Manchester Metropolitan University

e-space

Manchester Metropolitan University's Research Repository

Please cite the Published Version

Webb, EJD, Meads, D, Lynch, Y, Judge, S, Randall, N, Goldbart, J, Meredith, S, Moulam, L, Hess, S and Murray, J (2021) Attribute Selection for a Discrete Choice Experiment Incorporating a Best-Worst Scaling Survey. Value in Health, 24 (4). pp. 575-584. ISSN 1098-3015

DOI: https://doi.org/10.1016/j.jval.2020.10.025

Publisher: Elsevier

Version: Accepted Version

Downloaded from: https://e-space.mmu.ac.uk/628858/

Additional Information: This is an Author Accepted Manuscript of an article published in Value in Health by Elsevier.

Enquiries:
If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines)

Attribute Selection for a Discrete Choice Experiment Incorporating a Best-Worst Scaling Survey

Edward J.D. Webb, PhD, David Meads, PhD, Yvonne Lynch, PhD, Simon Judge, MEng, Nicola Randall, PhD, Juliet Goldbart, PhD, Stuart Meredith, Liz Moulam, PGDip, Stephane Hess, PhD, Janice Murray, phD

ABSTRACT

Objectives: Although literature exists on using qualitative methods to generate potential attributes for a discrete choice experiment (DCE), there is little on selecting which attributes to include. We present a case study in which a best-worst scaling case 1 (BWS-1) survey was used to guide attribute selection for a DCE. The case study's context was the decision making of professionals around the choice of augmentative and alternative communication (AAC) systems for children with limited natural speech.

Methods: BWS-1 survey attributes were generated from literature reviews and focus groups. DCE attributes were selected from BWS-1 attributes. The selection criteria were: include mostly important attributes; create coherent descriptions of children and AAC systems; address the project's research aims; have an appropriate respondent burden. Attributes' importance was judged using BWS-1 relative importance scores.

Results: The BWS-1 survey included 19 child and 18 AAC device/system attributes and was administered to N = 93 AAC professionals. Four child and five device/system attributes were selected for the DCE, administered to N = 155 AAC professionals.

Conclusions: In this case study BWS-1 results were useful in DCE attribute selection. Four recommendations are made for future studies: define selection criteria for DCE attributes a priori; consider the impact participant's perspective will have on BWS-1 and DCE results; clearly define key terminology at the start of the study and refine it as the study progresses to reflect interim findings; BWS will be useful when there is little existing stated preference work on a topic and/or qualitative work is difficult.

Keywords: attribute development, attribute selection, best-worst scaling, discrete choice experiment, methodology.

Introduction

Discrete choice experiments (DCEs) are an increasingly popular tool to investigate preferences in healthcare.[1-4] In a DCE respondents make choices between hypothetical objects (eg, treatments) described by a set of attributes (eg, efficacy, side effects). Each attribute can take a number of levels, and by varying the levels in each question, it is possible to analyze the trade-offs respondents make between them.

Developing a DCE requires constructing a set of attributes and levels that are meaningful to respondents, capture the most salient aspects of decision making, and can address the research question. It is usually best practice to generate attributes using qualitative methods (eg, focus groups) and by drawing on previous literature and expert input.[5-11] However, this process usually produces more potential attributes than the 5 to 6 typically included in a DCE.

The process of selecting final attributes from a list of many candidates is relatively neglected in the literature. There is no consensus on the optimal way to develop DCEs,[12] but Janssen et al[13] propose a general framework that has been used by several DCEs.[14-18] However, none of the framework's 5 steps (ie, evidence synthesis, expert consultation, stakeholder engagement, pre-test interviews and pilot testing) explicitly include selecting which attributes to include from a list of candidates. Bridges et al[5] present a checklist for good practice in DCEs that includes whether attribute selection was done appropriately (item 2.2). However, they do not discuss what criteria should be used for selection or give details about what methods are useful in practice.

We present a case study in which, as part of a single preplanned research project, a best-worst scaling case 1 (BWS-1) survey[19] (also known as best-worst scaling object case) was conducted followed by a DCE,[20] with BWS-1 results used to inform DCE attribute selection. This study sought to demonstrate that

BWS-1 results can provide a sound and quantitatively robust evidence base for researchers to draw on in DCE attribute selection. We present a case study in which we highlight issues that arose, discuss advantages and disadvantages of the process, and provide a list of recommendations for future studies.

BWS-1, and the related methods of BWS case 2 (BWS-2) and BWS case 3 (BWS-3) surveys, are common in healthcare.[21] Participants are shown 5 to 6 attributes drawn from a longer list and select which they consider the best and worst (or alternative terms depending on context). The relative importance of each attribute can then be estimated on a ratio scale. BWS-1 contrasts with DCE, as participants choose a single attribute in isolation, rather than selecting a choice object composed of multiple attributes. BWS-1 is complementary to DCE because it can accommodate many more attributes (eg, Hofstede et al[22] include 53), although at the cost of not gathering as much information on any single attribute. BWS-2 and BWS-3 may be seen as substitutes for DCE, as all 3 methods examine a similar number of attributes, with different levels for the attributes.

A literature search was carried out (see in Supplemental Materials found at https://doi.org/10.1016/j.jval.2020.10.025 for details). Several studies compare BWS and DCE results, the majority of which focus on BWS-2,[23-29] although some studies either compare or combine BWS-1 and DCE results.[30-33] However, no study used BWS-1 results in DCE attribute selection, thus we believe this is the first study to do so.

