questionnaires to assess the acceptability of a COVID-19 vaccine. Full ethical approval was obtained from King's College London Research Ethics Committee (REF:MRA-20/21–22,254).

Participants consisted of a convenience sample of 20 individuals recruited via Twitter. Ten individuals self-reported having received a first dose of the COVID-19 Vaccine (for the retrospective version of the acceptability questionnaire) and 10 self-reported an intention to have the COVID-19 Vaccine in the near future (for the prospective version of the acceptability questionnaire). Participants expressed their interest to take part in the study by contacting the primary researcher via e-mail. An information sheet and a consent form were then e-mailed to each participant. Participants signed the consent form and e-mailed back a scanned copy to the primary researcher.

Each participant completed the TFA questionnaire via a synchronous video call (Microsoft Teams), supported by a researcher (MS), who provided verbal instructions adapted from the think-aloud studies reported by French et al. (2007) [33] and Green and Gillhooly (1996) [34]. Participants were instructed to read each question, verbalise their thoughts (i.e., think-aloud) whilst completing the questionnaire, and to provide their response for each item. Participants were also instructed to state when a question item did not make sense, or if they were not sure what the item was asking. After the think-aloud

interview, the researcher asked participants to provide more details about any items identified as problematic and asked participants their opinions about the questionnaire in general (e.g., length; ease or difficulty of completing the questionnaire). The whole procedure took a maximum of half an hour per participant. Interviews were audio-recorded via the Microsoft Teams software, and transcribed verbatim.

Data were analysed by assigning each response to an adapted version of the categories applied by French et al. 2007 [33]:

- (1) No signification problems identified
- (2) Participant reread question, or seriously stumbled (i.e., stammered or stuttered because of misreading) in answering it (problems in understanding question),
- (3) Difficulty generating an answer
- (4) Questioned content of item (identified problems with how the question was worded, = did not understand the question), or
- (5) Answered a different question from the one that was asked or gave reasoning inconsistent with the answer given (problems in comprehending/answering question, misinterpretation of question).

