



Respondent Understanding in Discrete Choice Experiments: A Scoping Review

Alison Pearce¹ · Mark Harrison² · Verity Watson³ · Deborah J. Street⁴ · Kirsten Howard¹ · Nick Bansback⁵ · Stirling Bryan⁵

Accepted: 30 September 2020 / Published online: 3 November 2020
© The Author(s) 2020

Abstract

Introduction Despite the recognised importance of participant understanding for valid and reliable discrete choice experiment (DCE) results, there has been limited assessment of whether, and how, people understand DCEs, and how ‘understanding’ is conceptualised in DCEs applied to a health context.

Objectives Our aim was to identify how participant understanding is conceptualised in the DCE literature in a health context. Our research questions addressed how participant understanding is defined, measured, and used.

Methods Searches were conducted (June 2019) in the MEDLINE, EMBASE, PsychINFO and Econlit databases, as well as hand searching. Search terms were based on previous DCE systematic reviews, with additional understanding keywords used in a proximity-based search strategy. Eligible studies were peer-reviewed journal articles in the field of health, related to DCE or best-worst scaling type 3 (BWS3) studies, and reporting some consideration or assessment of participant understanding. A descriptive analytical approach was used to chart relevant data from each study, including publication year, country, clinical area, subject group, sample size, study design, numbers of attributes, levels and choice sets, definition of understanding, how understanding was tested, results of the understanding tests, and how the information about understanding was used. Each study was categorised based on how understanding was conceptualised and used within the study.

Results Of 306 potentially eligible articles identified, 31 were excluded based on titles and abstracts, and 200 were excluded on full-text review, resulting in 75 included studies. Three categories of study were identified: applied DCEs ($n = 52$), pre-testing studies ($n = 7$) and studies of understanding ($n = 16$). Typically, understanding was defined in relation to either the choice context, such as attribute terminology, or the concept of choosing. Very few studies considered respondents’ engagement as a component of understanding. Understanding was measured primarily through qualitative pretesting, rationality or validity tests included in the survey, and participant self-report, however reporting and use of the results of these methods was inconsistent.

Conclusions Those conducting or using health DCEs should carefully select, justify, and report the measurement and potential impact of participant understanding in their specific choice context. There remains scope for research into the different components of participant understanding, particularly related to engagement, the impact of participant understanding on DCE validity and reliability, the best measures of understanding, and methods to maximise participant understanding.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s40271-020-00467-y>) contains supplementary material, which is available to authorized users.

✉ Alison Pearce
Alison.pearce@sydney.edu.au

Extended author information available on the last page of the article

Key Points for Decision Makers

We consider understanding in the context of discrete choice experiments (DCEs) as an overarching concept that can be defined as participants choosing rationally based on comprehending both the choice context and choice task, and being willing to participate.

We identify a broad range of definitions, measurement approaches and proposed implications of participant understanding in the DCE literature.

While some common approaches emerge, there is little consensus around how to maximise participant understanding of DCEs in the health context.

There remains significant scope for further work on measuring and improving DCE participant understanding in health, which may draw on the use of DCEs in other fields, behavioural and experimental economics, and psychology.

1 Introduction

Discrete choice experiments (DCEs) are a quantitative method used to measure preferences for goods and services. The application of DCEs to measure preferences for health and health care is increasing in geographic scope, areas of application, and sophistication of the design and analysis [1].

The health context can present unique challenges that make DCEs complex to complete. They often include attributes related to risks or probability, and these concepts can be difficult for people to understand and accurately interpret [2–4]. Many health DCEs include unfamiliar medical terminology and some people, particularly those with lower health literacy, may have difficulty understanding the context of the choice they are asked to make [5]. The emotional intensity associated with making decisions about health and healthcare may result in heuristics that change people's choice behaviour [6] and many health scenarios may be difficult to imagine, such as being in a wheelchair or having to pay for health care in a universal healthcare system [7, 8]. Features of the DCE design such as the number of attributes, levels, alternatives and choice tasks [9–11], and the use of choices with utility balance or in the 'magic-p' range [12, 13], can also lead to increased choice difficulty.

If people have difficulty understanding a DCE this may impact the validity of their responses. Difficult questions in surveys can reduce response rates [14] and increase the use

of simplifying strategies [15], opt out selection [16], non-transitivity of choices [17], non-trading [18] and missing responses [19], and result in added variance and measurement error [20]. Each has the potential to reduce the quality or validity of the resulting preference data and therefore compromise results and interpretation of a DCE. An expert stakeholder panel of people doing and using DCEs suggested ensuring participant understanding was both the most desirable and most actionable characteristic of high-scientific-quality preference studies [21].

The concept of understanding relates to participants responding (choosing) rationally within a DCE based on comprehending both the choice context and choice task and being willing to participate. We conceptualise understanding as an overarching concept that is impacted by the task design and instructions, the attributes and levels and the choice context, as well as the personal characteristics of the respondents and the familiarity or relevance of the topic. There is no consistent or recommended approach to assess participant understanding of DCEs. Options include pretesting and piloting of the attributes and levels during the design phases [22, 23], including rationality or validity checks such as duplicate or dominated choice sets within the design [1, 24–27], using econometric approaches to assess rationality [28], or collecting self-reported participant understanding of the DCE [25]. However, most of these assessments are designed to assess a single component or indicator of understanding (such as consistent choices) rather than overall understanding in terms of choice context, choice task and/or task engagement.

There is limited previous research about whether and how people understand health DCEs and, more broadly, how 'understanding' is conceptualised. Ryan and Gerard [29] found "there did appear to be a significant minority who felt they had some difficulty, but given the nature of the question it was not clear whether this arose from the choosing per se or because the instructions were unclear", recommending further research around how comprehension is impacted by the number of attributes, number of levels, and presentation format. Others describe respondent understanding in terms of participant comprehension and minimising choice task ambiguity, with rigorous development processes and effective communication as strategies to achieve understanding [21]. A recent review considered pre-choice processes of decision making rather than broader understanding [30], describing methods such as eye-tracking, brain imaging, time to complete and think-aloud interviews to investigate the processes individuals follow in completing a choice task. The aim of this review was therefore to identify how participant understanding is conceptualised in the health DCE literature, including how understanding is defined, measured, and used. This can inform future research to improve ease of understanding of DCE choice tasks, approaches to maximise

validity of DCE findings, and how to appropriately interpret DCE results for use in changing practice and policy.

2 Methods

A scoping review was conducted following the methods of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) guidelines [31, 32] to capture the scope of the literature and conceptualisation of understanding in health DCEs [33]. As a scoping review, registration in PROSPERO was not applicable.

2.1 Review Questions

How is participant ‘understanding’ (1) defined, (2) measured, and (3) used within the health DCE literature?

2.2 Eligibility Criteria

Eligible papers included peer-reviewed journal articles in the field of health, related to a DCE or BWS3 (best–worst scaling case 3 studies, an extension of DCE where respondents are presented with three or more profiles and instead of choosing the single profile they most prefer, they choose both the best and the worst profiles [34]), and including some consideration or assessment of participant understanding. The concept of understanding was deliberately kept broad using words such as comprehension and misunderstanding to allow the full scope of how understanding is conceptualised in the health DCE literature to be captured.

Conference abstracts, review articles or non-English papers were excluded. Additionally, papers describing how preferences or choice responses differed by subgroups such as sex or age were excluded unless they assessed differences by level of understanding.

2.3 Information Sources, Search Strategy and Selection of Sources of Evidence

Searches were conducted in June 2019 in the MEDLINE, EMBASE, PsychInfo and Econlit databases, consistent with other recent systematic reviews of DCEs [35, 36]. Search terms were based on previous systematic reviews of DCEs [1, 35–37]. Keywords for the concept of understanding were used in a proximity-based search with keywords for patient, participant and respondent. The full search strategy for each database is presented in ‘Supplementary Material A: Database-Specific Search Terms’. Hand-searching reference lists of identified papers, a recent systematic review of DCEs in healthcare [1], and personal collections of authors was also conducted.

A single reviewer reviewed the title and abstract for each paper identified. All authors participated in the full-text review of any studies that appeared eligible, or which could not be identified as ineligible from the title and abstract, with each paper reviewed by two reviewers and uncertainty discussed with an additional reviewer to achieve consensus.

2.4 Data Extraction and Synthesis

A descriptive analytical approach was used to summarise the relevant data from each eligible paper (and online appendices). Data extraction was initially developed and trialled on six studies and circulated to all authors to ensure all relevant information was captured. Final data extraction included publication year, country, clinical area, subject group (e.g. general public, patients, or health professionals), sample size, study design (e.g. online survey, qualitative interviews), DCE or BWS3, number of attributes, the range of number of levels, number of choice sets, definition of understanding, how understanding was tested, results of the understanding tests, and how the understanding information was used. Each study was also categorised based on how understanding was conceptualised and used within the study. Consistent with the best practice guidelines for scoping reviews [32], methodological quality and risk of bias assessment were not undertaken for this study.

3 Results

The search strategy identified 306 unique articles for potential inclusion. Review of titles and abstracts excluded 31 papers and full-text review excluded a further 200 papers, resulting in 75 eligible studies (Fig. 1).

Tables 1, 2 and 3 show the characteristics of the eligible studies, while Tables 4, 5 and 6 show how understanding was defined, measured, and used within each eligible study. The majority of studies were DCEs (or studies about DCEs), with two BWS3 studies [38, 39] and three studies including both a DCE and a BWS3 design [40–42]. Most were conducted in Europe ($n = 24$) and the UK ($n = 17$), followed by the US ($n = 15$) and Australia ($n = 10$). Over half of the studies ($n = 42$) had a population of patients or carers, and 31 studies included the general population or target population (e.g. for prevention studies, the general population is the target population). Two studies examined clinician preferences [39, 43]. The mean sample size was 421 (range 18–4287, median 231). In the DCE studies, the number of attributes ranged from 3 to 12 (mean = 6) and the number of levels ranged from 2 to 6. On average, participants faced 11 choice sets (range 5–27).

