# Protocol Template for the Evaluation of Domain Match and Feasibility of XXXX for Use in Core Outcome Sets for XXXX

## Background

Selecting an appropriate outcome measurement instrument (hereafter referred to as "instrument") to represent a target domain requires evidence that it accurately measures that domain. This typically evaluating its measurement properties. However, before conducting such a review, some fundamental characteristics of the instrument must be assessed.

These include the first two questions outlines in the OMERACT Filter Instrument Selection Algorithm (OFISA) 1.

1. **Domain Match** – Is it a match with the target domain? This assessment considers the relevance and comprehensiveness of the instrument’s content and whether it fully captures the intended domain. The detailed domain definition endorsed prior to instrument selection serves as the "gold standard" for this evaluation 2,3.
2. **Feasibility -** Is it practical to use? Feasibility considerations include cost, accessibility, respondent and administrative burden, and any necessary training for implementation.

This protocol outlines the OMERACT approach to assessing domain match and feasibility, building on the work of Ken Tang, who synthesized frameworks for sensibility and concept match from the literature. The key articles that informed the OMERACT methodology are as follows:

|  |
| --- |
| Article Citation |
| Pakulis, P. Janine, et al. “Evaluating Physical Function in an Adolescent Bone Tumor Population.” Pediatric Blood & Cancer, vol. 45, no. 5, 2005, pp. 635–43, https://doi.org/10.1002/pbc.20383. |
| Smith ML, Sosa ET, Tisone CA, et al. Quality Enhancement Groups: A Qualitative Research Method for Survey Instrument Development. Vol 1.; 2010. https://www.researchgate.net/publication/266573781 |
| \* Wells, George, et al. “Updating the Omeract Filter: Discrimination and Feasibility.” Journal of Rheumatology, vol. 41, no. 5, 2014, pp. 1005–10, https://doi.org/10.3899/jrheum.131311. |
| Auger, Claudine, et al. “Making Sense of Pragmatic Criteria for the Selection of Geriatric Rehabilitation Measurement Tools.” Archives of Gerontology and Geriatrics, vol. 43, no. 1, 2006, pp. 65–83, https://doi.org/10.1016/j.archger.2005.09.004. |
| Terwee, Caroline B., et al. “Qualitative Attributes and Measurement Properties of Physical Activity Questionnaires: A Checklist.” Sports Medicine (Auckland), vol. 40, no. 7, 2010, pp. 525–37, https://doi.org/10.2165/11531370-000000000-00000. |
| Lawshe, C. H. “A Quantitative Approach to Content Validity.” Personnel Psychology, vol. 28, no. 4, 1975, pp. 563–75, https://doi.org/10.1111/j.1744-6570.1975.tb01393.x. |
| Rowe, Brian H., and Andrew D. Oxman. “An Assessment of the Sensibility of a Quality-of-Life Instrument.” The American Journal of Emergency Medicine, vol. 11, no. 4, 1993, pp. 374–80, https://doi.org/10.1016/0735-6757(93)90171-7. |
| Feinstein AR. The theory and evaluation of sensibility. In Feinstein AR Clinimetrics. Westford MA: Murray Printing Co. 1987:141-166. |
| \* Tang, Kenneth, et al. “Sensibility of Five At-Work Productivity Measures Was Endorsed by Patients with Osteoarthritis or Rheumatoid Arthritis.” Journal of Clinical Epidemiology, vol. 66, no. 5, 2013, pp. 546–56, https://doi.org/10.1016/j.jclinepi.2012.12.009. |
| Law, Mary. “Measurement in Occupational Therapy: Scientific Criteria for Evaluation.” Canadian Journal of Occupational Therapy (1939), vol. 54, no. 3, 1987, pp. 133–38, https://doi.org/10.1177/000841748705400308. |
| Nunnally, Jum C., and Ira H. Bernstein. Psychometric Theory. 3rd ed., McGraw-Hill, 1994. |

**Table 1** Literature used to formulate framework for the evaluation of domain match and feasibility. \*Key articles reflecting review and synthesis of the literature at OMERACT.

This study will assess whether the [NAME OF INSTRUMENT AND VERSION] meets the OMERACT Filter 4 criteria for domain match and feasibility.