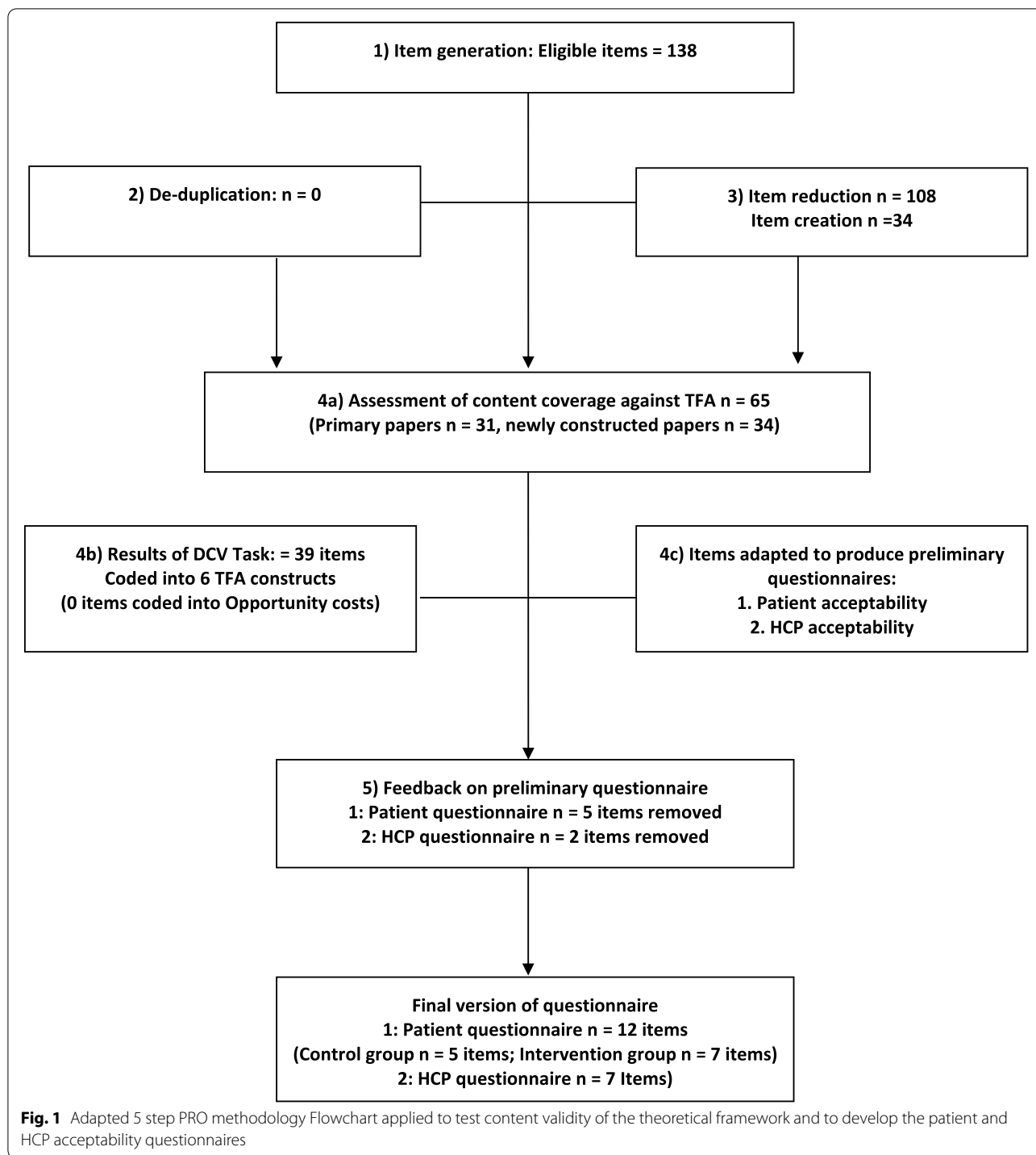
To assess the reliability of the researcher's (MS) coding, two additional researchers (JF and MC) each completed double coding on two transcripts (i.e., four transcripts in all). Agreement in coding between MS and MC, and between MS and JF was registered if the same part of a transcript was independently coded into the same category. Disagreement was registered where one researcher coded a section of text, but the other researcher did not, or else coded it into a different category. Percentage agreement rather than Cohen's Kappa was used to assess reliability because the items (i.e., sentences in transcripts) may have been coded into more than one of the categories [35, 36]. Any disagreements in coding were discussed and changes were agreed that would be applied to subsequent coding of the remaining transcripts.

Results

Figure 1 presents an overview of the adapted 5-Step PRO methodology applied to develop the two TFA questionnaires. Results of each of step are described in detail below.

Step 1: item generation

Twelve systematic reviews identified primary papers that had applied self-report assessment measures to investigate acceptability [7]. These included: three reviews assessing acceptability via measures of satisfaction [37,



38, 39]; four reviews describing reasons for discontinuation as indicators of acceptability [27, 40, 41, 42]; two reviews using participants’ perspectives and evaluations as assessments of intervention acceptability [29, 43]; two reviews synthesising evidence from open questions

to assess acceptability [44, 45], and one review assessing acceptability via participants’ attitudes [46].

Three hundred and forty-three primary papers included potential measures of acceptability. Of these, 325 (95%) papers were retrieved (18 were unavailable). Of the 325, 290 (89%) articles did not meet the inclusion

criteria for extraction of items. One hundred and thirty-eight items were extracted from 35 papers.

Step 2: De-duplication

There were no literal duplications of items, or differences in content (e.g. timeframe) or overlap with other generic items that were to be included in the questionnaires. Thus, no items were removed at this stage.