Studies were categorised as (1) applied DCEs ($n = 52$), which included an assessment of understanding in the

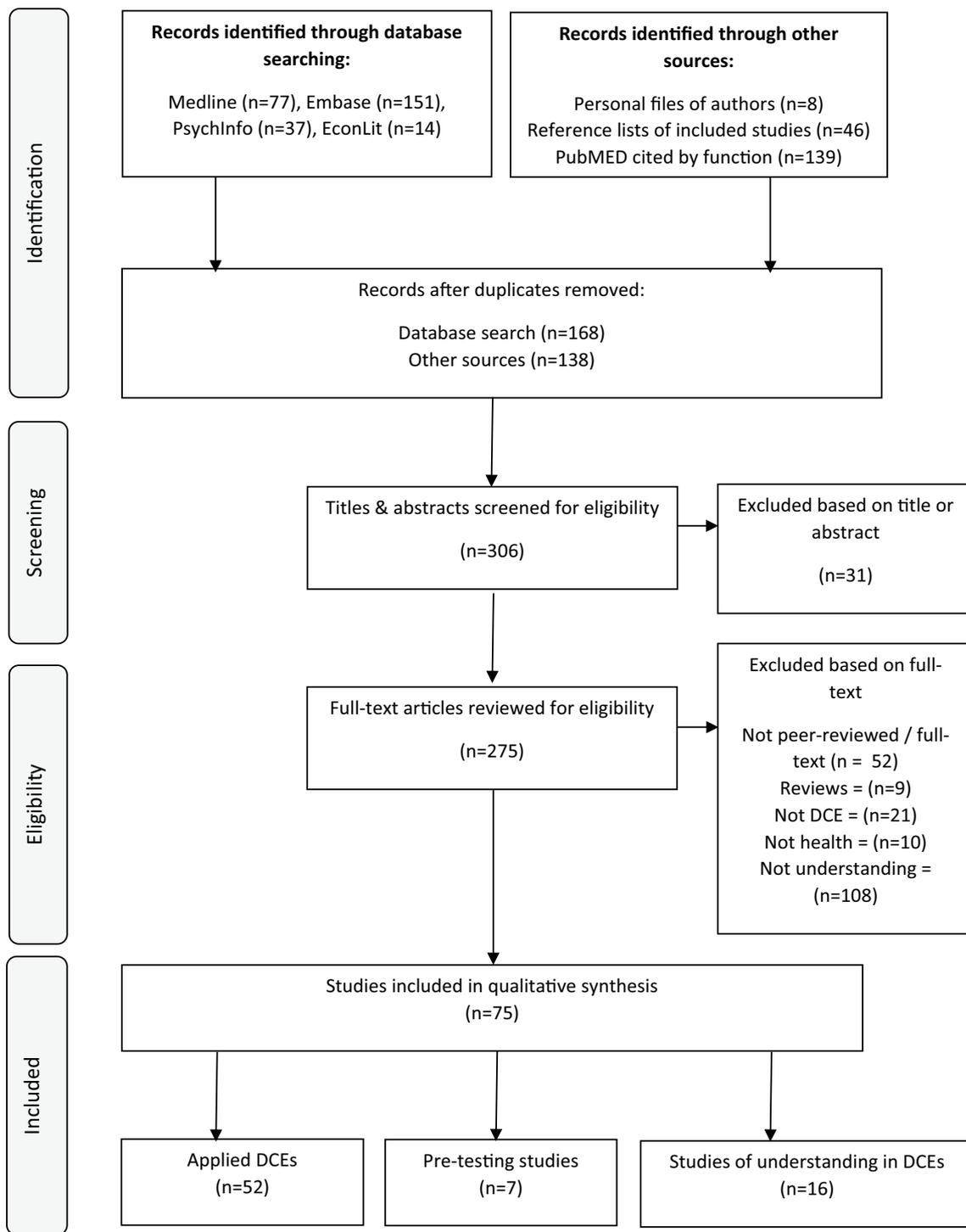


Fig. 1 PRISMA diagram for the scoping review, detailing the database searches and other methods of identifying articles, the number of abstracts screened, and full-text publications reviewed. *DCE* dis-

crete choice experiment, *PRISMA* Preferred Reporting Items for Systematic Reviews and Meta-Analyses

pretesting, analysis, or discussion; (2) pretesting studies ($n = 7$), which were stand-alone studies conducted prior to a DCE, designed to identify relevant attributes and levels of pilot DCE instruments; and (3) studies of understanding

($n = 16$), which were a mixture of qualitative studies and survey methods papers. Often these studies of understanding were a substudy to an applied DCE and aimed to assess the concept of understanding within the context of a choice

Table 1 Characteristics of publications included in the scoping review, by category: applied DCEs

First author, year	Country	DCE/BWS3	Clinical area	Subject group	Sample size	Main study design	Number of attributes	Number of levels	Number of choice sets
Adam, 2019 [82]	Europe	DCE	Complementary medicine	General (or target) population	263	Online survey	6	3	6
Becker, 2018 [79]	UK	DCE	Breastfeeding	General (or target) population	564	Paper-based survey	4	3–6	8
Bottomley, 2017 [57]	UK	DCE	Multiple sclerosis	Patients	350	Online survey	7	3–5	12
Bridges, 2012 [53]	UK	DCE	Non-small cell lung cancer	Patients	89 DCE, 6 pretesting interviews	Online survey	8	3	12
Brown, 2011 [38]	USA	BWS3	Haemophilia	Patients	53	Paper-based survey	12	2–4	12
Buchanan, 2016 [80]	UK	DCE	Genetic testing	Patients	219	Paper-based survey	6	4	16
Byun, 2016 [70]	Asia	DCE	Arthritis	Patients	100 patients, 60 doctors	Face-to-face survey	4	3	18
Cernauskas, 2018 [88]	India	DCE	Health care provider	General (or target) population	100	Face-to-face survey	5	2–3	8
Damen, 2011 [69]	Europe	DCE	Breast reconstruction	Patients	272	Paper-based survey	6	3	10
de Bekker Grob, 2013 [81]	Europe	DCE	Prostate-specific antigen testing	General (or target) population	459	Paper-based survey	5	4	16
de Freitas, 2019 [52]	Europe	DCE	Prostate cancer	Patients	152 DCE, 18 pretest interviews	Online survey	6	2–4	16
de Vries, 2015 [73]	Europe	DCE	Blood pressure	Patients	161	Paper-based survey	6	3–4	10
Fifer, 2018 [89]	Australia	DCE	Type II diabetes	Patients	171	Online survey	10	2–4	8
Gregor, 2018 [62]	Canada	DCE	Inflammatory bowel disease	Patients	586	Online survey	12	2–4	NS
Hauber, 2016 [44]	USA	DCE	Diabetes	Patients	1791	Online survey	5	2–3	10
Hauber, 2017 [48]	USA	DCE	End-stage renal disease	Patients	200	Online survey	7	NS	8
Heringa, 2018 [59]	Europe	DCE	Pharmacy	Patients	476	Online survey	5	2	12
Hofman, 2014 [75]	Europe	DCE	HPV vaccination	General (or target) population	302	Paper-based survey	4	3	9
Hofman, 2014 [76]	Europe	DCE	HPV vaccination	General (or target) population	500	Paper-based survey	4	3	16
Hol, 2010 [78]	Europe	DCE	Colorectal cancer screening	General (or target) population	1034	Online survey	3	3	12
Ivanova, 2019 [64]	USA	DCE	Soft tissue sarcoma	Patients	76 patients, 160 oncologists	Online survey	5	3	12
Jan, 2000 [71]	Australia	DCE	Public health research	General (or target) population	231	Paper-based survey	6	2–3	9

Table 1 (continued)

First author, year	Country	DCE/BWS3	Clinical area	Subject group	Sample size	Main study design	Number of attributes	Number of levels	Number of choice sets
Kistler, 2015 [74]	USA	DCE	Cancer	General (or target) population	116	Face-to-face survey	4	3	10
Laba, 2012 [77]	Australia	DCE	Medication non-adherence	General (or target) population	161	Online survey	8	2–4	10
Lee, 2008 [39]	USA	BWS3	Haemophilia	Doctors	30	Face-to-face survey	12	2–4	12
Liu, 2018 [67]	USA	DCE	GP appointments	Patients	132, 3 unknown	Online survey	5	2–5	8–10
Lloyd, 2011 [68]	UK	DCE	Diabetes	Patients	252	Paper-based survey	8	3	27
Lokkerbol, 2019 [91]	Europe	DCE	Anxiety	Patients	126	Online survey	4	3	12
Mansfield, 2016 [54]	USA	DCE	Renal cell carcinoma	Patients	378	Online survey	5	3	NS
Mansfield, 2017 [66]	Europe	DCE	Type II diabetes	Patients	474 Germany, 401 Spain	Online survey	7	2–4	8
Marshall, 2007 [63]	Canada	DCE	Colorectal cancer screening	General (or target) population	547	Paper-based survey	5	2–4	12
Marshall, 2018 [49]	Canada	DCE	Total joint replacement	Patients	422	Paper-based survey	5	2–4	12
Meads, 2017 [93]	UK	DCE	Cancer pain	Patients	248	Paper-based survey	7	2–3	6
Mohamed, 2015 [50]	Europe	DCE	Thyroid cancer	Patients	134	Online survey	4	3–4	12
Muhlbacher, 2011 [43]	Europe	DCE	Multiple myeloma	Doctors	243	Online survey	8	2–3	8
Muhlbacher, 2015 [56]	Europe	DCE	Chronic pain	Patients	1324 DCE, 24 pretest interviews	Paper-based survey	7	2	8
Muhlbacher, 2015 [86]	Europe	DCE	Acute coronary syndrome	Patients	68	Face-to-face survey	5	3	13
Naik-Panvelkar, 2012 [61]	Australia	DCE	Asthma services	Patients	80	Paper-based survey	8	2–3	9
Naik-Panvelkar, 2012 [85]	Australia	DCE	Asthma	Patients	80	Paper-based survey	8	2–3	9
Naunheim, 2017 [90]	USA	DCE	Ear, nose and throat clinic	General (or target) population	161	Online survey	5	3–5	14
Naunheim, 2018 [87]	USA	DCE	Subglottic stenosis	General (or target) population	162	Online survey	5	2–4	14
Qin, 2017 [47]	Europe	DCE	Type II diabetes	Patients	510 DCE, 50 interviews	Online survey	8	2–4	16
Schmidt, 2017 [55]	Europe	DCE	Lung and colon cancer	Patients	310	Online survey	5	3	10
Tada, 2019 [51]	Asia	DCE	Psoriasis	Patients	395	Online survey	6	2–4	16
Tinelli, 2012 [72]	UK	DCE	Cancer	Patients	183	Paper-based survey	4	3–5	16

Table 1 (continued)

First author, year	Country	DCE/BWS3	Clinical area	Subject group	Sample size	Main study design	Number of attributes	Number of levels	Number of choice sets
Veldwijk, 2014 [58]	Europe	DCE	Diabetes	Patients	781	Paper-based survey	5	3	9
Vennedey, 2016 [60]	Europe	DCE	Macular degeneration	Patients	86	Face-to-face survey	5	2–4	12
Whitaker, 2017 [83]	UK	DCE	Primary care	General (or target) population	601	Online survey	5	2–4	8
Whitty, 2013 [92]	Australia	DCE	Chronic heart failure	Patients	91	Face-to-face survey	5	2	5
Wong, 2014 [46]	Australia	DCE	Cancer	Patients	320 (yet to be recruited)	Protocol	6	3	8
Youssef, 2016 [45]	UK	DCE	HIV health services	Patients	1000 (yet to be recruited)	Protocol	7	2–4	12
Zanolini, 2018 [65]	Africa	DCE	HIV self-testing	General (or target) population	1617	Face-to-face survey	3	2–4	9

BWS3 best–worst scaling type 3, DCE discrete choice experiment, GP general practitioner, HIV human immunodeficiency virus, HPV human papillomavirus, NS not specified

experiment. Many focused on the factors influencing understanding or completion of a DCE.