DCE and BWS-1 results may have different properties, with one interpretation being that BWS elicits relative values rather than preferences.[34] However, data from each method can be analyzed using a random utility framework,[35] and there is evidence that BWS-1 results are consistent with participants' preferences in a DCE.[30-33]

The case study was carried out in the field of augmentative and alternative communication (AAC) provision for children as part of a wider research project called Identifying Appropriate Symbol Communication aids for children who are nonspeaking: enhancing clinical decision making (I-ASC).[36] Children with speech difficulties rely on AAC to express themselves. Speech difficulties may be related to several conditions (cerebral palsy, intellectual or developmental delays, autism spectrum condition). AAC encompasses many aids and techniques that supplement spoken communication, including low-tech options such as picture/letter boards and high-tech devices that may synthesize speech. The needs and abilities of children vary hugely, and AAC systems/devices must be carefully matched to children to take account of each individual's particular needs and circumstances. (The terms "system" and "device" are not fully synonymous, and "device" was used in the BWS-1 survey and "system" in the DCE survey. Further details are given in the Results section). In the current article, the word "system" is used except for in the specific context of the BWS survey.

No prior stated preference work existed in AAC to give guidance as to which potential attributes to select. In addition, the complex nature of the decision-making environment meant that we wished to present DCE choice tasks with participants choosing between which AAC system to provide for a hypothetical child vignette, constructed from a set of child attributes. We thus required attributes both for children and AAC systems, contributing to there being many potential attributes, with little evidence as to which to select. The above situation was the primary motivation to conduct a BWS-1 survey, and then incorporate the results into DCE attribute selection.

The present article focuses on survey instrument development. For details of the individual studies, see Webb et al[19] and Webb et al.[20] For more information on AAC terminology, Beukelman and Mirenda[37] is an introductory text.

Methods

Figure 1 gives a timeline of project events. Focus groups were conducted with 31 AAC clinical specialist stakeholders,[38,39] and reviews of the AAC literature were carried out.[40] Potential BWS-1 attributes were identified from this material and condensed into an initial list of child and AAC device attributes by authors (E.W., D.M., Y.L., N.R., S.J., J.G., L.M.) with expertise in AAC, speech and language therapy, qualitative research, and health economics. These lists were reduced by combining similar attributes (eg, age and educational stage). Feedback was received from other authors and a project advisory group comprising AAC professionals, people who use AAC, and relatives and support personnel of people who use AAC. Attribute identification and refinement was done by unstructured discussion among the authors until all were in agreement.

Two BWS-1 exercises, one for child and one for AAC device attributes, were administered online as a single questionnaire to AAC professionals who work with children. The questionnaire was piloted with N = 5 AAC professionals and minor changes made to language and formatting. Figure 2 shows an example task. Results were analyzed using mixed logit models, producing relative importance scores (RIS) for attributes.[19]

After the BWS-1, a DCE was carried out. Respondents chose an AAC device/system for a vignette describing a hypothetical child. Child and AAC system attributes were selected from BWS-1 attributes. This was achieved by consensus discussions among all authors.

Attributes were selected using the following criteria:

1. Most attributes should be anticipated to be highly important in participants' decision making.
2. DCE results would address the research project's specific aims.
3. DCE attributes should form coherent and realistic descriptions for both children and AAC systems.
4. The number of attributes should be small enough that the cognitive burden would be reasonable, and that models could be estimated given the expected sample size.

Anticipated attribute importance (criterion 1) was judged using attribute RIS from the BWS-1. Overall, attributes were selected pragmatically: no explicit weighting was given to the criteria, but individual attributes contributed more to satisfying some criteria than others.

Although BWS-1 RIS was important in selecting attributes, there was no preset threshold above which an attribute would automatically be selected (eg, if an attribute was in the top 3, it would be selected). The research project particularly aimed to investigate symbols used in AAC systems, thus attributes related to graphic symbols were prioritized (criterion 2). However, no a priori decisions were made to include any attributes in the DCE, regardless of BWS-1 results. Whether DCE attributes satisfied criteria 3 and 4 was assessed first by members of the research team with AAC expertise and then verified in piloting.

The attribute language was refined in the light of BWS-1 findings and to fit a different decision-making context. (For example, in a DCE levels are presented alongside attributes, unlike in BWS-1.)

Attribute selection and refinement was carried out by unstructured discussion among authors until consensus was

achieved. A DCE was created and piloted with N = 5 AAC professionals, with minor changes made to wording and formatting. Figure 3 shows an example choice task. The DCE was administered online to AAC professionals.[20] Ethical approval was received for the study from a National Health Service (NHS) Research Ethics Committee (REC reference 6/NW/0165), and informed consent was obtained from participants at the start of the survey.

Results

Initial lists of 31 child and 29 AAC device related attributes (Appendix Table 1 and Appendix Table 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2020.10.025) were condensed into lists of 19 child and 18 AAC device attributes (Table 1) used in the BWS-1. Completed responses from N = 93 participants were analyzed, producing meaningful results.[19]

Following the BWS-1, 4 child and 5 AAC system attributes were selected for the DCE, given in Table 2, as were attributes' rank in terms of BWS-1 RIS. The wording of 3 out of 4 child attributes and 3 out of 6 AAC system attributes was edited between BWS-1 and DCE.

Predicted future needs and abilities had the sixth highest RIS. However, it was decided that for vignettes to give a realistic summary of a child, it was necessary to include an indication of the child's predicted future trajectory.[41,42] "Children's predicted future needs and abilities" was changed to "predicted future skills and abilities" to make levels easier to create. Levels needed to capture a child progressing or regressing in ability and their needs changing due to different abilities. However, it was difficult to succinctly capture changing needs due to developing/regressing abilities, so the wording was changed to remove the reference to needs.

Graphic representation ranked 12th out of 18 in RIS but was included as 1 of 5 AAC system attributes, due to the wider research project's focus on graphic representation (the type of symbol a system uses).

Some changes in wording were made because of the different formats of BWS-1 and DCE. The layout of the DCE (Fig 3) displayed AAC system attributes as short "headlines" followed by longer descriptions below. This was possible owing to attribute labels remaining unchanged in each DCE question, so respondents did not have to read them every time. In the BWS-1, different attributes were shown in each question, and respondents had to read the whole list every time, making a similar visual layout to the DCE impractical. Thus the information contained in the DCE attribute descriptions had to be included in the attribute name for the BWS-1 survey. For example, "vocabulary or language packages" in the BWS-1 became "vocabulary sets" in the DCE.