Step 3: item reduction and refinement of item wording

Two authors read each of the 138 items, applied the inclusion criteria (presented in [Methods](#) section above), and agreed on decisions to remove 107 items. Based on the construct definitions, one author drafted 34 new items, 17 items relating to the AFFINITIE trial [23, 24] and 17 items relating to the standard care and patient-initiated appointment booking services within the BEB/HFS trial [26].

Step 4: assessment of content coverage against a pre-existing theoretical framework

Eight participants completed the DCV task on 65 items (31 identified from the primary reviews and 34 new items). Within DCV tasks, content validity is usually tested using single sample t-tests (based on judges' rating of their confidence that the item represents a specified construct) [22, 30]. In the current study this would have required 455 (i.e., 65 (no. of items) x seven (number of construct definitions)) one-sample t-tests based on data from eight judges. The likely number of Type I errors was a substantive threat to validity, therefore null hypothesis significance testing was deemed inappropriate.

Instead, the analysis focused on descriptive statistics (means, standard deviations and medians, interquartile ranges) for each item. A median confidence rating across judges of +5 or greater (on the scale of -10 to +10) was considered an appropriate threshold to avoid lengthy questionnaires and taken as an indication that the judges agreed that an item closely reflected a construct.

Thirty-nine of the 65 items had a median confidence rating of 5 or greater and were coded into six of the TFA constructs, with no identified items for the construct, opportunity costs. Six items had a median confidence rating of 5 or greater for more than one construct, thus these items did not achieve discriminant validity [22].

The remaining 33 items (Supplementary file 3) were reviewed by all three authors for adaptation and inclusion in a preliminary version of the AFFINITIE acceptability questionnaire and BEB/HFS questionnaire (Supplementary files 4 and 5). A key decision included selecting only

one item per construct, and the item that could be best adapted for both preliminary questionnaires taking into account the response scale. The research team applied the following criteria when selecting items for both preliminary versions of the questionnaire:

- (i) Degree to which each item reflected the core definition of the construct;
- (ii) Degree to which the wording of each item was clear and unambiguous;
- (iii) Degree to which the wording of the item was appropriate for the intervention.

Disagreements were resolved by discussion. Thus, for each intervention-specific questionnaire eight items were included, one item per TFA construct and one general acceptability item. The rationale for including only a single item per construct was to keep the TFA questionnaires as brief as possible to enhance the questionnaires' usability, ease and speed of completion, and to make it more feasible to include an assessment of acceptability alongside other outcome measures.

Step 5: feedback on preliminary version of acceptability questionnaire from key stakeholders

AFFINITIE trial

The two principal investigators of the AFFINITIE programme advised that the TFA items would be more user-friendly and reduce participant burden for completion if the response anchors on the TFA items could be adapted to reflect the same 5-point Likert scales for the other items in the broader process evaluation questionnaire (i.e., strongly agree - strongly disagree).

Considering this feedback, the TFA response scale were changed to 5-point Likert scales, rather than the original scales. Supplementary file 4 displays the original TFA questionnaire (version 1) and the modified TFA questionnaire (version 2) applied in the AFFINITIE trial.

BEB and HFS trial

Feedback from the two patient representatives on the draft version of the BEB/HFS questionnaire resulted in the re-wording of three items to improve clarity. Both patient representatives suggested incorporating an option for additional comments for the intervention coherence item. Supplementary file 5 displays the final version of the control group and intervention group TFA-informed acceptability questionnaires for the BEB/HFS trial.