3.1 Applied Discrete Choice Experiments (DCEs)

Fifty-two applied DCEs were identified. Most were conducted with patients ($n = 35$) and sample sizes ranged from 30 [39] to 1791 [44]. The most common clinical areas were diabetes and cancer. Twenty-four of the DCEs were conducted online, 18 were paper-based, and 8 were face-to-face interviews. An additional two DCEs were protocol papers describing DCEs yet to be conducted [45, 46].

Understanding was assessed in three ways: using pretesting to test or confirm participant understanding ($n = 18$); including tests of understanding within the DCE survey, such as a repeated or dominated choice set ($n = 22$); and referring to the concept of understanding in the discussion ($n = 4$). There were eight additional studies that addressed understanding in both the pretesting phase and with tests in the DCE survey.

3.1.1 Applied DCEs Addressing Understanding in the Pretesting Phase

The 18 applied DCEs addressing understanding in the pretesting phase all used the concept of understanding as a way to modify or revise the DCE instrument before final roll out [44–61]. Most used pretesting to test or confirm understanding in relation to the terminology of attributes and medical concepts. For example, "... comprehension

and relevance of the attributes ... as well as the survey instructions" [52] and "wording used in the questionnaire was correct and understood by the target population" [58]. However, many also took a broader view of understanding and used the pretesting process to assess understanding of the choice task; for example, whether participants "understood the attribute definitions, accepted the hypothetical context of the survey, and were able to complete the choice questions as instructed" [50]. Similarly, Bridges et al. [53] completed six pretesting interviews to assess whether participants "understood the survey and were willing to trade off among the attributes and levels". Despite each of the studies in some way describing the use of pretesting to assess or confirm understanding, nearly one-third ($n = 7$) did not report the results of their assessment. Others gave minimal information, such as "minor changes were made to the wording to improve respondent comprehension" [50].

All except one study implied that understanding was established to be good enough to continue with the DCE. The exception was a protocol paper that described focus groups and 1:1 interviews with cancer survivors and oncology health professionals to develop a DCE about preferences for cancer care among people with cancer [46]. They found "Half of the participants felt that the questionnaire was confusing and difficult to interpret, which suggested that the questionnaire might be particularly burdensome and cognitively demanding to our group of patients", and subsequently made changes to the text and layout to improve understanding [46].

Table 2 Characteristics of publications included in the scoping review, by category: pretesting studies

First author, year	Country	DCE/BWS3	Clinical area	Subject group	Sample size	Main study design	Number of attributes	Number of levels	Number of choice sets
Abdel-All, 2019 [96]	India	DCE	Health workforce retention	General (or target) population	20	Face-to-face interviews and focus groups	5	NS	NS
Coast, 2007 [98]	UK	DCE	Dermatology	Patients	19	Qualitative interviews	4	2–4	NA
Helter, 2016 [97]	NS	DCE	Alcohol misuse	General (or target) population	12 expert interviews, 5 focus groups	Literature review and qualitative focus groups	6–9	NS	NS
Janssen, 2016 [40]	USA	DCE and BWS3	Diabetes	Patients	25	Literature review, stakeholder engagement and qualitative interviews	6	3	NS
Katz, 2018 [95]	USA	DCE	Tobacco treatment	Patients	48	Qualitative face-to-face interviews	9	24	15
Klojgaard, 2012 [94]	Denmark	DCE	Spinal disease	Patients and doctors	2 clinician interviews, 3 patient interviews	Stepwise series of qualitative approaches to aid the design of a DCE	NS	NS	NS
McGrady 2018 [99]	USA	DCE	Cancer	Patients	20	Literature review, cognitive interviews and pretesting	4	3	9

BWS3 best–worst scaling type 3, DCE discrete choice experiment, NS not specified, NA not applicable

3.1.2 Applied DCEs Including a Test of Understanding Within the DCE Task

The 22 applied DCEs that included a test of understanding within the DCE task were spread across year, country, clinical area, sample size and survey administration, and the definition of understanding was similarly varied [62–83]. While many studies referred to relatively generic concepts of understanding, such as the task or attributes and levels [e.g. 82], other studies took a more targeted approach and related the concept of understanding back to the underlying assumptions of rational decision making, referring to concepts of consistency [62–64, 77], dominance or non-satiation [67] and rationality [77, 78]. Some studies referred to engagement with the task as an indicator of understanding [70, 71, 79], while others referred to participants understanding the concept of making a choice [68, 72].

Despite these different definitions of understanding, methods to assess understanding were relatively consistent. Over half used a dominated choice set [63, 64, 68–71, 73–78, 82], with other approaches including a repeated choice set [64, 77, 79], analysis of dominant preferences [63, 67], inclusion of debriefing questions [65, 76, 80, 83], inclusion of quiz questions [65, 66], inclusion of a numeracy scale [66, 75] and use of the root likelihood approach in Sawtooth [62]. Fourteen studies used the understanding results to exclude participants from the analysis [62–64, 67, 68, 70, 71, 73, 74, 77–80, 82], two used them to describe the sample [65, 69], one included them only in the discussion section [72], and the remaining did not describe how they used the results of their understanding assessment [66, 75, 76, 81, 83].

Where reported, rates of passing the various tests of understanding varied and were spread across the studies that excluded participants based on the results of these

Table 3 Characteristics of publications included in the scoping review, by category: studies of understanding in DCEs

First author, year	Country	DCE/BWS3	Clinical area	Subject group	Sample size	Main study design	Number of attributes	Number of levels	Number of choice sets
Bech, 2011 [10]	Denmark	DCE	Dental care	General (or target) population	1053	Online survey with substudy examining the impact of number of choice sets on choice difficulty	6	2–4	5, 9, or 17
Cheraghi-Sohi, 2007 [103]	UK	DCE	Primary care	General (or target) population	20	Qualitative interviews using the think-aloud technique	6	2–4	8
Kenny, 2003 [100]	Australia	DCE	Varicella vaccination	General (or target) population	34	Qualitative telephone interviews after completion of a DCE	7	NS	16
Kenny, 2017 [106]	Australia	DCE	Primary care	General (or target) population	1208	Online survey with participants randomised to two different presentation formats	10	3	18
Maddala, 2003 [107]	USA	DCE	HIV testing	General (or target) population	353	Paper-based survey with alternative presentations	6	2–4	11
Ryan, 2009 [27]	UK	DCE	Bowel cancer screening	General (or target) population	18	Qualitative face-to-face think-aloud interviews	5	3	10
Ryan, 2010 [108]	UK	DCE	Rheumatology	Patients	189	Paper-based survey with substudy examining rationality	6	2–4	8
San Miguel, 2005 [101]	UK	DCE	Primary care	Patients	1343	Paper-based survey with substudy examining factors influencing rationality	5	NS	6, 8, or 10
Scott, 2002 [18]	UK	DCE	Primary care	General (or target) population	4287	Substudy examining frequency of, and factors relating to, dominant preferences	4	2–4	15
Severin, 2013 [42]	Europe	DCE and BWS3	Genetic testing	General (or target) population	31 DCE, 26 BWS3	Qualitative interviews comparing completion of a DCE and BWS3	6	2–3	12

Table 3 (continued)

First author, year	Country	DCE/BWS3	Clinical area	Subject group	Sample size	Main study design	Number of attributes	Number of levels	Number of choice sets
Skedgel, 2013 [109]	UK	DCE	Healthcare priority setting	General (or target) population	604	Online survey comparing DCE vs. constant-sum paired comparison for societal preferences	6	3	10
Spinks, 2016 [9]	Australia	DCE	Complementary medicine	General (or target) population	32	Eye-tracking laboratory study exploring how complexity impacts attribute non-attendance	3–8	2–4	8
Van Dijk, 2016 [41]	USA	DCE and BWS3	Hip replacement	General (or target) population	429	Online survey comparing preferences obtained from a DCE and BWS3 simultaneously	5	4	8 DCE or 12 BWS3
Vass, 2019 [105]	UK	DCE	Breast cancer screening	General (or target) population	19	Qualitative think-aloud interviews	3	NS	5
Veldwijk, 2015 [104]	Europe	DCE	Rotavirus vaccination	General (or target) population	959	Online survey with two versions	5	3	9
Veldwijk, 2016 [102]	Europe	DCE	Rotavirus vaccination and prostate cancer screening	General (or target) population	70	Face-to-face or telephone qualitative interviews after completion of a DCE	5	2–4	9 or 16

BWS3 best–worst scaling type 3, DCE discrete choice experiment, HIV human immunodeficiency virus, NS not specified

Table 4 Conceptualisation of understanding in the publications included in the scoping review, by category: applied DCEs