## Research Questions

1. How well does [INSTRUMENT NAME AND VERSION] capture the salient issues of [DOMAIN] in patients with [DISEASE].
2. How feasible is it to use [INSTRUMENT NAME AND VERSION] in clinical trials of persons with [DISEASE] including patient research partners (PRPs) and other collaborators involved in the selection and use of outcome measures for these disorders?

## Objectives

This study aims to:

1. Assess the comprehensiveness, comprehensibility and relevance of the items in [INSTRUMENT NAME AND VERSION]
2. Evaluate item-level performance of [INSTRUMENT NAME AND VERSION] based on the responses of the relevant PRPs with [DISEASE] including response distributions, missing data patterns, item-total correlations, and, if applicable, Cronbach’s alpha coefficient.
3. Examine the feasibility of using [INSTRUMENT NAME AND VERSION] by analyzing completion rates, time required for completion, and potential language or cultural differences affecting accurate and complete responses (e.g., missing data, respondent comments on item relevance).

## Design

This study will use a cross-sectional survey approach, involving two participant groups:

1. Patient Research Partners (PRPs)
2. Other Collaborators, including clinicians, researchers, industry representatives, payors, and policymakers from the OMERACT community.

## Participants

A total of at least 60 participants will be recruited, including:

* PRPs: A minimum of 30 individuals with lived experience of [DISEASE]. Participants will have undergone training on OMERACT methodologies and the role of PRP input in instrument selection, using online educational materials. Their evaluation will focus on domain match, content validity, and feasibility. Efforts will be made to recruit PRP’s from diverse geographical regions, ideally spanning at least three continents.
* All Other Collaborators: A minimum of 30 individuals from the OMERACT community, including clinicians, researchers, industry representatives, payors, and policymakers. These participants will have experience with the disease of interest and understand the OMERACT core outcome set development process.

**Inclusion criteria:**

PRPs

1. Aged 18 or older
2. Has lived experience with [DISEASE]
3. Has completed or is willing to complete the e-learning module on Instrument Selection
4. Provides consent to participate in the survey

All Other collaborators.

1. Aged 18 or older
2. Engaged in OMERACT core outcome set work as a researcher, payor, provider, policymaker, or industry representative
3. Has completed or is willing to complete the e-learning module on Instrument Selection
4. Provides informed consent to participate in the survey

**Exclusion criteria (Either group):**

1. Unwilling or unable to complete the survey in the available language(s) and administration modes.

## Sample size justification

There are no standardized calculations for determining sample size in studies evaluating domain match, or feasibility. However, a minimum of 30 participants per group is considered sufficient to capture diverse perspectives, assess the breadth and depth of the instrument’s content, and examine response patterns. This sample size allows for meaningful qualitative and quantitative feedback while ensuring adequate representation of key collaborator groups.

## Recruitment

PRPs and other collaborators will be recruited using approved electronic and paper-based postcards that include a URL link and a matching QR code. These materials will provide potential participants with study details, emphasizing its anonymous and voluntary nature.

To facilitate participation, individuals will complete a brief eligibility screening before accessing the survey. If they meet the inclusion criteria, they will proceed to complete the survey.

The URL/QR Code link will be distributed via email to PRPs and other collaborators through the OMERACT mailing list, which is regularly used for communication by the central office, committees, and working groups. Recipients will be responsible for initiating their participation by using the provided link.

Study personnel contact information will be included in all communications, ensuring that participants can reach out via email or phone if they have any questions.

## Consent

Upon accessing the survey link, invitees will be directed to a study description page outlining the purpose, procedures, and voluntary nature of participation.

Participants will be informed that:

* No identifying information will be collected.
* They may skip any questions they do not wish to answer.
* Their responses will be used solely for the purposes described in the study.

By proceeding beyond this page and entering the survey, participants will provide implied consent to take part in the study.

## Confidentiality

All survey responses will be anonymous, identified only by a study-specific ID number. No personal identifying information will be collected, and survey results will be downloaded without any links to IP addresses or email addresses.

Data will be reported in aggregate at the group level. To further protect participant anonymity, any response categories with fewer than five individuals will either be combined with another category or suppressed from reporting.

## Survey Content

The survey will be administered online via [PLATFORM NAME] via [SERVERS NAME] servers, with two distinct versions:

1. One for PRPs
2. One for All Other Collaborators (clinicians, researchers, industry representatives, payors, and policymakers)

As outlined in Table 2, some survey questions will be specific to PRPs.