Generic TFA questionnaire

Table 1 displays a generic form of each of the items that can be adapted to assess intervention acceptability. There

Table 2 Frequency and type of problems identified for each of the prospective and retrospective TFA items adapted to assess acceptability of the COVID-19 vaccine

TFA Construct	Prospective/ retrospective	Questionnaire item	No problems	Re-read or stumbled in reading	Difficulty generating an answer	Questioned content of item	Misinterpreted item
			n/10 %	n/10 %	n/10 %	n/10 %	n/10 %
Affective attitude	Prospective	How comfortable do you feel to receive the COVID-19 Vaccine?	10 100%	0	0	0	0
	Retrospective	How comfortable did you feel to receive the COVID-19 Vaccine?	10 100%	0	0	0	0
Burden	Prospective	How much effort will it take to receive the COVID-19 Vaccine?	10 100%	0	0	0	0
	Retrospective	How much effort did it take to receive the COVID-19 Vaccine?	10 100%	0	0	0	0
Ethicality	Prospective	Receiving the COVID-19 Vaccine is fair to all:	9 90%	0	0	1 10%	0
	Retrospective	Receiving the COVID-19 Vaccine is fair to all:	5 50%	0	1 10%	5 50%	0
Ethicality	Prospective	There are moral or ethical consequences associated with receiving the COVID-19 vaccine:	7 50%	1 10%	0	2 20%	0
	Retrospective	There are moral or ethical consequences associated with receiving the COVID-19 vaccine:	9 90%	0	1 10%	0	0
Perceived effectiveness	Prospective	The COVID-19 Vaccine is likely to improve my chances of not becoming seriously ill:	10 100%	0	0	0	0
	Retrospective	The COVID-19 Vaccine has improved my chances of not becoming seriously ill :	10 90%	0	0	0	0
Intervention coherence	Prospective	It is clear to me how receiving the COVID-19 Vaccine will reduce my risk of becoming severely ill form Coronavirus	9 90%	1 10%	0	1 10%	0
	Retrospective	It is clear to me how receiving the COVID-19 Vaccine has reduced my risk of becoming severely ill from Coronavirus	9 90%	0	0	1 10%	0
Self-efficacy	Prospective	How confident do you feel that you can do what is required to receive the COVID-19 Vaccine?	6 60%	1 10%	0	2 20%	2 20%
	Retrospective	How confident did you feel that you could do what was required to receive the COVID-19 Vaccine?	8 80%	0	0	2 20%	0
Opportunity costs	Prospective	Receiving the COVID-19 Vaccine will interfere with my other priorities:	10 100%	0	0	0	0
	Retrospective	Receiving the COVID-11 Vaccine interfered with my other priorities:	10 100%	0	0	0	0
General acceptability	Prospective	How acceptable is the COVID-19 Vaccine to you?	10 100%	0	0	0	0
	Retrospective	How acceptable was the COVID-19 Vaccine to you?	10 100%	0	0	0	0

are two example items for the constructs of ethicality to reflect the differences in patient and healthcare professional interventions. Each item can be adapted to reflect a specific healthcare intervention by inserting a description of ‘the intervention’ (e.g. “feedback materials”) or to include the specific behaviour required to engage with the intervention (e.g. “booking own treatment appointment”). For some constructs, the item may include a reference to both the specific behaviour required to engage with the intervention and description of the intervention. For example, for the AFFINTIE trial, the generic burden item ‘How much effort did it take to [engage with intervention]?’ was adapted to ‘How much effort did it take to read the feedback materials?’, where ‘read’ is the behaviour and ‘feedback materials’ is the description of the intervention.

Think-aloud study

Table 2 presents an overview of the problems identified for each of the prospective and retrospective TFA items adapted to assess acceptability of the COVID-19 vaccine. Stemler (2004) [36] suggests that when using percentage

agreements to assess inter-rater reliability, values from 75 to 90% indicate an acceptable level of agreement. The agreement between the primary researcher (MS) and each of the additional researchers (MC, JF) was good (80–100%).