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Adam, 2019 [82]	"Patients' understanding of the attributes and levels"	To exclude participants	"The questionnaire contained one choice task with a dominated alternative"	"Five were defined as 'irrational' responses. The 'irrational' responses were given by patients from the GM group. These respondents were mostly male (3 men, 1 woman, and 1 'not stated'), with a mean age of 48.8 years (SD 16.8) and a mean duration of education of 11.6 years (SD 1.1). Three of these patients reported having a chronic condition" Two questionnaires were excluded
Becker, 2018 [79]	"Respondents had read, understood and engaged constructively with the choice task"	To exclude participants	"Validity checks ... the chosen options were found to be the same across all choice tasks indicating that respondents had not necessarily considered trade-offs between the presented scenarios"	
Bottomley, 2017 [57]	"Participants had little trouble in understanding the DCE task"	To revise DCE instrument	"One-on-one interviews" method to elicit understanding not reported	"Participants had little trouble in understanding the DCE task"
Bridges, 2012 [53]	"Understood the survey and were willing to trade off among the attributes and levels"	To revise DCE instrument	"Guidelines for patient-reported assessment of outcomes were used to facilitate comprehension by respondents, which was tested using open-ended interviews"	"Patients from the UK diagnosed with NSCLC who confirmed that they understood the survey and were willing to trade-off among the attributes and levels"
Brown, 2011 [38]	"Complexity of the choice task exercises ... length of the instrument"	In the discussion, to justify choice of DCE instrument	Not tested	"Lee et al. reported no respondent complaints about the complexity of the choice task exercise (using a similar exercise) or length of the instrument"
Buchanan, 2016 [80]	"Difficulty, indicating ... understood the choice questions"	To exclude participants	Self-reported difficulty of the DCE task	"Most respondents (97%) passed the rationality check, with 9% rating the DCE as difficult"
Byun, 2016 [70]	Patients and physicians gave their attention to the questionnaire	To exclude participants	Dominant choice set	"98.8% of respondents rationally preferred the dominant COX-2 inhibitor when informed regarding the trade-offs between the benefit-risk attributes, meaning that patients and physicians gave their attention to the questionnaire"

Table 4 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Cernauskas, 2018 [88]	"Understood the task and made consistent choices"	To revise DCE instrument	Pilot testing and repeated choice set	"Pretest showed that all respondents understood the task and provided valid answers, but it also led to changing the wording of several attributes to eliminate any ambiguity in meaning and enhance respondents' comprehension" AND "seven responses were excluded from the analysis due to the lack of consistency between the first and last choice sets which were identical. This implied these respondents did not perhaps understand the choice task"
Damen, 2011 [69]	"Passed the dominant question, demonstrating their understanding of the DCE task"	To describe sample	"A dominant choice set (i.e. a choice set in which both alternatives used implant material but one was characterised by logically preferable levels on all other attributes) and was included to test for rationality"	"In total, 270 of 272 patients (99%) passed the dominant question, demonstrating their understanding of the DCE task"
de Bekker-Grob, 2013 [81]	"Experienced difficulty of the questionnaire"	Not reported	Not reported	"93% of the respondents passed the rationality test that was included in the questionnaire, and 76% of the respondents did not find the DCE questions difficult"
De Freitas, 2019 [52]	"... comprehension and relevance of the attributes ... as well as the survey instructions"	To revise DCE instrument	"Pretest interviews"	"Participants ... were able to answer all the questions, and ... the instructions were clear and easy to understand. Some participants provided feedback about the wording of some of the attribute levels ...", and "no changes to the number of choice tasks or to the experimental design were made on the basis of the pretest interviews"
De Vries, 2015 [73]	"Understand the task"	To exclude participants	"This dominant choice set was added to identify the responders who may not understand the task"	NS—10 participants were excluded (assumed because they failed the dominant choice set)

Table 4 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Fifer, 2018 [89]	Online pilot with targeted feedback to ensure logic and understanding. Measured: time to complete survey, response patterns, question regarding understanding of the task	To revise the DCE instrument and to exclude participants	Online pilot with targeted feedback to ensure logic and understanding. Measured: time to complete survey, response patterns, question regarding understanding of the task	Pretesting: “No changes were made to the overall survey and DCE as participants reported the attributes were relevant and the survey easy to understand” AND in the discussion—“Literacy and language skills were not tested and participants may not have had an adequate understanding of the choice tasks”. No results given for validity checks and it is not clear if anyone was excluded from the analysis
Gregor, 2018 [62]	“... consistency of responses was verified with root likelihood approach”	To exclude participants	Root likelihood approach (referenced to Sawtooth software)	93 of 744 respondents were excluded from the analysis due to inconsistent survey responses
Hauber, 2016 [44]	“Ability to understand and accept the attributes and levels ... and ... willingness to trade among attributes in choosing between hypothetical device profiles”	To revise the DCE instrument	“The draft survey was pretested using in-person interviews ... asked to complete a paper-and-pencil version of the survey instrument”	“Based on the pretest interviews, the survey instrument was finalised”
Hauber, 2017 [48]	“Clear and comprehensible” survey questions—understanding of definitions and instructions and hypothetical context of survey	To revise the DCE instrument	Qualitative pretests—“Interview participants were asked to think aloud as they completed the draft survey. Participants were also asked a series of debriefing questions to determine whether they understood the definitions and instructions, accepted the hypothetical context of the survey, and successfully completed the choice questions as instructed”	Not specified
Heringa, 2018 [59]	Not specified, but mention health literacy, risk understanding and task understanding in the discussion	Adjust the wording of the DCE	“In the third step, the questionnaire was pretested by patients and pharmacists for understanding, feasibility and wording”	“The questionnaire was adapted according to the feedback in a cyclic process until after three cycles no new issues were identified and good understanding was reached”
Hofman, 2014 [75]	“Understanding of the DCE task” and “to gain more insight into respondents’ understanding of the DCE task, i.e. comparing risks and percentages, we included the Subjective Numeracy Scale (SNS)”	NS	“We assessed respondents’ understanding of the DCE task by including a dominant choice set as a rationality test. To gain more insight into respondents’ understanding of the DCE task, i.e. comparing risks and percentages, we included the Subjective Numeracy Scale”	“The relatively high score for subjective numeracy score indicates that our sample probably did understand the risks and percentages they had to compare in the DCE task”

Table 4 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Hofman, 2014 [76]	“Understanding of the DCE task”	Not specified	“To assess respondents’ understanding of the DCE task we included a dominant choice set as a rationality test. Also, we included four items on a 5-point Likert scale to evaluate whether respondents considered the DCE questions ‘clear-unclear’, ‘difficult-easy’, ‘annoying-pleasant’, and the number of questions as ‘too many-not too many’.”	“The dominant choice set was answered correctly by 490/500 of the respondents ... the mean evaluation of the DCE questions were unclear-clear mean 3.48, difficulty-easy mean 3.53, annoying-pleasant mean 2.82 and ‘too many questions-not too many’ mean 2.56”
Hol, 2010 [78]	“A rationality test was included in the questionnaire to determine the understanding of the questionnaire by each subject”	To exclude participants	Dominant choice task	A significantly higher proportion of the previously screened subjects (91%) passed the rationality test compared with the screening-naïve subjects (82%, $p < 0.001$)
Ivanova, 2019 [64]	“... consistency of individual responses: a test–retest exercise and a dominant alternative test ... [to determine] ... whether respondents sufficiently understood the evaluation task”	To exclude participants	“... to test for consistency of individual responses: a test–retest exercise and a dominant alternative test; the test–retest exercise tested that subjects who preferred Treatment A to Treatment B at one point in the sequence of choice exercises should also prefer Treatment A to Treatment B at any subsequent point. The dominant alternative test determined whether respondents sufficiently understood the evaluation task to indicate a preference for an unambiguously better treatment profile ...”	76 of the 78 patients who completed the DCE passed the dominant alternative test, and 160 of the 206 oncologists who completed the DCE passed the dominant alternative test. Sensitivity analyses for the subpopulation defined by respondents who also passed the test–retest test were consistent with those from the main analyses
Jan, 2000 [71]	“It is important to ensure that individuals being surveyed understand the tasks being presented to them and that they are taking the exercise seriously”	To exclude participants	“Respondents’ understanding was tested by including a choice in which one scenario ‘dominated’ another”	“Two respondents provided inconsistent responses. These were removed from subsequent analysis”
Kistler, 2015 [74]	“Understood the DCE selection process”	To exclude participants	“A participant who was unable to understand that one test was superior or dominant to the other was considered unable to complete the remainder of the choice questions”	“Of 277 potentially eligible participants who were approached ... 9 were unable to complete the dominant choice question”

Table 4 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Laba, 2012 [77]	A repeated choice set and a dominant choice set. "Rather than test for survey understanding, these tests are included to investigate whether responses are 'rational' according to the axioms of preference-based consumer theory"	To exclude participants	"Two additional choice sets were added to each survey version to check for consistency (a repeated choice set) and internal validity (a choice set with a dominated alternative according to a priori expectations)"	"91% passed the monotonicity test generally indicating 'rational' interpretation of the factors as intended ... For the repeated choice task, 77% of the respondents passed and the kappa statistic was 0.48 representing moderate agreement"
Lee, 2008 [39]	Missing data and time to complete, as well as presence of lexicographic preferences (choosing based on one attribute)—described only in the discussion	In the discussion, to confirm choice of DCE instrument	Missing data, time to complete, presence of lexicographic preferences—described only in the discussion	"That all surveyed haematologists completed the self-administered discrete choice exercise without any missing data suggests a clear understanding of the issues involved in this experiment... Furthermore, the patterns of responses revealed no lexicographical issues among respondents"
Liu, 2018 [67]	Dominant preference and non-satiation (assessed through a dominated choice set)	To exclude participants	Tested for dominant preferences (i.e. made a choice based on one attribute only)	Dominant preferences: "16.7% of the subjects showed dominant preferences in Study 1, and this number was 17.0%, 14.7%, and 16.1% in studies 2–4 respectively"
Lloyd, 2011 [68]	"If they did not understand the concept of trading—that is, if they failed a test question that asked them to choose between 2 treatments, of which one was clearly better than the other, and in which, therefore there was no trade-off"	To exclude participants	Dominated choice set	Dominated choice set: "In total, 7.6% of the respondents who faced choice sets with a dominated alternative preferred the dominate one in Study 1, and this percentage was 4.0%, 4.1% and 5.7% in studies 2–4 respectively. These proportions of 'irrational' choices are well within the acceptable standard in a DCE (Johnson, 2007). These percentages are relatively small, so it is unlikely that subjects were confused by the task"
				"Four people failed the test question and were excluded from the WTP analysis"