The term "AAC device" used in the BWS-1 was changed to "AAC system." Although nomenclature is not completely rigid, participants perceived "AAC device" to imply high-tech powered electronic devices, whereas "AAC system" encompassed both powered high-tech devices as well as paper- and person-based AAC strategies—for example, boards with letters, words, or pictures on them. The change was possible as attributes specific to high-tech powered devices (eg, battery life) were not included in the DCE.

DCE pilot testing revealed a need for a slightly different target population than for the BWS-1. Testing showed some AAC professionals, although they had clinical experience, had little input into decisions about which AAC system to choose for a child (for example, occupational therapists who help children with adjustments to AAC systems selected by other members of a clinical team). These individuals struggled to give meaningful answers drawing on their own experience. No similar problems were found in BWS-1 testing. A screening question was added to the survey that redirected individuals without the necessary experience for completing the DCE tasks to answer only demographic questions. The specific wording of the questions was: "I confirm my work involves assessing children for AAC systems and I contribute to the decision making in relation to the language and vocabulary organisation within AAC systems (Yes/No)."

The survey was completed by N = 172 participants, of which 155 answered yes to the screening question and answered DCE questions. Analysis produced meaningful results.[20]

Discussion

The DCE attribute selection criteria detailed in the Methods section were similar to those implicitly or explicitly used in developing most DCEs (see, eg, item 6 of the International Society for Pharmacoeconomics and Outcomes Research Checklist for Conjoint Analysis in Health[5]). The difference here is that candidate attributes' anticipated importance was informed by BWS-1 results, and it is instructive to examine how this influenced the application of the criteria.

Figure 1. Timeline of project events.

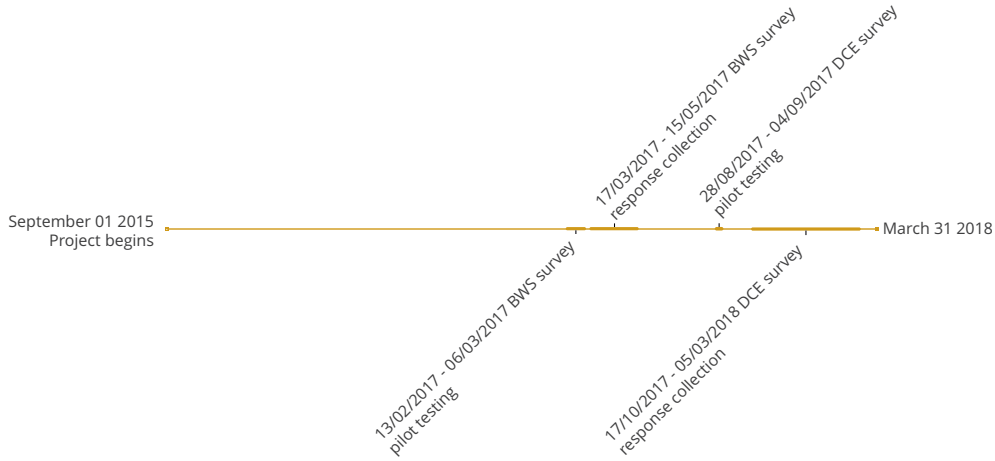
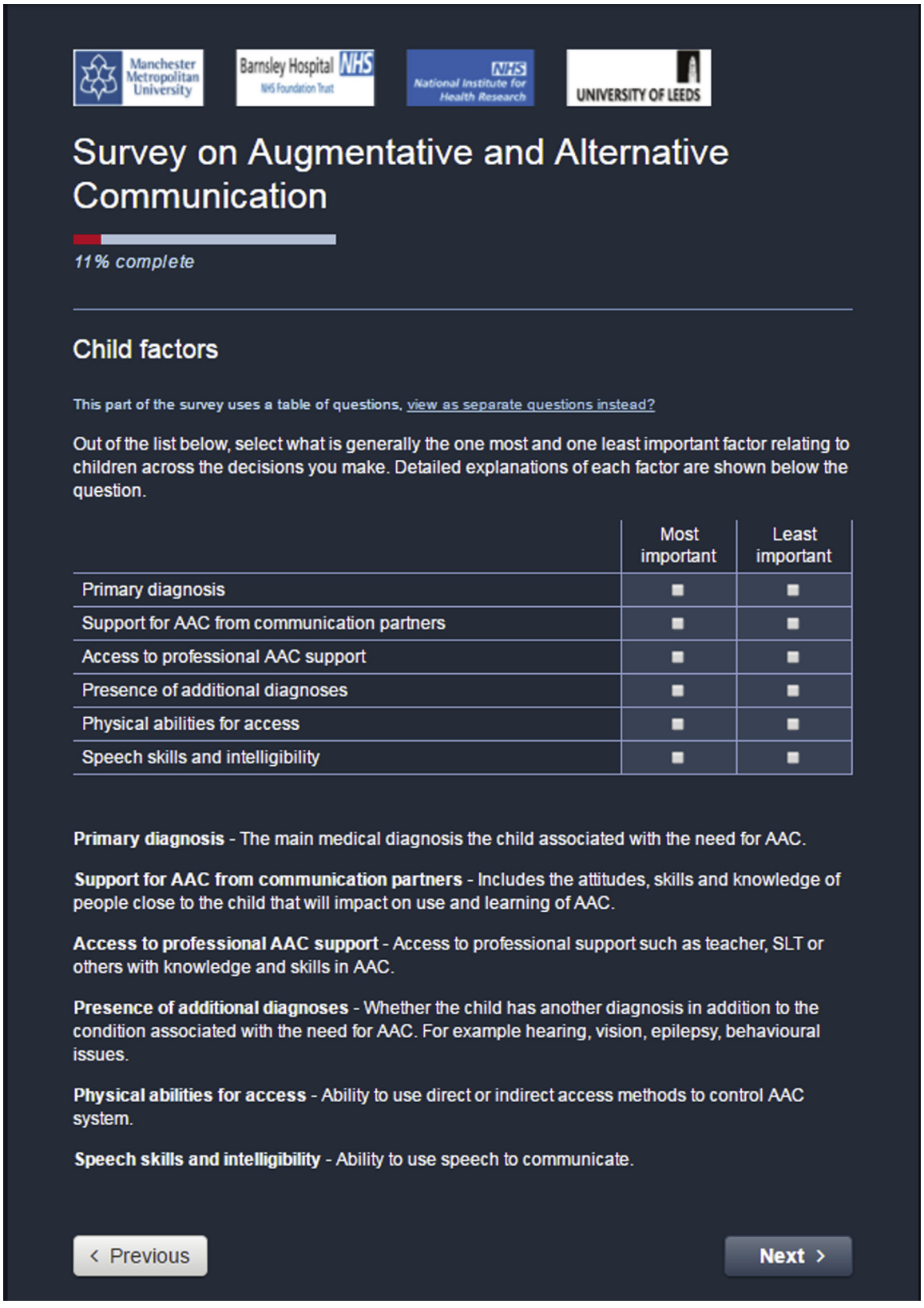