Supplementary file 6 provides illustrative examples of the problems that participants reported for the ethicality, intervention coherence and self-efficacy items. No participant identified problems with understanding and answering five out of nine items reflecting overall acceptability and four constructs: affective attitude, burden, perceived effectiveness, opportunity costs. Most problems that participants encountered included questioning the content of both prospective ethicality items (fairness item (n = 1); moral consequences item (n = 2)), retrospective ethicality (moral consequences item (n = 5)), both intervention coherence items (prospective (n = 1), retrospective (n = 2)) and self-efficacy items (prospective item (n = 2); retrospective item n = 2)). Participants re-read three items for the prospective questionnaire. This included the items for the constructs of ethicality (fairness) (n = 1), intervention coherence (n = 1) and self-efficacy (n = 1). Some participants also had difficulties

in generating an answer for the retrospective ethicality items (fairness $n = 1$), moral consequences $n = 1$). Misinterpretations occurred for the prospective self-efficacy item ($n = 2$).

Discussion

This study has described the development of a generic questionnaire, based on the TFA, for assessing the acceptability of healthcare interventions. We adapted the generic TFA-based questionnaire to assess prospective and retrospective acceptability of the COVID-19 vaccine in the UK context.

Our think-aloud findings indicate that the generic items assessing affective attitude, burden, perceived effectiveness and opportunity costs, adapted to assess acceptability of the COVID-19 vaccine, were comprehensible and answerable to all participants in our sample. Our think-aloud study, however, did identify some issues with the items representing the constructs ethicality and self-efficacy. This suggests that each of the ethicality items may not make sense in relation to some interventions. On reflection, the self-efficacy items may have been clearer if we had specified the behaviour more clearly (e.g. booking your COVID-19 vaccine appointment or attending your booked appointment). We advise researchers to pilot test their entire questionnaire with members of their intended population in relation to any specific intervention and adapt it, if participants find the items difficult to answer.

Strengths and limitations

The 5-step pre-validation method [20] provided a systematic approach in which both inductive (existing items from the overview of reviews) [7] and deductive methods (definitions of each of the seven TFA component constructs) were applied to develop two intervention-specific acceptability questionnaires. For each version of the preliminary questionnaire, the TFA was the basis for development of new items to reflect the construct, opportunity costs, as none of the items assessing this construct in the DCV exercise achieved discriminant content validity. Explicit TFA construct definitions were also important in re-wording existing items for inclusion in the acceptability questionnaires. Another strength of this study is the application of the DCV method to assess the content validity and discriminant validity of the identified existing and newly generated items across all seven component constructs in the TFA. As recommended, the DCV method [22] was completed in the early phase of developing the TFA-informed acceptability questionnaires. A further strength was the seeking of stakeholder

feedback for each of the two preliminary TFA questionnaires (one involving healthcare professionals and one involving patients), which resulted in minor amendments to the questionnaires.

The present study also demonstrates strengths in its comprehensive think-aloud study to assess the adaptation of both the prospective and retrospective acceptability items to assess acceptability of the COVID-19 vaccine. The study included an adequate sample size of 20 participants (10 participants per questionnaire), which allowed us to assess the face validity and answerability of the items and identify specific problems for each of the items.

There were several limitations. In this study, the large pool of items and the limited pool of judges meant it was not possible to complete the recommended statistical analysis for the DCV method [30] as multiple hypothesis testing would have generated too many false positives. However, the use of descriptive statistics and a threshold confidence rating to determine eligibility for inclusion in the questionnaire, identified an adequate number of items with good evidence of discriminant content validity. The pre-validation methodology recommends completing a think-aloud study with the target population on a newly developed questionnaire [20]. Whilst efforts were made to gain feedback from two stakeholders per preliminary questionnaire, it was not possible to complete think-aloud interviews with participants from both trial contexts that may have provided further information with regards to the comprehensibility, relevance and answerability of the draft questionnaires. Lastly, whilst efforts were made to recruit an adequate number of participants for the generic acceptability questionnaire think-aloud interviews, participants were recruited from a convenience sample, thus limiting the generalisability of the findings.