Table 4 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Lokkerbol, 2019 [91]	<p>Pretesting: “understandability and usefulness of the attributes/levels”</p> <p>DCE: “the difficulty of the choice task”</p>	To revise the DCE instrument and to describe participants	<p>Pretesting: “Pre-test with three patients”—methods for eliciting understanding not described</p> <p>DCE: “Respondents were asked to scale the difficulty of the choice tasks on a five-point Likert scale”</p>	<p>Pretest: “The understandability and usefulness of the attributes/levels was judged to be appropriate”</p> <p>DCE: “Respondents rated the difficulty of the DCE task on a five-point scale with a 2.81 average, which lies between easy (=2) and neutral (=3); 72.2% of the completers rated the questionnaire as not difficult (score ≤ 3)”</p>
Mansfield, 2016 [54]	“The pretest interviews evaluated the survey instrument for patient comprehension and burden”	To revise the DCE instrument	Pretest interviews	“During the pretest interviews, most survey features worked smoothly. In general, respondents understood the attributes and could answer the choice questions. Minor changes to attribute descriptions were made to improve comprehension and readability. However, several difficulties were identified regarding comprehension of the information treatment (specifically, the description of the correlation between efficacy and toxicities). Survey revisions, including reducing the number of potentially correlated attributes to one in each DCE question, were made to help respondents understand the correlation”
Mansfield, 2017 [66]	Understanding of the attributes	To revise the DCE instrument	“Each of the seven attributes was described separately, and the descriptions were followed by questions about the attribute to ensure respondents’ understanding Respondents who answered incorrectly were presented with the correct answer and an explanation before proceeding with the survey, while respondents who answered correctly were presented with the correct answer to reinforce their understanding”	Not specified

Table 4 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Marshall, 2007 [63]	<p>“... consistency of responses to assess the reliability of the responses” ... “consistency tests showed that respondents were considering trade-offs between attributes”</p>	To exclude participants	<p>“... individual responses were assessed for violation of the tenet of monotonic preferences”—assessed using two dominant pair tests. Second, “respondents with preferences whose responses seem to indicate unwillingness to trade a given attribute against other attributes. A respondent was considered unwilling to trade if they made the same choices in at least 8 or 10 scenarios with the attribute” ... “if a respondent failed either of these consistency tests, he was considered inconsistent”</p>	<p>“Coefficient signs were unaffected, and differences in the coefficients were insignificant using Wald tests ($p < 0.05$). Thus, we decided to retain all respondents in the final analysis”</p>
Marshall, 2018 [49]	“Understanding of the questionnaire content”	To revise the DCE instrument	Not specified	“Based on pretesting results, the questionnaire was modified”
Meads, 2017 [93]	“We did not test patient’s understanding of the DCE method, although the survey was carefully explained to individuals”	In the discussion, to interpret results	Not tested	<p>“For a small minority, there was a suggestion of counter-intuitive preferences ... this may be because of a lack of participant understanding or engagement or the result of an artefact of the survey design”</p>
Mohamed, 2015 [50]	“Understood the attribute definitions, accepted the hypothetical context of the survey, and were able to complete the choice questions as instructed”	To revise the DCE instrument	<p>15 face-to-face semi-structured interviews ... during these interviews, patients were asked to ‘think aloud’ as they completed the draft survey instrument and a series of debriefing questions to ascertain that they understood the attribute definitions, accepted the hypothetical context of the survey, and were able to complete the choice questions as instructed”</p>	<p>“Minor changes were made to the wording to improve respondent comprehension”</p>

Table 4 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Muhlbacher, 2011 [43]	DCE: "the degree of difficulty ... in performing the paired comparisons in the DCE"	To revise the DCE instrument and describe the sample	Pretesting: "30 physicians participated in this preliminary survey by filling in the online questionnaire" DCE: "At the end of the questionnaire, the subjects were asked to give an assessment of the degree of difficulty they had"	Pretesting: "The evaluation showed that the questions were clear and comprehensible to the subjects" DCE: "Only 2% found it 'very difficult', 17% said it was 'rather difficult', 46% gave a medium ranking ('some comparisons more difficult than others'), 27% thought it was 'not very difficult', and 8% claimed they had 'no problems at all' in completing the DCE. This means that overall, the DCE can be considered as feasible from the physicians' point of view"
Muhlbacher, 2015 [56]	"Comprehensibility and practicality"	To revise the DCE instrument	"Two free text questions on content validity and comprehensibility (practicality) were included in the questionnaire"	None reported
Muhlbacher, 2015 [86]	Pretesting: clarity of the questionnaire design, the quality of the scales used, and the understandability of the attributes, levels, and trade-offs DCE: "consistency of responses in each set" as well as a repeated choice set and dominant choice set	To exclude participants	Pretest interviews, repeated choice set, dominated choice set	Results of pretesting interviews not presented. 60 respondents were excluded "after internal consistency tests and the exclusion of all invalid questionnaires"
Naik-Panvelkar, 2012 [61]	"Comprehension, interpretation, format and face validity"	To revise the DCE instrument	Pilot testing on a convenience sample	"No major changes were made to the questionnaire following the pilot"
Naik-Panvelkar, 2012 [85]	Main study: "Ease and time of completion of the DCE" Pretesting: "comprehension, interpretation, language, flow and format"	To revise the DCE instrument	Questions about ease and time to complete the DCE	Pretesting—"None of the respondents raised any issues with respect to understanding the questionnaire and so the pilot did not lead to any major modifications in the questionnaire." Main study: The mean time for completion of the questionnaire was 15 minutes." Results of 'ease of completion' questions not reported

Table 4 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Naunheim, 2017 [90]	Pretesting: "questions were well understood" DCE: 'stated comprehension'	To revise the DCE instrument	Pretesting interviews and in the DCE 'included a question on 'stated comprehension'—by asking whether patients understood the survey	"All participants indicated that they understood the format and content of the DCE... the reliability of responses was mixed: in 42.3% of cases, the most important attribute as individually measured with the DCE matched the stated importance as assessed by simple rankings of attributes by the participant"
Naunheim, 2018 [87]	"Understanding of the experimental design ... stated comprehension and consequentiality ... [and] ... monotonic responses"	To revise the DCE instrument	"The survey was initially piloted in 10 patients ... and responses were elicited to assess understanding through both directed questions and open-ended comments" AND "stated comprehension and consequentiality ... and response consistency... were included as measures of internal validity"	Pretesting: "Nine of 10 demonstrated full understanding of the experimental design, and responses were monotonic across attributes (i.e. there was a clear best-to-worst ordinal ranking for levels within each attribute)" DCE: "All participants indicated that they understood the format and content of the DCE" AND "Results for consequentiality and response consistency not reported"
Qin, 2017 [47]	"... wording and comprehension as well as cognitive and overall burden"	To revise the DCE instrument	"... survey was pilot tested with respect to wording and comprehension, as well as cognitive and overall burden, with five patients in the UK"	Not specified
Schmidt, 2017 [55]	"The first choice set enabled us to test patients' understanding of the DCE method because it included a dominant profile" AND "We conducted a pretest to ensure that the final questionnaire could be understood easily by the patients"	To revise the DCE instrument	Pretest interviews, dominated choice set	"The pretest showed that most patients could not answer questions about their disease state or therapy goals. Therefore, this question was excluded from the questionnaire". Results of the dominant choice set not reported (even in the supplementary material)
Tada, 2019 [51]	"The questionnaire's understandability"	To revise the DCE instrument	"A pilot study was executed"	Not specified
Tinelli, 2012 [72]	Understand the concept of choosing between hypothetical situations	In the discussion section	Not specified	"Many found it difficult to understand the concept of choosing between hypothetical situations and hence took up to an hour to complete. The research nurse helped them understand what they had to do"

Table 4 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Veldwijk, 2014 [58]	“wording used in the questionnaire was correct and understood by the target population”	Adjust the wording of the DCE	When sent pilot versions (by post) “respondents were asked to mark every question or answering category that they did not understand or found hard to grasp and they were asked to provide suggestions for improvement”	Not reported
Vennedy, 2016 [60]	“Understanding the task or the attribute and level descriptions”	Adjust the wording of the DCE	“Focus group interviews”	“Two additional pilot tests did not reveal any further problems in understanding the task or the attribute and level descriptions” AND “interviewer assistance provides the opportunity to identify misunderstandings of the task at early stages in questionnaire administration”
Whitaker, 2017 [83]	“Perceived difficulty completing the discrete choice experiment”	Not reported	“All participants were asked to rate how easy/difficult they found the scenarios to complete, and were given the opportunity to provide free-text comments on the survey”	“Most of the 601 participants found the discrete choice scenarios easy ($n = 291$; 48%), or very easy ($n = 240$; 40%), to complete, with a smaller proportion reporting they found the section difficult ($n = 68$; 11%) or very difficult ($n = 2$; 0.3%). The open-text comments also did not reveal any concerns about the discrete choice experiment”
Whitty, 2013 [92]	“Understand the DCE task and the hypothetical concept of paying for a service”	In the discussion, to interpret results	Not tested	Not tested
Wong, 2014 [46]	“Face validity ... comprehension, options and wording of the DCE”	To revise the DCE instrument	“Focus group participants, of whom 10 out of 11 questionnaires were returned”	“Half of the participants felt that the questionnaire was confusing and difficult to interpret, which suggested that the questionnaire might be particularly burdensome and cognitively demanding to our group of patients”
Youssef, 2016 [45]	“The DCE was piloted ... to ascertain patient acceptability and understanding”	To revise the DCE instrument	Pilot testing	Results not reported

Table 4 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Zanolini, 2018 [65]	“Asked how confident they were that they understood how to do the test”	To describe the sample	“Asked how confident they were that they understood how to do the test” ... “a 9-item questionnaire was used to objectively assess their understanding”	“Responses to the knowledge assessment questions also indicated good understanding of HIVST (median score 8 out of 9 correct, range 2–9)” ... “supplementing the instruction sheet with a brief video increased participant confidence slightly, but did not improve (and slightly worsened) understanding of how to perform the test”

BWS3 best-worst scaling type 3, DCE discrete choice experiment, GM general medicine, HIVST human immunodeficiency virus self testing, SD standard deviation, NSCLC non-small cell lung cancer, COX cyclooxygenase, NS not specified, WTP willingness to pay

tests and those that did not. Many studies reported rates of < 5% failing their tests [64, 68–71, 74, 76, 79, 80, 82]; however, some were much higher, for example, 12.5% provided inconsistent responses [62], up to 7.6% failed a dominated choice set, and up to 17% displayed dominant preferences [67], while 18% failed a rationality test [78] and 33% failed a repeated choice task [77]. Liu et al. [67] stated “These proportions of irrational choices are well within the acceptable standard in a DCE” and cited Johnson et al. [84], although it is not clear where in this work a threshold for irrational choice prevalence is given. Other studies simply stated when testing the impact of including or excluding those with inconsistent/irrational responses, there were no significant differences in the results [63, 64]. Some studies gave only descriptive summaries of the results, such as “Responses to the knowledge assessment questions also indicated good understanding” [65] and “Many found it difficult to understand the concept of choosing between hypothetical situations and hence took up to an hour to complete” [72].