Figure 2. Example best-worst scaling task.



Manchester Metropolitan University | Barnsley Hospital NHS Foundation Trust | NHS National Institute for Health Research | UNIVERSITY OF LEEDS

Survey on Augmentative and Alternative Communication

11% complete

Child factors

This part of the survey uses a table of questions, view as separate questions instead?

Out of the list below, select what is generally the one most and one least important factor relating to children across the decisions you make. Detailed explanations of each factor are shown below the question.

	Most important	Least important
Primary diagnosis	■	■
Support for AAC from communication partners	■	■
Access to professional AAC support	■	■
Presence of additional diagnoses	■	■
Physical abilities for access	■	■
Speech skills and intelligibility	■	■

Primary diagnosis - The main medical diagnosis the child associated with the need for AAC.

Support for AAC from communication partners - Includes the attitudes, skills and knowledge of people close to the child that will impact on use and learning of AAC.

Access to professional AAC support - Access to professional support such as teacher, SLT or others with knowledge and skills in AAC.

Presence of additional diagnoses - Whether the child has another diagnosis in addition to the condition associated with the need for AAC. For example hearing, vision, epilepsy, behavioural issues.

Physical abilities for access - Ability to use direct or indirect access methods to control AAC system.

Speech skills and intelligibility - Ability to use speech to communicate.

‹ Previous | Next ›

Figure 3. Example discrete choice experiment task.

Child B has delayed expressive and receptive language and able to use aided AAC for a few communicative functions. Child B is only motivated to communicate through methods other than symbol communication systems. Child B is predicted to maintain current skills and abilities (plateau).

	System 1	System 2	System 3
Graphic Representation Primary type of graphic symbol used	Ideographic symbols	Photo symbols	Pictographic symbols
Consistency of layout Consistency of layout of symbols on pages, including when navigating through pages to select desired output.	Highly consistent layout	Somewhat consistent layout	Inconsistent layout
Vocabulary sets Pre-determined vocabulary or language package provided	Commercially provided sets without language progression	Commercially provided sets with language progression	Commercially provided sets without language progression
Size of vocabulary The size of the output vocabulary available within the aided AAC system.	More than 1000 vocabulary items	50-1000 vocabulary items	Up to 50 vocabulary items
Type of vocabulary organisation Primary format used to organise the vocabulary within the aided AAC system	Pragmatic organisation	Semantic syntactic organisation	Visual scene display
For this child I would choose:	○	○	○

On a scale from 1 to 7, how good a match is your chosen device for this child? (1=very unsuitable, 7=very suitable)

1 = very unsuitable	2	3	4	5	6	7 = very suitable
○	○	○	○	○	○	○

35%

We did not simply select the top ranked BWS-1 attributes, because this would not have resulted in an appropriate DCE instrument. The attributes graphic representation and battery life illustrate how different selection criteria were more relevant for different attributes. There was no significant difference in the attributes' RIS, being ranked 12th and 13th, respectively, out of 18. Although neither was expected to be among the most important in DCE participants' decision-making (criterion 1), graphic representation was the attribute most suited to address the project's focus on symbols used in AAC systems. Thus it was selected for inclusion under criterion 2. Battery life was not useful to address the project's aims, nor was it necessary for a realistic description of an AAC system (criterion 3). Thus it was excluded from the DCE, reducing the number of attributes in line with criterion 4.

There was less reliance on the research team's prior beliefs when applying criterion 1 than would otherwise have been the case. The authors' prior belief was that battery life would be important, and there were even worries that it might dominate other attributes. However, the BWS-1 provided detailed quantitative information from many AAC professionals indicating battery life did not fulfill the inclusion criteria.

Table 1. Best-worst scaling case 1 attributes.

Child attributes	Included in DCE?
Access to professional AAC support	
Attention level	
Child's determination and persistence	Yes
Child's receptive and expressive language abilities	Yes
Communication ability with aided AAC	Yes
Educational stage	
Functional visual skills	
History of aided AAC use	
Insight into own communicative skills	
Level of fatigue	
Level of learning ability	
Literacy ability	
Mobility	
Physical abilities for access	
Predicted future needs and abilities	Yes
Presence of additional diagnoses	
Primary diagnosis	
Speech skills and intelligibility	
Support for AAC from communication partners	
AAC device attributes	
Additional assistive technology functions	
Appearance	
Battery life	
Consistency of layout and navigation	Yes
Cost	
Durability and reliability	
Ease of customization	
Ease of mounting on a range of equipment	
Graphic representation	Yes
Number of cells per page	
Number of key presses required to generate symbol or text output	
Portability	
Range of access methods	
Size of output vocabulary	Yes
Supplier support	
Type of vocabulary organization	Yes
Vocabulary or language package(s)	Yes
Voice	

AAC indicates augmentative and alternative communication; DCE, discrete choice experiment.