Recommendations for future work

Whilst systematic methods have been applied to develop the two TFA-based questionnaires, and the generic items have been applied to assess the acceptability of the COVID 19 vaccine, with a think aloud study completed, further work will be needed to establish further psychometric properties of both questionnaires. This is true of all pre-validation phases in developing new measures [20, 22]. We provide a supplementary file (Supplementary file 7) with guidance on how to adapt each of the items, and some notes on suggestions for analysing the TFA generic questionnaire.

Conclusion

This study has described the systematic methods applied to develop two intervention-specific acceptability questionnaires based on the recently developed Theoretical Framework of Acceptability (TFA), which have been adapted to develop a generic TFA-based questionnaire. We offer the generic TFA-based questionnaire as a brief and adaptable tool for researchers and clinicians to measure intervention acceptability across a range of healthcare interventions and to contribute to establishing an evidence base for psychometric properties of the items.

Abbreviations

TFA: Theoretical Framework of Acceptability; PRO: Patient Reported Outcomes; DCV: Discriminant Content Validation; BEB: Benign Essential Blepharospasm; HSF: Hemifacial Spasm; AFFINITIE: Audit and Feedback Interventions to Increase evidencebased Transfusion practice; COVID-19: Coronavirus disease.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12913-022-07577-3>.

Additional file 1.
Additional file 2.
Additional file 3.
Additional file 4.
Additional file 5.
Additional file 6.
Additional file 7.

Acknowledgements

Not applicable.

Authors' contributions

JF and MC conceived the study and supervised part of this work. MS completed the data collection and analysis for the preliminary acceptability questionnaires and wrote the main body of the manuscript. MS conducted the think-aloud interviews and analysis. MC completed reliability checks and contributed to the writing and editing of the manuscript. JF also completed reliability and contributed to the writing and editing of the manuscript. All authors contributed intellectually to the development of the generic theoretical framework of acceptability questionnaire. All authors read and approved the final version of the manuscript.

Funding

No grant funding was received to complete the work reported in this manuscript.

Availability of data and materials

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Think aloud study: This study was carried out in accordance with the Declaration of Helsinki, and full ethical approval was obtained from King's College London Research Ethics Committee (REF:MRA-20/21–22254). All protocols

were carried out following relevant guidelines and regulations. The research was classified as minimal risk. Participants expressed their interest to take part in the study by contacting the primary researcher via e-mail. An information sheet and a consent form were then e-mailed to each participant. Participants signed the consent form and e-mailed back a scanned copy to the primary researcher.

Consent for publication

'Not Applicable' in this section.

Competing interests

We declare that the authors have no competing interests as defined by BMC, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Author details

¹School of Health Sciences, City University of London, Northampton Square, London EC1V 0JB, UK. ²Department of Population Health Sciences, Faculty of Life Sciences and Medicine, Kings College London, London SE1 1UL, UK. ³Melbourne School of Health Sciences, University of Melbourne, Melbourne, VIC 3010, Australia. ⁴Centre of Implementation Research, Ottawa Hospital Research Institute – General Campus, 501 Smyth Road, Ottawa, ON K1H 8L6, Canada.