Two studies compared irrational responses by sociodemographic characteristics [78, 82]. Adam et al. [82] found that the five participants (of 263) who gave irrational responses were more likely to be male, older, and less educated than the average respondent in their DCE [82]. Hol et al. [78] identified that participants who had previous experience of the health test under study (colorectal cancer screening) were significantly more likely to pass the rationality test than previously unscreened participants.

In studies where debriefing questions were used, the rates of people reporting the survey difficult to complete were higher than the proportion who failed the rationality test. For example, one study reported 3% failed, while 9% rated the survey as difficult [80], and, in another study, 7% failed and 24% reported the survey to be difficult [81]. In a study that only included debriefing questions, 12% of respondents reported they found the scenarios difficult to complete [83]. When numeracy was tested rather than rationality, it was reported that “The relatively high score for subjective numeracy score indicates that our sample probably did understand the risks and percentages they had to compare in the DCE task” [75].

Eight additional studies assessed understanding in both pretesting and within the DCE survey [43, 85–91]. These studies were similar to the other reported studies, with pretesting results used to refine the wording and formatting of the surveys, and the rationality tests using dominant or repeated choice tests and debriefing questions. Insufficient details were available to assess whether the rates of people failing the rationality/consistency tests were lower when pretesting for understanding had also been conducted.

Table 5 Conceptualisation of understanding in the publications included in the scoping review, by category: pretesting studies

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Abdel-All, 2019 [96]	“Comprehension (ability to understand the question as intended), retrieval of information (thinking about the question and drawing conclusions), judgement and selection of response to the question”	To revise DCE instrument	Cognitive ‘think aloud’ interviews and focus group discussions	“The DCE was well received by the ASHAs and ... they did not find it difficult to understand the choice sets presented to them. The introductory statement with the choice set example helped to explain the hypothetical nature of the DCE and confirmed the cognitive understanding of the ASHAs of the DCE” Not tested
Coast, 2007 [98]	Understanding of the technical concepts in the scenario and “understanding of specific terms”	To revise DCE instrument	Not tested—qualitative interviews to elicit participant definitions of different concepts	Not tested
Helter, 2016 [97]	“The final step in the process of attribute development aims to ensure that the desired meaning is evoked and that the terminology is understandable for respondents”	To revise DCE instrument	“The literature suggests qualitative techniques (part of pretesting and piloting, cognitive interviews, think-aloud technique), or researchers judgement based on data already available”	Not reported
Janssen 2016 [40]	Understanding of the attributes in terms of names and descriptions	To revise DCE instrument	Not tested—qualitative interviews to elicit participant definitions of different concepts	“Names and descriptions of the attributes were adapted throughout the pretesting process to maximise participants’ understanding”
Katz, 2018 [95]	“Two qualitative analysts analysed the think aloud data to evaluate subjects’ overall comprehension of, and any difficulties in completing, the choice tasks.” Also assessed monotonicity (dominant choice set) and stability (duplicate choice set), and non-continuity	To revise DCE instrument	Cognitive ‘think aloud’ interviews	“Most subjects completed the DCE questionnaire without difficulty; however, not all subjects were fully engaged, and in some cases the interviewer documented signs of cognitive overload or exhaustion. In a few cases, subjects struggled with the idea of choosing between two hypothetical medications.” ... “Most subjects showed good comprehension of the choice tasks and clearly considered two or more attributes in choosing between medications”

Table 5 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Klojgaard, 2012 [94]	“Understand the content of the attribute and in a clear and concise manner”, especially risk-based attributes	To revise DCE instrument	During patient interviews, included discussion of ‘Framing: Perception and understandability of the given text and explanations’ and ‘... complexity and overall experience with filling out the pilot’	“... it became very clear that the included 6 attributes were too many. The respondents had problems remembering the first attribute when looking at the later attributes, and two respondents ended up answering lexicographically or using heuristics with emphasis on the last or the two last attributes, despite not ranking these the highest.” ... “Furthermore, two respondents read the word ‘ability’ as ‘aim’ because these two words are close in Danish. This fact proved to make the attributes difficult to understand and pointed to the need for reframing them.” ... “None of the respondents had difficulty understanding or interpreting the [risk] attribute” ... “It was also clear that the number of assigned levels should be limited because the cognitive ability of the respondents was limited, and the assignment of making choices had already proven to be difficult to grasp.” ... “They were also able to explain the task ahead from reading the text only. Nevertheless, none of the respondents were able to perform the task when it was given to them, and they needed much more explanation and help to get started”
McGrady, 2018 [99]	“Meaning of each attribute” and “participant comprehension”	Adjust the wording of the DCE	Looked for ‘clarifying questions’ during DCE administration in the cognitive interviews	Participants in the final round “did not ask any clarifying questions during DCE administration. Of note, all Wave II participants also described the task as ‘easy’ or ‘pretty easy.’”

ASHA Accredited Social Health Activist (community health worker), *BWS3* best–worst scaling type 3, *DCE* discrete choice experiment

Table 6 Conceptualisation of understanding in the publications included in the scoping review, by category: studies of understanding in DCEs

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Bech, 2011 [10]	“Certainty about choices” and “perceived difficulty of the DCE questions”	As an explanatory factor for DCE completion	Self-reported certainty of answers and perceived difficulty of the DCE questions	“About two-thirds of the respondents in each group reported that they were very certain of their answers to the DCE questions. This result is very satisfactory and suggests that respondents had a good general understanding of the choice task. With respect to choice difficulty, significantly more of the respondents presented with 5 choice sets stated that the choice tasks were very easy compared to the respondents presented with 9 and 17 choice sets ... Overall, however, it would seem that self-assessed choice complexity is only weakly related to the number of choices”
Cheraghi-Sohi, 2007 [103]	“‘Rationality’ (i.e. respondents who behave in line with the axioms of economic theory) ... They should also exhibit ‘compensatory decision-making’ (i.e. respondents are willing to trade decreases in one attribute for increases in another). However, there is evidence that individuals use simplifying heuristics when faced with complex decisions ... It is not clear whether behaviour which does not correspond to the assumptions underlying economic theory reflects limitations in the decision maker or task related demand characteristics”	Verify validity of the DCE approach and revise the DCE instrument	Think aloud during completion of two written DCE questionnaires	Extensive! Covers reinterpretation of attributes, reactions to the cost attribute (from dismissing its relevance to it being a dominant attribute), evidence of non-trading, effects of previous experience, and other considerations
Kenny, 2003 [100]	“The interpretation of technical information and understanding the task”	Verify validity of the DCE approach	Semi-structured interview, including prompts/questions about understanding (definitions) of specific attributes, difficulty with the survey and choice strategy	The technical information used to describe the programme attributes appeared to be used appropriately by participants, although their explanations indicated that their understanding did not always come from the questionnaire information. Only one participant appeared to misunderstand the stated preference task, and a small number thought that the complexity and length should be reduced

Table 6 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Kenny, 2017 [106]	“The assessment of accuracy of understanding through the analysis of responses where one alternative is objectively better or equivalent across all attributes”	Verify validity of the DCE approach	Dominated choice sets (x2) and “questions about how easy or difficult it was to understand the information in the choice questions”	“10.4% of respondents chose the worse alternative in at least one of the two dominated choices seen ... the probability of choosing the worse alternative was significantly higher when the presentation format was style 4 (percentage only)” AND “in terms of consumer ranking, style 1 was more frequently identified as the better presentation format for choosing ...” AND “the majority of respondents reported they found the information easy to understand, and this did not differ substantially by presentation style”
Maddala, 2003 [107]	“We examine whether the increased overlap survey was easier by comparing response consistency, dominated responses, perceived difficulty and fatigue between the two surveys”	Verify validity of the DCE approach	Dominated choice set, duplicate choice set, all choices based on one attribute, self-reported rating of difficulty of the survey, compared first half of responses with the second half	“Respondents completing the increased overlap survey did not exhibit significantly fewer inconsistent responses. The increased overlap design did not result in fewer dominant responses. Respondents did not perceive the increased overlap survey as easier. The fatigue effects were similar in both treatments and there was no evidence of a learning effect”
Ryan, 2009 [27]	Understanding of the task, the attributes, and risk information all discussed	Verify validity of the DCE approach	“Qualitative data from the verbal protocol analysis provide rich information about respondent’s preferences” AND “Qualitative data allow the validity of these assumptions to be explored ... this allows us to examine whether respondents made a mistake, misunderstood the attribute, or interpreted the attribute differently from the researchers’ assumptions”	“Respondent 14 ... completed the DCE with ease after some initial misunderstandings that she asked the reviewer to clarify” AND “... respondents make non-monotonic choices because they misunderstand an attributes meaning” AND “Respondent 1 may have failed the quantitative completeness test 1 because he struggled to understand the task initially” AND “there is some evidence that respondents had difficulty understanding risk information”
Ryan, 2010 [108]	Rationality and “misunderstood the DCE”	To exclude participants	“The statistical design resulted in two choices that, on any ‘rational’ choice process, should be preferred to others. Individuals who answered one or both of these choices ‘irrationally’ were assumed either to have misunderstood the questionnaire or not be taking it seriously”	56/189 failed at least one rationality test (41 failed one and 15 failed both)