The survey will remain open for two weeks. Participants will receive:

* A reminder email at the end of week one, including a survey link, closing date reminder, and a thank-you note for those who have already responded.
* A final reminder on the day before the survey closes 5.

**Survey Components**

All participants will:

1. Complete [INSTRUMENT NAME AND VERSION], the instrument under evaluation.
2. Reflect on its alignment with the target domain and its practicality of use, including:

* Ease of understanding
* Accuracy in capturing the impact of their experience (instrument as a whole and individual items)
* Perceived length and difficulty, and whether content covered important issues for them

1. Provide demographic data, including age, gender, and country of residence.
2. PRPs will additionally answer questions related to their lived experience with arthritis.

**Assessment of Construct Validity**

To evaluate construct validity, additional items will be included to examine expected correlations between the target instrument and related constructs. Anticipated relationships will be pre-specified before survey administration.

| **Section** | **Content** | **PRPs** | **Other OMERACT collaborators** |
| --- | --- | --- | --- |
| Background information | Brief Description of survey including domain definition | **X** | **X** |
| Instrument of interest | XXXX | **X** | **Answer thinking of patients with disease** |
| Opinions about instrument of interest (following each Instrument of interest) | Difficulties/opinions for each of the standardized questionnaires overall | **X** | **X** |
| Content validity at the items level: relevance of each item to the experience (Content validity index) | **X** | **X** |
| Open ended response: Considering the domain of interest, is there anything missing from this questionnaire? | **X** | **X** |
| Thinking of the changes you\* experience with your arthritis, do you think this set of questions would be able to capture that change? (\*modified to people with disease for other collaborators) | **X** |  |
| Feasibility questions: see appendix for sample questions about the feasibility of using this instrument (time to complete, complexity, wording and instructions clear etc.) | **X** | **X** |
| Demographic  Information | Age range | **X** | **X** |
| Gender | **X** | **X** |
| Language spoken at home | **X** | **X** |
| Group you are representing in your survey responses.  Patient, clinician, researcher other. | **X** | **X** |
| Country of residence | **X** | **X** |
| Questions related to current arthritis experience for those with the disorder. | Rating of current state of condition. | **X** |  |
| Duration of your arthritis/musculoskeletal disorder. | **X** |  |
| Patient global health assessment (16, 17) | **X** |  |
| Use of supportive devices (18) | **X** |  |
| Comorbidity (20) | **X** |  |

**Table 2.** Survey content.

## Data Management

All data will be collected through a secure online survey platform. Upon survey completion, data will be reviewed to ensure no retention of identifying information, such as IP addresses, before being downloaded in CSV format. Data files will be securely stored on an institutional server, with access restricted to the study team to maintain confidentiality and data integrity.

## Analytic Plan

Data analysis will be completed using Excel and statistical software (SAS, R).

**Description of the sample and the scales**

Descriptive statistics will be used to describe the sample demographics (age, sex, etc.) for both groups: PRPs and all other collaborators. Distribution of the instrument as completed by the PRP group will be used to describe the range of scores.

**Item level analysis**

The instrument will be analyzed at the item level (i.e., response distributions, missing items, item total correlations, Cronbach’s alpha coefficient) using the PRP responses only (See table XX as a template for this analysis). Relevance of the item will also be rated by calculating the average relevance rating for each item, and across the items. An overall relevance across items exceeding 0.70 will suggest relevant content in this set of items 6.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Response: Amount of difficulty doing item | | | |  |  |
| Item # | Item label | Missing (n) | 0 | 1 | 2 | 3 | Mean (0-3) item score | Item to total correlation |
|  |  |  | No difficulty | Some difficulty | A lot of difficulty | Not able to do |  |  |
| 1 | Reach | 1 | 148 | 83 | 17 | 1 | 0.48 | 0.43 |
| 2 | Sit | 0 | 127 | 105 | 16 | 1 | 0.56 | 0.38 |
| 3 | Lift | 1 | 117 | 101 | 28 | 2 | 0.66 | 0.49 |
| 4 | work | 3 | 75 | 105 | 57 | 9 | 1.00 | 0.33 |
| 5 | Carry | 2 | 49 | 121 | 70 | 7 | 1.14 | 0.51 |
| 6 | Pain | 3 | 102 | 111 | 33 | 0 | 0.72 | 0.53 |
| 7 | Dress | 4 | 37 | 94 | 80 | 34 | 1.45 | 0.46 |
| 8 | Transport | 2 | 127 | 93 | 26 | 1 | 0.60 | 0.43 |
| 9 | Walk | 1 | 175 | 63 | 10 | 0 | 0.33 | 0.67 |
| 10 | Run | 5 | 153 | 78 | 13 | 0 | 0.43 | 0.70 |
| 11 | Sports | 2 | 130 | 96 | 21 | 0 | 0.56 | 0.52 |
| Instrument: XXXXX  Mean Score:\_\_\_\_\_\_\_\_\_\_\_\_\_\_ SD\_\_\_\_\_\_\_\_\_\_\_\_\_ Median\_\_\_\_\_\_\_\_\_\_\_  Cronbach’s Alpha coefficient in this data:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | | | | | | |  |