Received: 27 July 2021 Accepted: 1 February 2022

Published online: 01 March 2022

References

- Haynes B. Can it work? Does it work? Is it worth it?: the testing of health-care interventions is evolving. *BMJ. Br Med J.* 1999;319:652.
- Say RE, Thomson R. The importance of patient preferences in treatment decisions—challenges for doctors. *Bmj.* 2003;327:542–5.
- Sidani SaB CJ. Testing the acceptability and feasibility of interventions. Design, evaluation, and translation of nursing interventions. West Sussex: Wiley; 2011.
- Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *Bmj.* 2008;337:a1655.
- Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *Bmj.* 2015;350:h1258.
- Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *Bmj.* 2021;30:374.
- Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. *BMC Health Serv Res.* 2017;17:88.
- Arrowsmith ME, Aicken C, Saxena S, Majeed A. Strategies for improving the acceptability and acceptance of the copper intrauterine device. *Cochrane Database Syst Rev.* 2012;3.
- Berlim MT, McGirr A, Van den Eynde F, Fleck MP, Giacobbe P. Effectiveness and acceptability of deep brain stimulation (DBS) of the subgenual cingulate cortex for treatment-resistant depression: a systematic review and exploratory meta-analysis. *J Affect Disord.* 2014;159:31–8.
- Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet.* 2009;373:746–58.
- Gillies K, Bower P, Elliott J, MacLennan G, Newlands RS, Ogden M, et al. Systematic techniques to enhance rEtenion in randomised controlled trials: the STEER study protocol. *Trials.* 2018;19:197.
- Hughes-Morley A, Young B, Hempel RJ, Russell IT, Waheed W, Bower P. What can we learn from trial decliners about improving recruitment? Qualitative study. *Trials.* 2016;17:494.
- Koller I, Levenson MR, Glück J. What do you think you are measuring? A mixed-methods procedure for assessing the content validity of test items and theory-based scaling. *Front Psychol.* 2017;8:126.