Table 6 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
San Miguel, 2005 [101]	“Complexity of the choice task faced by individuals, defined by factors such as the number of alternatives, number of attributes, similarity between alternatives etc. Increased complexity may lead individuals to develop simplifying heuristics or to make more errors”	As an explanatory factor in the rationality of DCE completion	Difference in the dispersion of the standard deviation among attribute levels across alternatives, as an indicator of the relative ease of the choice task. In addition, asked participants if it was more difficult to complete questions with three choices compared with two	“Around 18% of respondents found choices with three alternatives more difficult to complete and around 22% said this depended on the choices” AND “as choice sets become less similar, choices become easier” AND “those who thought that the questionnaire was more difficult and those who thought that choices with three alternatives were more difficult were less likely to satisfy the tests”
Scott, 2002 [18]	“Ease of completion and the version of the questionnaire”, also “time to complete”	As an explanatory factor for DCE completion	Unclear—self-reported ease of completion and time for completion	Those who found the questionnaire difficult to complete and who had a university education were less likely to have a dominant preference. Those who found it difficult to complete and were educated to below secondary level were more likely to have a dominant preference. Respondents who filled out questionnaire two and found it difficult to complete were less likely to have a dominant preference. Finally, older respondents who found the questionnaire difficult to complete were less likely to have a dominant preference
Severin, 2013 [42]	“Understanding of the DCE format” and “understanding of the choice format”	Verify validity of the DCE approach	Dominant choice set and self-reported difficulty in understanding and answering the choice formats	“The majority of respondents to the DCE reported that they found the choice tasks not very difficult to understand. However, the respondents reported that the DCE questions were difficult to answer. In contrast, more than half the participants in the BWS exercise found the BWS tasks difficult or very difficult to understand and nearly half of the participants found the questions very difficult to answer
Skedgel, 2013 [109]	“Difficulty understanding the tasks and of answering the tasks”	Verify validity of the DCE approach	“Respondents were asked to rate the ... the difficulty of understanding the tasks and of answering the tasks on 7-point scales ranging from extremely easy to extremely difficult”	“There was no significant difference between the two questionnaires in the proportion that rated the tasks ‘somewhat difficult’ or ‘extremely difficult’ to understand”, overall or within subgroups

Table 6 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Spinks, 2016 [9]	“Processing additional information”	Not specified	Not specified	“In general participants reported being engaged with the survey and although many stated that the choice sets with more information took longer to process, the information itself was not difficult to understand”
Van Dijk, 2016 [41]	“Respondents rated the difficulty of the choice tasks associated with BWS and DCE”	Verify validity of the DCE approach	“Respondents rated the difficulty of the choice tasks associated with BWS and DCE with a number ranging from 1 (very easy) to 5 (very difficult)”	The mean completion difficulty of the DCE was lower than for BWS but not significant. In addition, no significant differences were found between the education groups rating the choice set difficulty
Vass, 2019 [105]	“How respondents complete choice tasks ... [and] ... whether their behaviour aligns with a priori expectations or whether it results in violation of the underlying utility theories”	Verify validity of the DCE approach	Think-aloud interviews to observe decision making in a DCE and in-depth discussion of risk presentations	“The findings of this study suggested most of the women included in the sample were able to trade-off the attributes presented and understand the choice tasks. Although sometimes unwilling to trade-off the cost attribute, this was clearly a preference rather than an expression of cognitive burden or confusion”
Veldwijk, 2015 [104]	“Consistency of choices was used as a proxy for direct attribute level understanding” and “respondents’ level of understanding of the choice tasks in words and graphics, their preference for either presentation of the choice tasks and their interpretation of the words and graphics”	Verify validity of the DCE approach and revise DCE instrument	“Respondents’ level of understanding of the choice tasks in words and graphics, their preference for either presentation of the choice tasks and their interpretation of the words and graphics (i.e. either as a numeric value or as a low-medium-high categorized variable) were measured” AND “Consistency of choices was used as a proxy for direct attribute level understanding”	“These results suggest that the use of words in choice tasks provided more accurate responses than the use of graphics.” Results of questions about ‘level of understanding of the choice task’ and ‘preferences for presentation of the choice tasks’ not presented

Table 6 (continued)

First author, year	Definition of understanding	How they used understanding	How they tested understanding	Results of understanding tests
Veldwijk, 2016 [102]	Understanding the information provided and using complex decision strategies	Verify validity of the DCE approach	Asked in semi-structured interview about: ability to repeat the definition of the risk attributes, mentioning at least three attributes as driving their choices (continuity axiom), use of trading off	“The majority of the participants seemed to have understood the provided information about the choice tasks, the attributes, and the levels. They used complex decision strategies (continuity axiom) and are therefore capable to adequately complete a DCE. However, based on the participants’ age, educational level and health literacy additional, actions should be undertaken to ensure that participants understand the choice tasks and complete the DCE as presumed”

BWS3 best-worst scaling type 3, DCE discrete choice experiment

3.1.3 Applied DCEs that Referred to Understanding in the Discussion

Of the 52 applied DCEs identified, four simply referred to the concept of understanding within the discussion [38, 39, 92, 93]. In these studies, understanding was referred to as an indicator of validity of the results. For example, “That all surveyed haematologists completed the self-administered discrete choice exercise without any missing data suggests a clear understanding of the issues involved in this experiment” [39], and “For a small minority, there was a suggestion of counter-intuitive preferences ... this may be because of a lack of participant understanding or engagement or the result of an artefact of the survey design” [93].

3.2 Pretesting Studies

Seven studies described the process to develop a DCE—either identifying attributes and levels and/or pretesting the draft DCE instrument with an assessment of understanding [40, 94–99]. Most were published after 2015 and half were conducted in the US. DCE topics covered a range of clinical areas, such as diabetes [40], cancer [99], dermatology [98], and workforce retention [96]. All used qualitative interviews [40, 94–96, 98, 99] and/or focus groups [96, 97].

Most studies focused their assessment of understanding on participants’ understanding of the terminology in the attributes, choice sets, or survey. For example, in their focus group discussions, Helter and Boehler [97] defined understanding as “The final step in the process of attribute development aims to ensure that the desired meaning is evoked and that the terminology is understandable for the respondents”, and Coast and Horrocks [98] report testing participants’ ‘understanding of specific terms’. These types of studies usually presented results descriptively, such as “Names and descriptions of the attributes were adapted throughout the pretesting process to maximise participants’ understanding” [40].

However, some studies also included an assessment of understanding of the choice task itself. In a think-aloud study, Katz et al. [95] found “Most subjects ... completed the DCE questionnaire without difficulty; however, not all subjects were fully engaged, and in some cases the interviewer ... documented signs of cognitive overload or exhaustion. In a few cases, subjects struggled with the idea of choosing between two hypothetical medications ... Most subjects showed good comprehension of the choice tasks and clearly considered two or more attributes in choosing between medications”. Similarly, in their cognitive interviews and focus group discussions, Abdel-All et al. [96] examined multiple aspects of understanding, which they describe as “comprehension (ability to understand the questions as intended), retrieval of information (thinking about the question and

drawing conclusions), judgement and selection of response to the question”. They found “The DCE was well received ... and ... they did not find it difficult to understand the choice sets presented to them” [96].

One study did not specifically define understanding, but detailed the issues identified during their patient interviews [94]. These included respondents’ understanding of terminology, i.e. “two respondents read the word ‘ability’ as ‘aim’ because these two words are close in Danish. This fact proved to make the attributes difficult to understand and pointed to the need for reframing them”, and understanding the choice task, i.e. “... it became very clear that the included six attributes were too many ... and two respondents ended up answering lexicographically or using heuristics with emphasis on the last or the last two attributes, despite not ranking these the highest” [94].

3.3 Studies of DCE Understanding

Sixteen studies were specifically about the concept of understanding in the context of DCEs [9, 10, 18, 27, 41, 42, 100–109]. Published between 2002 and 2019, they were primarily from the UK [18, 27, 101, 105, 108, 109] and Europe [10, 42, 102, 104], and were conducted in a range of clinical areas, including vaccination [100, 102, 104], screening [27, 102, 105] and primary care [18, 101, 103, 106]. All were conducted in the general (or target) population, with the exception of the study by Veldwijk et al. [102], who included both the general population and patient groups, and Ryan and Bate [26] and San Miguel et al., [101] who used only a patient sample. The sample sizes ranged from 19 [105] to 4287 [18]. All included a DCE, with two also including a BWS3 task to compare the two approaches [41, 42].

3.3.1 Qualitative Studies of DCE Understanding

There were six qualitative studies in this group [27, 42, 100, 102, 103, 105]. Two were retrospective investigations of understanding in participants who had previously completed a DCE, and aimed to verify the appropriateness of DCEs as a method in their clinical area [100, 102]. They considered understanding to relate to both the information presented, such as definitions of attributes and levels, and the ‘complex decision strategies’ used when the task is understood. Both studies found that participants appeared to understand the technical information within the choice task, and the choice task itself. For example, “The majority of the participants seemed to have understood the provided information about the choice tasks, the attributes, and the levels. They used complex decision strategies (continuity axiom) and are therefore capable to adequately complete a DCE. However, based on the participants’ age, educational level and health literacy additional, actions should be undertaken to ensure

that participants understand the choice tasks and complete the DCE as presumed” [102].

Three papers were prospective think-aloud studies, to assess decision-making processes [103] within a DCE and confirm participants were trading off [105] and choosing rationally [27]. All reported that most, but not all, participants were able to understand the concept of making a choice and make rational trade-offs in line with the assumptions of decision making. For example, one study found some participants reinterpreted the attributes, reacted in different ways to the cost attribute (from dismissing it to considering it a dominant attribute), showed evidence of non-trading, and brought their own experiences to the choice task [103].

The last qualitative study [42] compared the feasibility of DCE and BWS in their clinical context (genetic testing) by interviewing participants about their understanding of the choice format, as well as including a dominant choice set and self-reported difficulty questions. They found respondents completing the BWS task were more likely to report difficulties in both understanding and answering the choice task than those completing the DCE task [42].

3.3.2 Quantitative Studies of DCE Understanding

Among the 16 studies of DCE understanding, 9 were online or paper-based surveys [10, 18, 41, 101, 104, 106–109]. Four of these were designed to compare two different presentation versions of a DCE instrument, such as words or graphics [104], level overlap [107], general presentation format [106], and the impact of number of choice sets on choice difficulty [10]. The other studies examined various assumptions of decision making, such as rationality [101, 108] or dominant preferences [18], or compared a DCE choice task with another choice task—BWS3 [41] or Constant Sum Paired Comparison [109]. These studies measured understanding through consistency of preferences [101, 104, 107], dominated preferences [106, 107], duplicate choice set [107], participant perception of difficulty or format preferences [10, 18, 41, 101, 104, 106, 107, 109], time to complete the survey instrument [18], and survey fatigue [107].