For some attributes, criterion 3 was most relevant. For example, predicted future skills and abilities, with the sixth highest RIS, was included as 1 of 4 child attributes, as it was necessary to create realistic child vignettes. Because 4 child attributes was believed to be a maximum to not overburden participants (criterion 4), this necessitated not including 1 of the top 4 ranked child attributes. Support for AAC from communication partners, ranked second, was omitted because it was thought to be the least necessary to form realistic child vignettes or to address the project's aims.

The previous discussion highlights that selecting DCE attributes remains a pragmatic and subjective process after incorporating a BWS-1. However, BWS-1 results can give clarity to both researchers and the reader. Because the BWS-1 results were also published, the reader can observe researchers' selection decisions, in line with recommendations[43] that studies report greater detail on instrument development.

It was assumed that the relative importance participants ascribed to BWS-1 attributes was informative about how DCE participants would respond. BWS-1 and DCE have different properties. It is argued that they elicit different sorts of information, with DCEs eliciting preferences and BWS values,[24,34] and there is evidence they measure different constructs.[23] However, a random utility framework is usually used to analyze both BWS-1 and DCE data,[35,44] and there is evidence that BWS-1 RIS corresponds to utility weights in a DCE.[30-33]

BWS-1 results gave additional information to update prior beliefs about which candidate attributes fulfilled the criteria for inclusion in the DCE. In this light, methodological differences seem less important, since the assumption that BWS RIS is informative about the relative importance of DCE attributes is both reasonable and supported by empirical evidence,[33] not least the current project's results. For example, graphical representation had the lowest RIS of any BWS-1 attribute included in the DCE. Although it had a significant effect on decision making in the DCE, it also had the lowest relative importance.

BWS-1 is an alternative to existing methods that also have significant differences in their properties and the constructs they measure. Studies using nominal group technique often perform a ranking exercise at the end.[6] Others studies use Q-methodology, in which participants sort candidate attributes from most to least important on a board.[45,46] BWS-1 has advantages over these alternatives; for example, it imposes a lower cognitive burden than ranking,[47] potentially giving more reliable results, as well as allowing statistical significance testing of the difference in importance of 2 attributes.

The conclusions drawn from BWS-1 and DCE depend on their context and on what perspective participants are asked to complete tasks from.[48,49] The current study examined a particular perspective with a professional population, whereas the majority of healthcare DCEs target a patient or general public population,[1,4] and it is possible to elicit preferences from many different perspectives.[48,49] Although in principle incorporating BWS-1 into DCE attribute development should be possible in other contexts, a single study can demonstrate feasibility in only one. We discuss first methodological lessons from the current context, and then potential issues that might arise in other areas.

In the current study, BWS-1 and DCE decisions were made in different contexts. The BWS-1 measured the relative importance of child and AAC device attributes in general, that is, averaged over the case mix AAC professionals encounter in practice. In contrast, the DCE examined their decision making for a specific child vignette. These correspond to the social ex ante case 1 and proxy ex post perspectives, respectively, in Tsuchiya and Watson's[49] classification system. Having different perspectives had benefits, in that each addressed the research question of how AAC professionals make decisions from different angles. The DCE examined trade-offs made between AAC system attributes for a child with specific characteristics. However, the set of child vignettes shown to DCE participants was not representative of the

Table 2. Discrete choice experiment attributes and levels and corresponding best-worst scaling case 1 attributes.

BWS attribute	BWS rank	DCE attribute	DCE levels
Child attributes			
Child's receptive and expressive language abilities	1	Receptive and expressive language	Delayed
			Receptive language exceeding expressive language
Communication ability with aided AAC	3	Communication ability with AAC	No previous AAC experience
			Abe to use AAC for a few communicative functions
			Able to use AAC for a range of communicative functions
Child's determination and persistence	4	Child's determination and persistence	Does not appear motivated to communicate through any methods and means
			Motivated to communicate through symbol communication systems
			Only motivated to communicate through methods other than symbol communication
Predicted future needs and abilities	6	Predicted future skills and abilities	Regression
			Plateau
			Progression
AAC device/system attributes			
Vocabulary or language package(s)	1	Vocabulary sets	No vocabulary set
			Fixed vocabulary set
			Vocabulary set with staged progression
Consistency of layout and navigation	2	Consistency of layout	Consistency of some aspects of layout
			Consistency of all aspects of layout
			Idiosyncratic layout
Type of vocabulary organization	5	Type of vocabulary organization	Visual scene
			Taxonomic
			Semantic-syntactic
			Pragmatic
Size of output vocabulary	7	Size of vocabulary	Up to 50 vocabulary items
			50-1000 vocabulary items
			More than 1000 vocabulary items
Graphic representation	12	Graphic representation	Photos
			Pictographic symbol set
			Ideographic symbol system (with rules or encoding)
			Text

AAC indicates augmentative and alternative communication; BWS, best-worst scaling; DCE, discrete choice experiment.

population of children seen by AAC professionals. The BWS-1 measured how important attributes were in typical practice, which the DCE could not. The disadvantage of different perspectives in the BWS-1 and DCE was a greater risk that BWS-1 RIS would be misleading as to attributes' importance in a DCE.