**Table 3.** Sample table summarizing item level response frequencies and score summaries for a fictitious instrument.

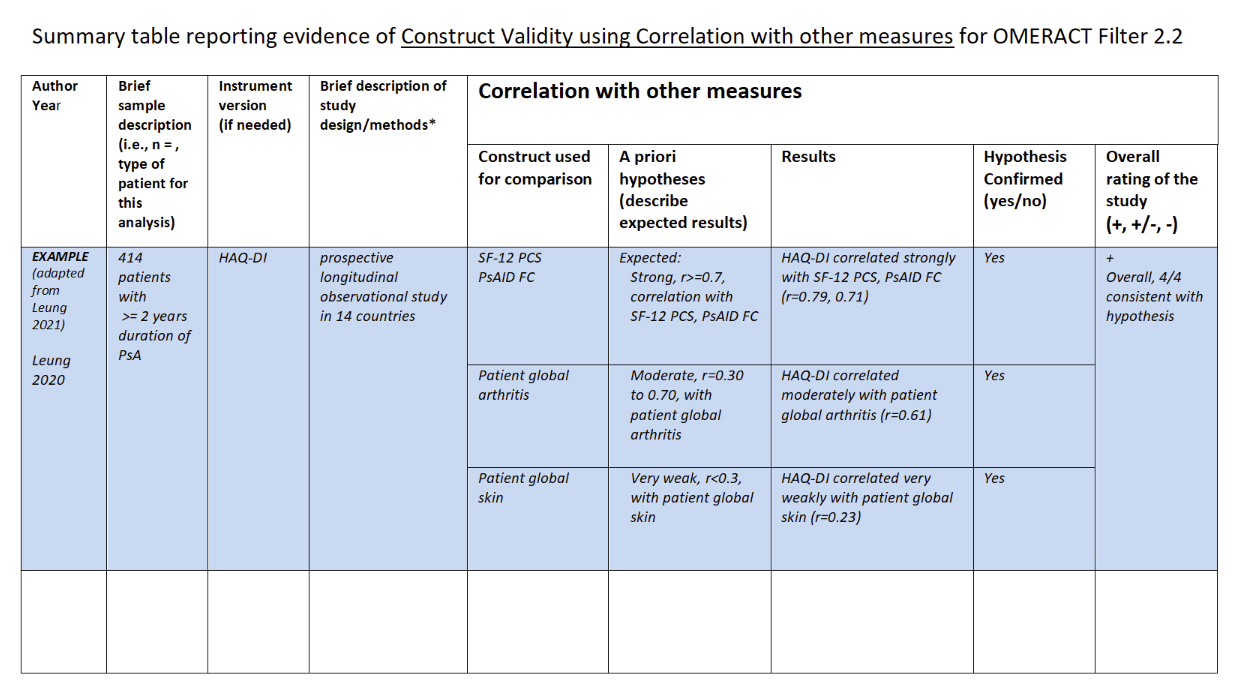
**Score Analysis**

**Score Calculation & Distribution**

* Overall scores will be computed according to the instrument developers’ guidelines.
* Score distribution will be examined for validity and representativeness.
* Floor and ceiling effects will be assessed, with a threshold set at ≤15%.
* Complete collection rates of ≥80%.

**Construct Validity**

* Construct validity will be tested comparing instrument scores with constructs fielded in the survey for comparison.
* A priori hypotheses for each comparison, including expected direction and magnitude, will be documented in the OMERACT construct validity narrative table (sample shown below).
* Concurrent comparison of two instruments will be carried out using the Spearman rank correlation coefficient, which can be used with ordinal data.
* Known group comparisons will be tested with an unpaired analysis of variance with statistical significance set to p < 0.05.