14. Sekhon M, Cartwright M, Francis JJ. Acceptability of health care interventions: A theoretical framework and proposed research agenda. *Br J Health Psychol.* 2018;23:519-31. <https://doi.org/10.1111/bjhp.12295>.
15. Wiering B, de Boer D, Delnoij D. Patient involvement in the development of patient-reported outcome measures: a scoping review. *Health Expect.* 2017;20(1):11–23.
16. Black N. Patient reported outcome measures could help transform health care. *BMJ.* 2013;346:f167.
17. Renko E, Knittle K, Palsola M, Lintunen T, Hankonen N. Acceptability, reach and implementation of a training to enhance teachers' skills in physical activity promotion. *BMC Public Health.* 2020;20(1):1–3.
18. Keyworth C, O'Connor R, Quinlivan L, Armitage CJ. Acceptability of a brief web-based theory-based intervention to prevent and reduce self-harm: mixed methods evaluation. *J Med Internet Res.* 2021;23(9):e28349.
19. Koller I, Levenson MR, Glück J. What Do You Think You Are Measuring? A Mixed-Methods.
20. Prior ME, Hamzah JC, Francis JJ, Ramsay CR, Castillo MM, Campbell SE, et al. Pre-validation methods for developing a patient reported outcome instrument. *BMC Med Res Methodol.* 2011;11:112.
21. Turner RR, Quittner AL, Parasuraman BM, Kallich JD, Cleeland CS, Group MFPROC. Patient reported outcomes: instrument development and selection issues. *Value Health.* 2007;10:S86.
22. Johnston M, Dixon D, Hart J, Glidewell L, Schröder C, Pollard B. Discriminant content validity: a quantitative methodology for assessing content of theory-based measures, with illustrative applications. *Br J Health Psychol.* 2014;19:240–57.
23. Gould NJ, Lorencatto F, Stanworth SJ, Michie S, Prior ME, Glidewell L, et al. Application of theory to enhance audit and feedback interventions to increase the uptake of evidence-based transfusion practice: an intervention development protocol. *Implement Sci.* 2014;9:92.
24. Hartley S, Foy R, Walwyn RE, Cicero R, Farrin AJ, Francis JJ, et al. The evaluation of enhanced feedback interventions to reduce unnecessary blood transfusions (AFFINITIE): protocol for two linked cluster randomised factorial controlled trials. *Implement Sci.* 2017;12:84.
25. Lorencatto F, Gould NJ, McIntyre SA, During C, Bird J, Walwyn R, et al. A multidimensional approach to assessing intervention fidelity in a process evaluation of audit and feedback interventions to reduce unnecessary blood transfusions: a study protocol. *Implement Sci.* 2016;11:163.
26. Wickwar S, McBain H, Newman SP, Hirani SP, Hurt C, Dunlop N, et al. Effectiveness and cost effectiveness of a patient-initiated botulinum toxin treatment model for blepharospasm and hemifacial spasm compared to standard care: study protocol for a randomised controlled trial. *Trials.* 2016;17:129.
27. Kulier R, Helmerhorst FM, Maitra N, Gülmezoglu AM. Effectiveness and acceptability of progestogens in combined oral contraceptives—a systematic review. *Reproductive health.* 2004;1:1.
28. Caldeira D, Goncalves N, Ferreira JJ, Pinto FJ, Costa J. Tolerability and acceptability of non-vitamin K antagonist oral anticoagulants in atrial fibrillation: systematic review and meta-analysis. *Am J Cardiovascular Drugs.* 2015;15(4):259-65.
29. Glenton C, Khanna R, Morgan C, Nilsen ES. The effects, safety and acceptability of compact, pre-filled, auto-disable injection devices when delivered by lay health workers. *Tropical Med Int Health.* 2013;18:1002-16.
30. Dixon D, Johnston M, McQueen M. The Disabilities of the Arm, Shoulder and Hand Questionnaire (DASH) can measure the impairment, activity limitations and participation restriction constructs from the International Classification of Functioning, Disability and Health (ICF). *BMC musculo-skeletal disorders.* 2008;9:114.
31. Lynn MR. Determination and quantification of content validity. *Nursing research.* 1986.
32. Waltz CF, Bausell RB. *Nursing research: Design, statistics, and computer analysis.* FA Davis company; 1981.
33. French DP, Cooke R, Mclean N, Williams M, Sutton S. What do people think about when they answer Theory of Planned Behaviour questionnaires? A think aloud study. *Journal of health psychology.* 2007;12(4):672-87.
34. Green C, Gilhooly K. Protocol analysis: Practical implementation. *Handbook of qualitative research methods for psychology and the social sciences.* 1996:55-74.
35. Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin.* 1968;70(4):213.
36. Stemler SE. A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research, and Evaluation.* 2004;9(1):4.
37. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS one.* 2010;5:e13196.
38. Blenkinsopp A, Hassey A. Effectiveness and acceptability of community pharmacy-based interventions in type 2 diabetes: a critical review of intervention design, pharmacist and patient perspectives. *International Journal of Pharmacy Practice.* 2005;13:231–40.
39. Kedge EM. A systematic review to investigate the effectiveness and acceptability of interventions for moist desquamation in radiotherapy patients. *Radiography.* 2009;15:247–57.
40. Halpern V, Lopez LM, Grimes DA, Stockton LL, Gallo MF. Strategies to improve adherence and acceptability of hormonal methods of contraception. *Cochrane database of systematic reviews.* 2013(10).
41. Kaltenthaler E, Sutcliffe P, Parry G, Beverley C, Rees A, Ferriter M. The acceptability to patients of computerized cognitive behaviour therapy for depression: a systematic review. *Psychological Med.* 2008;38:1521-30.
42. Koesters M, Guaiana G, Cipriani A, Becker T, Barbui C. Agomelatine efficacy and acceptability revisited: systematic review and meta-analysis of published and unpublished randomised trials. *The British Journal of Psychiatry.* 2013;203:179-87.
43. Muftin Z, Thompson AR. A systematic review of self-help for disfigurement: Effectiveness, usability, and acceptability. *Body Image.* 2013;10:442-50.
44. Robinson L, Hutchings D, Dickinson H, Corner L, Beyer F, Finch T, et al. Effectiveness and acceptability of nonpharmacological interventions to reduce wandering in dementia: a systematic review. *International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences.* 2007;22:9-22.
45. Newman PA, Logie C. HIV vaccine acceptability: a systematic review and meta-analysis. *Aids.* 2010;24:1749–56.
46. Diepeveen S, Ling T, Suhrcke M, Roland M, Marteau TM. Public acceptability of government intervention to change health-related behaviours: a systematic review and narrative synthesis. *BMC Public Health.* 2013;13:756.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