A different perspective in the DCE was essential in this area owing to the large amount of heterogeneity in the population of children who would benefit from AAC. In areas outside AAC this may not be necessary. For example, DCEs may be used to investigate clinician preferences for screening programs.[50,51] In that case, it would be straightforward to have the same perspective by asking in a BWS-1 what factors are considered most and least important for screening programs, and then in a DCE asking to choose which of 2 programs to implement. This perspective is classified as proxy ex ante by Tsuchiya and Watson.[49] When studying patient preferences or willingness-to-pay for treatments it would also be natural to have the same perspective for a BWS-1 ("What is the most/least important aspect of treatment?") and a DCE ("What is your preferred treatment?"). This perspective is classified as personal ex post by Tsuchiya and Watson.[49]

Because perspective may influence how participants respond to surveys,[48,52-54] a possible advantage of using the same perspective for BWS-1 and DCE is greater correspondence between the participants' responses in the 2 formats. This may be especially important if the research is eliciting WTP. In this case, external validity could be compromised if a highly valued attribute

is omitted from the DCE. Having the same perspective would also enable the estimation of joint models with BWS-1 and DCE data,[33] although care should be taken in interpreting the results of such models given the methods' different properties.[34]

There are also drawbacks to using the same perspective in BWS-1 and DCE. Using different perspectives gave complementary findings, giving insight both into AAC professionals' priorities over the population of children they see in practice, as well as the detailed trade-offs made between AAC attributes when deciding for a specific child.

Depending on a study's context and target population, there may be a choice between using the same or differing perspectives in BWS-1 and DCE. In such cases, researchers should carefully consider the trade-offs between potentially greater correspondence between BWS-1 RIS and DCE attribute importance and being able to address research questions from different perspectives. Even where there is little freedom to choose different perspectives, it is still recommended to consider the impact of participants' perspectives on the interpretation of results.

In addition to theoretical considerations, this study highlighted practical issues in incorporating a BWS-1 into DCE attribute selection. First, some necessary and desirable changes to terminology were made in translating BWS-1 attributes to DCE attributes. The shift from AAC device to AAC system was discussed above, and changes in attributes' wording are shown in Table 1. Changing terminology will have altered, to some extent, the constructs that attributes were capturing. However, the different perspectives discussed above meant BWS-1 and DCE attributes would inevitably be conceptually different.

The experience of this study was that language and terminology were important to consider and refine throughout the project. A recommendation for future studies is to carefully define key terminology at the project's inception, and to continuously reassess and refine it as the project progresses.

Many DCEs gather information on attributes' relative importance from existing DCEs in the area. However, in AAC, no stated preference studies existed prior to this project, making BWS-1 results especially valuable. Our experience was that having BWS-1 results available in DCE attribute selection was extremely helpful. On this basis, it is recommended that incorporating BWS-1 results in DCE attribute selection will be most useful in similar areas with many potential attributes and little or no previous stated preference work.

Even in fields with several prior DCEs, BWS-1 results could still be useful, as there may be a risk of "path dependency" in the literature. If the first stated preference study in a given field omits attributes, subsequent studies are more likely to omit them as well, leading to only a narrow range of attributes being examined. For example, Janssen et al[13] found only 10 distinct attribute categories in the area of type 2 diabetes, and Webb et al[55] found 13 in the field of multiple sclerosis. In addition, new clinical and social developments may decrease the current relevance of past studies.

Adding a BWS-1 to DCE instrument development means a greater workload and longer project timelines than a DCE alone. In the current project, it was an advantage that it examined professionals' views, so several members of the research team were part of the target population. If patient preferences were being elicited, then stakeholder involvement would have been more resource intensive, both in developing BWS-1 attributes and translating them to DCE attributes. For example, when creating levels for DCE attributes, the research team was able to verify they were meaningful and in an appropriate range. With a patient population this may require greater resources—for example, focus groups to obtain patients' input.

It is a disadvantage of the current study that it presents only 1 case study in 1 clinical area. It also targets a professional population, whereas DCEs with patient respondents are more common.[4] Future research could usefully investigate the feasibility of the method with patients.

The population of AAC professionals who work with children in the United Kingdom is small, and many will have responded to both surveys (though data is not available on how many). It is not desirable to have DCE respondents who previously may have completed the BWS-1, since it is not known whether and in what way the earlier survey may have influenced their DCE choices. Thus a disadvantage of this method is that with a small potential respondent pool, there will inevitably be an overlap in participation between the 2 surveys.

Conclusion

We have demonstrated that a BWS-1 can be useful in informing attribute selection for a DCE. This study has also shown that using BWS-1 and DCE as part of a predefined research project can produce synergies, giving greater insight than carrying out each study singly, especially in a complex decision-making environment where no prior stated preference work has been conducted.

We propose that incorporating a BWS-1 into the process of DCE instrument development is a valid and robust method, and we finish with a set of recommendations for future studies:

1. Have a clear plan of how DCE attribute selection will proceed, including what selection criteria will be used in addition to BWS-1 results. For example, researchers may establish a priori that to address the research question some attributes will be included in the DCE regardless of BWS-1 results.
2. Clearly define the decision-making context and participants' perspectives in both the BWS-1 and DCE, and consider the impact perspective will have on findings.
3. Clearly define key terminology at the beginning of the project and refine it as necessary as it progresses in the light of interim findings.
4. Consider using a BWS-1 to inform DCE attribute development and selection, especially in areas with little prior stated preference work and/or extensive qualitative work is difficult. However, there may still be many advantages to the approach where this is not the case.

Randall, Goldbart, Moulam, Murray); Institute for Transport Studies, University of Leeds, Leeds, England, UK (Judge).

Correspondence: Edward J.D. Webb, Leeds Institute of Health Sciences, University of Leeds, Leeds, England, United Kingdom. Email: e.j.d.webb@leeds.ac.uk

Author Contributions: *Concept and design*: Webb, Meads, Lynch, Judge, Randall, Goldbart, Meredith, Moulam, Hess, Murray
Acquisition of data: Webb, Meads, Lynch, Judge, Randall, Goldbart, Meredith, Moulam, Hess, Murray
Analysis and interpretation of data: Webb, Meads, Lynch, Judge, Randall, Goldbart, Meredith, Moulam, Hess, Murray
Drafting of the manuscript: Webb, Meads, Lynch, Judge, Randall, Goldbart, Meredith, Moulam, Hess, Murray
Critical revision of the paper for important intellectual content: Webb, Meads, Lynch, Judge, Randall, Goldbart, Meredith, Moulam, Hess, Murray
Obtaining funding: Judge

Conflict of Interest Disclosures: The authors reported no conflicts of interest.