**Figure 1:** Sample summary table reporting evidence of construct validity using correlation with other measures for OMERACT Filter 2.2

**Feasibility**

Feasibility data will be analyzed descriptively, with responses categorized as Yes or No for appraisal. An agreement threshold of 70% will be used to determine endorsement of an item by the working group.

* **Agreement Criteria:**
  + An item is considered endorsed when both subgroups (PRPs and Other Collaborators) report at least 70% agreement.
  + If the difference in endorsement rates between the two subgroups is ≤10%, and the combined average remains ≥70%, the item will be considered agreed upon.
  + If the difference between subgroups exceeds 10%, the item will be classified as having dissensus.
* **Data Presentation:**
  + Results will be reported separately for PRPs and All Other Collaborators, as well as in a combined dataset.
  + Uncertain responses and missing data will be summarized for each item.

The full list of feasibility assessment questions can be found in the Appendix of this protocol.

## Final decision about the instrument.

The results of the domain match and feasibility assessment will be presented to the working group for evaluation. Based on the findings, the group will determine whether the instrument is suitable for further consideration using a traffic light rating system:

A close-up of a traffic light

AI-generated content may be incorrect.

**Figure 2.** Meaning of the traffic light scoring in the instrument selection part A domain match & feasibility

Only instruments receiving a Green or Amber rating will advance to a full evaluation of measurement properties.

# References

1. Maxwell LJ, Beaton DE, Boers M, et al. The evolution of instrument selection for inclusion in core outcome sets at OMERACT: Filter 2.2. Semin Arthritis Rheum. 2021;51(6):1320-1330. doi:10.1016/j.semarthrit.2021.08.011

2. D’Agostino MA, Beaton DE, Maxwell LJ, et al. Improving domain definition and outcome instrument selection: Lessons learned for OMERACT from imaging. Semin Arthritis Rheum. 2021;51(5):1125-1133. doi:10.1016/j.semarthrit.2021.08.004

3. Maxwell LJ, Jones C, Bingham CO, et al. Defining domains: developing consensus-based definitions for foundational domains in OMERACT core outcome sets. Semin Arthritis Rheum. 2024;66. doi:10.1016/j.semarthrit.2024.152423

4. Boers M, Beaton DE, Shea BJ, et al. OMERACT filter 2.1: Elaboration of the conceptual framework for outcome measurement in health intervention studies. Journal of rheumatology. 2019;46(8):1021-1027. doi:10.3899/jrheum.181096

5. Faubion CW, Andrew JD. Book Review: Dillman, D. A. (2000). Mail and Internet Surveys: The Tailored Design Method (2nd ed.). New York: Wiley 464 pp., $47.50 (hardcover). Rehabil Couns Bull. 2001;44(3):178-180. doi:10.1177/003435520104400309

6. LAWSHE CH. A QUANTITATIVE APPROACH TO CONTENT VALIDITY 1. Pers Psychol. 1975;28(4):563-575. doi:10.1111/j.1744-6570.1975.tb01393.x

# Appendix. A - Survey Of PRPs With Lived Experience On Instrument Domain Match & Feasibility

Instrument Name:

|  |  |
| --- | --- |
| **Match to Domain:** Thinking about the content of the actual questions/items in the instrument, based on experience of this domain. | **Respondents Answer** |
| Are the items in this instrument relevant to you and your experience?  **Comments:** | Yes  Uncertain  No |
| Do you think there should be any additional items (i.e., were there things that were missed)?  **Comments:** | Yes  Uncertain  No |
| Do you think that there should be any items taken out of the instrument?  If yes, tell us why.  **Comments:** | Yes  Uncertain  No |
| Were there overlapping, sensitive, or embarrassing items?  **Comments:** | Yes  Uncertain  No |
| Does the instrument overall reflect your experience of your [domain]?  **Comments:** | Yes  Uncertain  No |
| Did you find that all the items were easy to read? If not, which items were not easy to read?  **Comments:** | Yes  Uncertain  No |
| Did you feel that all the items were clear and understandable? Could you understand what all the questions were trying to ask? If not, which items did you feel were unclear?  **Comments:** | Yes  Uncertain  No |
| Did you think that the response options were clear and understandable (i.e. did the possible answers match well with the questions)? If not, which items did you feel had a mismatched response scale?)  **Comments:** | Yes  Uncertain  No |
| Were the instructions for answering the items clear?  **Comments:** | Yes  Uncertain  No |
| Does the timing of the recall period seem reasonable to you (e.g. over the past week, last 24 hours) (if applicable)?  **Comments:** | Yes  Uncertain  No  Not applicable |