Funding/Support: This independent research was funded by the National Institute for Health Research, UK (Health Services and Delivery Research Project: 14/70/153—Identifying appropriate symbol communication aids for children who are nonspeaking: enhancing clinical decision-making). The views expressed in this article are those of the authors and not necessarily those of the NHS, the National Institute for Health Research, or the Department of Health. Stephane Hess acknowledges additional support by the European Research Council through the consolidator grant 615596-DECISIONS.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Ethical Approval: Ethical approval was received for the study from an National Health Service Research Ethics Committee (REC reference 6/NW/0165) and informed consent was obtained from participants at the start of the survey.

Data Availability: Survey data is not publicly available as respondent consent was not obtained for this. However, it is available on request to the corresponding author or to Leeds Institute of Health Sciences if a formal data sharing agreement is entered into.

Acknowledgment: The authors thank Muireann McCleary and the Speech and Language Therapy team at the Central Remedial Clinic, Dublin, who piloted and gave feedback on the survey; Helen Whittle; and to participants who responded to the survey.

REFERENCES

1. Clark MD, Determann D, Petrou S, Moro D, de Bekker-Grob EW. Discrete choice experiments in health economics: a review of the literature. *Pharmacoeconomics*. 2014;32:883–902.
2. de Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. *Health Econ*. 2012;21:145–172.
3. Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Appl Health Econ Health Policy*. 2003;2:55–64.
4. Soekhai V, de Bekker-Grob EW, Ellis AR, Vass CM. Discrete choice experiments in health economics: past, present and future. *Pharmacoeconomics*. 2019;37:201–226.
5. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health—a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value Health*. 2011;14:403–413.
6. dosReis S, Castillo WC, Ross M, Fitz-Randolph M, Vaughn-Lee A, Butler B. Attribute development using continuous stakeholder engagement to prioritize treatment decisions: a framework for patient-centered research. *Value Health*. 2016;19:758–766.
7. Hollin IL, Young C, Hanson C, Bridges JF, Peay H. Developing a patient-centered benefit-risk survey: a community-engaged process. *Value Health*. 2016;19:751–757.
8. Michaels-Igbokwe C, Lagarde M, Cairns J, Terris-Prestholt F. Using decision mapping to inform the development of a stated choice survey to elicit youth preferences for sexual and reproductive health and HIV services in rural Malawi. *Social Sci Med*. 2014;105:93–102.
9. Kløjgaard ME, Bech M, Søgaard R. Designing a stated choice experiment: the value of a qualitative process. *J Choice Modelling*. 2012;5:1–18.
10. Coast J, Al-Janabi H, Sutton EJ, et al. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. *Health Econ*. 2012;21:730–741.
11. Coast J, Horrocks S. Developing attributes and levels for discrete choice experiments using qualitative methods. *J Health Services Res Policy*. 2007;12:25–30.
12. Wittenberg E. Instrument development in choice experiments. Commentary on: "Applying a framework for instrument development of a choice experiment to measure treatment preferences in type 2 diabetes.". *Patient*. 2016;9:379–381.
13. Janssen EM, Segal JB, Bridges JFP. A framework for instrument development of a choice experiment: an application to type 2 diabetes. *Patient*. 2016;9:465–479.
14. Seo J, Smith BD, Estey E, Voyard E, O'Donoghue B, Bridges JFP. Developing an instrument to assess patient preferences for benefits and risks of treating acute myeloid leukemia to promote patient-focused drug development. *Curr Med Res Opinion*. 2018;34:2031–2039.
15. Bridges JFP, Oakes AH, Reinhart CA, Voyard E, O'Donoghue B. Developing and piloting an instrument to prioritize the worries of patients with acute myeloid leukemia. *Patient Preference Adherence*. 2018;12:647–655.
16. Beusterien K, Chan E, Such P, et al. Development of a stated-preference instrument to prioritize treatment goals in recent onset schizophrenia. *Curr Med Res Opinion*. 2017;33:2129–2136.
17. Katz DA, Stewart KR, Paez M, et al. Development of a discrete choice experiment (DCE) questionnaire to understand veterans' preferences for tobacco treatment in primary care. *Patient*. 2018;11:649–663.
18. Barber S, Bekker H, Marti J, Pavitt S, Khambay B, Meads D. Development of a discrete-choice experiment (DCE) to elicit adolescent and parent preferences for hypodontia treatment. *Patient*. 2019;12:137–148.
19. Webb EJ, Meads D, Lynch Y, et al. What's important in AAC decision making for children? Evidence from a best–worst scaling survey. *Augment Altern Comm*. 2019:1–15.
20. Webb EJ, Lynch Y, Meads D, et al. Finding the best fit: examining the decision-making of augmentative and alternative communication professionals in the UK using a discrete choice experiment. *BMJ Open*. 2019;9.
21. Cheung KL, Wijnen BF, Hollin IL, et al. Using best-worst scaling to investigate preferences in health care. *Pharmacoeconomics*. 2016;34:1195–1209.
22. Hofstede SN, van Bodegom-Vos L, Wentink MM, et al. Most important factors for the implementation of shared decision making in sciatica care: ranking among professionals and patients. *PLoS One*. 2014;9:e94176.
23. Whitty JA, Gonçalves ASO. A systematic review comparing the acceptability, validity and concordance of discrete choice experiments and best–worst scaling for eliciting preferences in healthcare. *Patient*. 2018;11:301–317.
24. Krucien N, Watson V, Ryan M. Is best–worst scaling suitable for health state valuation? A comparison with discrete choice experiments. *Health Econ*. 2017;26:e1–e16.
25. Flynn TN, Peters TJ, Coast J. Quantifying response shift or adaptation effects in quality of life by synthesising best-worst scaling and discrete choice data. *J Choice Modelling*. 2013;6:34–43.
26. Whitty JA, Ratcliffe J, Chen G, Scuffham PA. Australian public preferences for the funding of new health technologies: a comparison of discrete choice and profile case best-worst scaling methods. *Med Decision Making*. 2014;34:638–654.
27. Whitty JA, Walker R, Golenko X, Ratcliffe J. A think aloud study comparing the validity and acceptability of discrete choice and best worst scaling methods. *PloS One*. 2014;9:e90635.
28. Potoglou D, Burge P, Flynn T, et al. Best–worst scaling vs. discrete choice experiments: an empirical comparison using social care data. *Soc Sci Med*. 2011;72:1717–1727.
29. Severin F, Schmidtke J, Mühlbacher A, Rogowski WH. Eliciting preferences for priority setting in genetic testing: a pilot study comparing best-worst scaling and discrete-choice experiments. *Eur J Human Genetics*. 2013;21:1202–1208.
30. Mansfield C, Ndife B, Chen J, Gallaher K, Ghate S. Patient preferences for treatment of metastatic melanoma. *Future Oncol*. 2019;15:1255–1268.
31. Marshall T, Pugh A, Fairchild A, Hass S. Patient preferences for device-aided treatments indicated for advanced Parkinson disease. *Value Health*. 2017;20:1383–1393.
32. Turk D, Boeri M, Abraham L, et al. Patient preferences for osteoarthritis pain and chronic low back pain treatments in the United States: a discrete-choice experiment. *Osteoarthritis Cartilage*. 2020;08:1202–1213.
33. Zhang J, Johnson FR, Mohamed AF, Hauber AB. Too many attributes: a test of the validity of combining discrete-choice and best–worst scaling data. *J Choice Modelling*. 2015;15:1–13.
34. Coast J, Flynn TN, Natarajan L, et al. Valuing the ICECAP capability index for older people. *Social Sci Med*. 2008;67:874–882.
35. Flynn TN, Louviere JJ, Peters TJ, Coast J. Best–worst scaling: what it can do for health care research and how to do it. *J Health Econ*. 2007;26:171–189.
36. Murray J, Lynch Y, Goldbart J, et al. Identifying appropriate communication aids for children and public involvement in research: a mixed methods study.

In press. Report No. 14/70/153. https://www.journalslibrary.nihr.ac.uk/hsdr/hsdr08450/#/abstract
37. Beukelman DR, Mirenda P. *Augmentative and Alternative Communication: Supporting Children and Adults with Complex Communication Needs*. Paul H. Brookes; 2013.
38. Murray J, Lynch Y, Meredith S, et al. Professionals' decision-making in recommending communication aids in the UK: competing considerations. *Augment Altern Comm*. 2019;35(3):167–179.
39. Lynch Y, Murray J, Moulam L, et al. Decision-making in communication aid recommendations in the UK: cultural and contextual influencers. *Augment Altern Comm*. 2019;35(3):180–192.
40. Judge S, Randall N, Goldbart J, et al. The language and communication attributes of graphic symbol communication aids: a systematic review and narrative synthesis. *Disabil Rehabil Assist Technol*. 2019;15(6):652–662.
41. Thistle JJ, Wilkinson KM. Building evidence-based practice in AAC display design for young children: current practices and future directions. *Augment Altern Commun*. 2015;31:124–136.
42. Dada S, Alant E. The effect of aided language stimulation on vocabulary acquisition in children with little or no functional speech. *Amer J Speech Lang Pathol*. 2009;18:50–64.
43. Janssen EM, Bridges JF. *Art and Science of Instrument Development for Stated-Preference Methods*. Springer; 2017.
44. Lancsar E, Fiebig DG, Hole AR. Discrete choice experiments: a guide to model specification, estimation and software. *Pharmacoeconomics*. 2017;35:697–716.
45. Armatas CA, Venn TJ, Watson AE. Applying Q-methodology to select and define attributes for non-market valuation: a case study from Northwest Wyoming, United States. *Ecological Econ*. 2014;107:447–456.
46. Kerr GN, Swaffield SR. Identifying cultural service values of a small river in the agricultural landscape of Canterbury, New Zealand, using combined methods. *Soc Nat Resour*. 2012;25:1330–1339.
47. Louviere JJ, Street D, Burgess L, Wasi N, Islam T, Marley AA. Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information. *J Choice Modelling*. 2008;1:128–164.
48. Dolan P, Olsen JA, Menzel P, Richardson J. An inquiry into the different perspectives that can be used when eliciting preferences in health. *Health Econ*. 2003;12:545–551.
49. Tsuchiya A, Watson V. Rethinking the different perspectives that can be used when eliciting preferences in health. *Health Econ*. 2017;26:e103–e107.
50. Sicsic J, Krucien N, Franc C. What are GPs' preferences for financial and non-financial incentives in cancer screening? Evidence for breast, cervical, and colorectal cancers. *Social Sci Med*. 2016;167:116–127.
51. Boone D, Mallett S, Zhu S, et al. Patients' & healthcare professionals' values regarding true- and false-positive diagnosis when colorectal cancer screening by CT colonography: discrete choice experiment. *PloS One*. 2013;8:e80767.
52. Damschroder LJ, Roberts TR, Zikmund-Fisher BJ, Ubel PA. Why people refuse to make tradeoffs in person tradeoff elicitations: a matter of perspective? *Med Decision Making*. 2007;27:266–280.
53. Bobinac A, Van Exel NJA, Rutten FF, Brouwer WB. Valuing QALY gains by applying a societal perspective. *Health Econ*. 2013;22:1272–1281.
54. Gyrd-Hansen D. Willingness to pay for a QALY. *Pharmacoeconomics*. 2005;23:423–432.
55. Webb EJD, Meads D, Eskyte I, et al. A systematic review of discrete-choice experiments and conjoint analysis studies in people with multiple sclerosis. *Patient*. 2018;11:391–402.