|  |  |
| --- | --- |
| **Feasibility:** Questions about the practical considerations about this instrument. | **Respondents Answer** |
| Was it easy enough to complete?  **Comments:** | Yes  Uncertain  No |
| Did it take a reasonable amount of time to complete?  **Comments:** | Yes  Uncertain  No |
| Did the format seem appropriate (how it looked on the page, font size, how items and responses were organized)?  **Comments:** | Yes  Uncertain  No |
| Do you think there was too much equipment and training needed before you could be able to respond to this instrument?  **Comments:** | Yes  Uncertain  No |
| Is it easy for respondents to understand (considering reading level, instructions, health, and literacy needed)?  **Comments:** | Yes  Uncertain  No |
| Can it be completed within a reasonable amount of time given your study context?  Comments: | Yes  Uncertain  No |
| Is the method of administration feasible for your application (i.e., computer-based, paper, equipment needs)?  **Comments:** | Yes  Uncertain  No |
| Are the costs feasible? (consider licensing fees, equipment and training costs).  **Comments:** | Yes  Uncertain  No |
| Are the copyright issues (if any) reasonable and manageable?  **Comments:** | Yes  Uncertain  No |
| Are the equipment, space and training needs feasible for you to carry out?  **Comments:** | Yes  Uncertain  No |
| Is it available in the right language/culture for your intended application?  **Comments:** | Yes  Uncertain  No |

# Appendix. B - Working Group Assessment Survey On Instrument Domain Match

INSTRUMENT NAME:

|  |  |
| --- | --- |
| Is this instrument (think about items, response, domain capture for PROs; for imaging, think about match with domain components) measuring what YOU want to measure? Are the items relevant to your concept, as experienced by your targeted patients’ experiences and for the intended application? Consider sources of variability you identified, are any of those criteria that were considered in the definition of the domain? For example, is using assistive devices permitted in your concept of independence in ADL functioning? Or do you want to specify a particular time of day when you define your concept of pain intensity – night pain, or morning pain for example?  **Comments:** | Yes  Uncertain  No |
| Have all important the elements of the target domain for this population, and intended application been included (consider breadth and depth needed)?  **Comments:** | Yes  Uncertain  No |
| Is the instrument free of redundant, unnecessary, or potentially inappropriate or sensitive items?  **Comments:** | Yes  Uncertain  No |
| Are the items phrased in a clear and understandable way?  **Comments:** | Yes  Uncertain  No |
| Are the items written at a level that will be understood by the target population?  Comments: | Yes  Uncertain  No |
| Are the instructions for completing items and selecting responses for the items clear?  **Comments:** | Yes  Uncertain  No |
| Are the response options clear and appropriate for each item (consider match with the question, ordering of responses)?  **Comments:** | Yes  Uncertain  No |
| Is the recall period in the instrument appropriate given the population, domain and intended application, i.e, over the past week, last 24 hours (if applicable)?  **Comments:** | Yes  Uncertain  No |
| Is the method of scoring appropriate (consider any weighted responses)?  **Comments:** | Yes  Uncertain  No |

# Appendix. C - Working Group Survey On Instrument Feasibility

Instrument Name::

|  |  |
| --- | --- |
| **Feasibility:** Questions about the practical considerations about this instrument. | **Respondents Answer**  **Options** |
| Is it easy for respondents to understand (considering reading level, instructions, health, and literacy needed)?  **Comments:** | Yes  Uncertain  No |
| Can it be completed within a reasonable amount of time given your study context?  **Comments:** | Yes  Uncertain  No |
| Is the method of administration feasible for your application (i.e., computer-based, paper, equipment needs)?  **Comments:** | Yes  Uncertain  No |
| Are the costs feasible? (consider licensing fees, equipment and training costs).  **Comments:** | Yes  Uncertain  No |
| Are the copyright issues (if any) reasonable and manageable?  **Comments:** | Yes  Uncertain  No |
| Are the equipment, space and training needs feasible for you to carry out?  **Comments:** | Yes  Uncertain  No |
| Is it available in the right language/culture for your intended application?  **Comments:** | Yes  Uncertain  No |