The results were mixed for how task presentation impacts consistency and dominance. While some studies found alternative task presentations could improve consistency [104, 106], others did not [107]. It was noted that participants who reported the task as being difficult were less likely to satisfy tests of rationality such as dominance or consistency [18, 101]. Neither of the studies comparing the ease of DCE completion with other study designs found a significant difference in self-assessed difficulty of the different tasks [41, 109].

Finally, there was one laboratory-based study among the 16 studies of DCE understanding, in which 32 participants undertook a DCE within an eye-tracking system [9]. The

study explored how the number of attributes in the choice scenario impacted attribute non-attendance. Understanding was conceptualised as ‘processing additional information’, and the study found “participants reported being engaged with the survey and although many stated that the choice sets with more information took longer to process, the information itself was not difficult to understand” [9]. However, this paper does not directly link the processing of information or attribute non-attendance with the concept of ‘understanding’ within a DCE.

4 Discussion

This scoping review aimed to identify how the concept of participant understanding is conceptualised in the DCE literature in the health context. We identified 75 DCE and BWS3 studies that included some consideration or assessment of participant understanding. They included 52 applied DCEs, 7 pretesting studies and 16 studies specifically to examine some aspect of DCE understanding.

4.1 How is Participant Understanding Defined Within the Health DCE Literature?

As with many key terms in the DCE literature, there is inconsistency in how participant understanding is defined. It appears that there are two primary ways ‘understanding’ is conceptualised within the health DCE literature. The first is the understanding of the general choice scenario, such as the medical terminology or levels of risk used in the attributes and levels. There has been consistent evidence in the general survey and DCE literature that specialised medical terminology [110] and calculations or presentations of risk and probability [2, 3] are difficult for many people to accurately interpret. This misinterpretation can negatively impact survey responses, including reduced response rates [14] and increased use of simplifying strategies [15], opt out selection [16], non-transitivity of choices [17], and non-trading [18] and missing responses [19]. Our review finds this type of understanding is typically referred to in applied DCE studies and tested using qualitative methods in the pretesting phase of DCE development. However, a growing body of work is examining the influence of alternative presentation formats for concepts and choices, to aid participant understanding and increase the validity of participant responses. Additionally, although not identified in this review, there is also interest in the use of novel presentation formats, such as video [111], to improve the instructions and definitions given to DCE participants. It is possible that having personal experience of the choice scenario may assist with understanding of attributes and complex concepts such as risk, as found in a study of preferences for colorectal cancer screening where

previously screened subjects were more likely to pass the rationality test compared with screening-naïve subjects [78]; however, to date there is insufficient literature to confirm this.

The second way ‘understanding’ is conceptualised within the DCE literature is the need to understand the concept of making a choice. Our review suggests this is seen in the literature as a key component of the rational decision-making process. This is reliant on participants feeling like the choice is realistic and one that they are ‘qualified’ to make [112], and although it can also be linked to the clarity of instructions and formatting within the survey instrument, it is more often related to the face validity of the decision-making process. Our review found that while this aspect of understanding is addressed in applied DCEs using tests such as duplicate or dominant choice sets, or studied specifically through both qualitative and quantitative methods, there is little consensus of the best way to measure or maximise this component of understanding. Assumptions that participants with good numeracy or health literacy will understand a DCE task also require further investigation.

Rarely seen in the literature in this review was the concept of engagement as a key component to understanding. Both understanding the choice scenario and understanding the concept of making a choice are reliant on participants being willing to engage with the task and process the information. It is clear from general survey literature that people may be more or less willing to engage with survey tasks due to feelings about the relevance or importance of the task [113, 114], personal characteristics (e.g. intellectual capability, attention span) [115], or familiarity with the concepts (such as having done DCE surveys before) [78]. While DCE response rates are often reported and have been associated with DCE complexity and relevance [116], there has been relatively little exploration of a broader conceptualisation of participant engagement as a component of understanding in the context of DCEs, such as dropout rates, time taken, topic relevance, and use of incentives.

4.2 How is Participant Understanding Measured and Used Within the Health DCE Literature?

There appear to be three primary methods to measure or assess participant understanding in the health DCE literature: qualitative methods in pretesting, rationality tests within the DCE, and self-reporting within the DCE. We did not find any studies including econometric approaches, such as the analysis of attribute non-attendance [28] or controlling for inconsistency and fatigue by allowing for unobserved preferences or scale heterogeneity [17, 117]. Regardless of the method used to assess understanding, the reporting of these methods, their results, and the implications for analysis and interpretation are not consistent.

While the use of qualitative methods, particularly in the pretesting phase, are important in developing high-quality DCEs [22], their use does not guarantee understanding during choice task completion. Data collection modes or participant characteristics may differ from pretesting to final rollout [118]. For example, pretesting of attributes and levels is often recommended or conducted using face-to-face interviews, even for paper-based or online DCEs [23]. Similarly, pilot testing may use a sample recruited from a different pool of potential respondents.

There were insufficient details in the included studies to assess whether studies including pretesting for understanding had lower rates of participants failing rationality tests in the final DCE.

Within the DCE instrument itself, quantitative approaches such as duplicate or dominant choice sets are often used to assess rational decision making. Rationality tests or validity checks are frequently cited to assess particular axioms of human decision making [1, 24–26], and exclude irrational, inconsistent or incomplete responses [24, 26, 27]. However, in this review, the meaning and implications of these assessments is described in various ways by different authors, suggesting a lack of clarity around the true purpose and interpretation of these tests. Liu et al. [67] suggested that their proportion of ‘irrational choices’ was “well within the acceptable standard in a DCE”, however these standards are not widely known or used. It also appears that while many studies state they will use the results for exclusion of ‘irrational’ respondents, this is rarely done in the primary analysis. Using these checks as a sensitivity analysis is consistent with both qualitative [27, 119] and quantitative [24] work, suggesting that ‘failing’ tests such as these does not necessarily reflect irrational or uninformative responses, or a lack of understanding of the task [27, 100, 101, 119].

Finally, self-reported ‘debriefing’ questions are often used to assess understanding within the DCE instrument. Ryan and Gerard [29] found 10 of 34 studies reported the results of debriefing questions around ease of completion; 40% of these studies reported that participants expressed difficulty, ranging from 20 to 35% of participants. More recently, a survey of health DCE authors found about half included debriefing questions, but the wording varied widely and the results were often not analysed or reported [25]. There appears to be limited literature to establish whether self-reported difficulty with a DCE is related to response validity or reliability.

4.3 Implications and Next Steps

There is inconsistency in how participant understanding is defined and characterised. This review provides a standardised conceptualisation of participant understanding and how it fits with other key components of DCE reliability

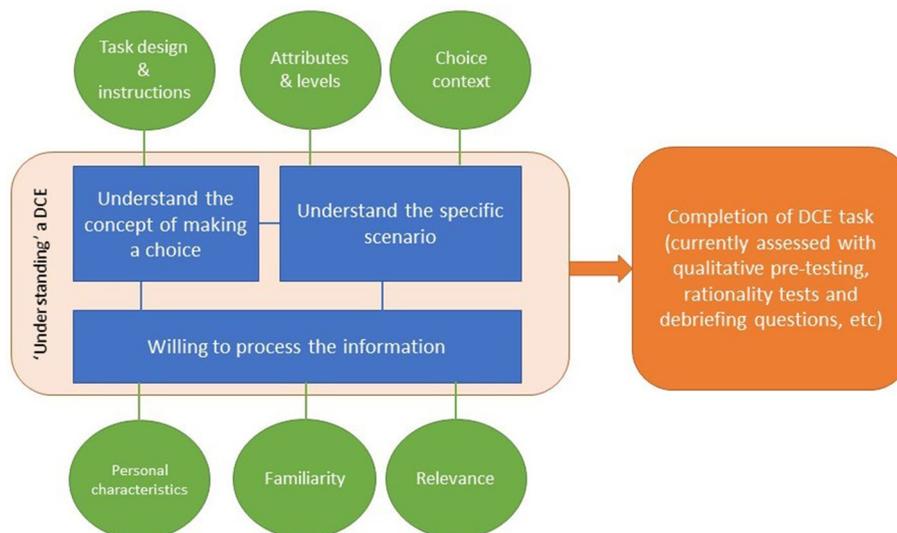
and validity (summarised in ‘Supplementary Material B: Definitions of Key Terms in Relation to Participant Understanding’).

It would appear the concept of ‘understanding’ is commonly discussed as an important component of establishing the validity of DCE instruments, and numerous DCEs use pretesting, rationality checks and debriefing questions to measure or improve understanding. Despite this, there are relatively few studies specifically examining the concept of understanding within DCEs and how it can be measured or improved. The studies that do assess the impact of understanding on some aspect of DCE completion are primarily around alternative presentation formats, or design aspects such as the number of alternatives or choice sets. The existing assessments are primarily designed to assess a single aspect of understanding, such as debriefing questions asking about understanding the concept of making a choice, or an indicator for a lack of understanding, such as a dominated choice set. While assessments of rationality in completing a choice task are important, many may capture some aspects of understanding as well as some aspects of task completion that are unrelated to understanding itself, and more work is required to tease apart these concepts. It is likely that to truly capture overall understanding in terms of the choice context, choice task and task engagement, a suite of tests or assessments will be required.

For those conducting DCEs, our results suggest a need to carefully consider which aspects of understanding are critical to the choice context and ensure the methods to improve and assess understanding can be justified. Ensuring DCEs can be understood by participants will allow individuals and groups who may not participate in DCEs if the tool is too complex to have their preferences recorded. Similarly, improved understanding of a DCE instrument will increase the robustness of the results and ideally lead to improved health or healthcare experience for the community under study. For those using the results of DCEs to inform policy or practice, understanding and identifying the importance, potential risks and implications of participant understanding in DCEs may be an important component of interpreting the results, and improved participant understanding may increase the impact of these studies through increased confidence in their results.

There remain several avenues of future research to improve how we conceptualise, measure, and improve participant understanding in DCEs. Figure 2 illustrates how the three components of DCE understanding identified in this review, i.e. understanding the concept of making a choice, understanding the scenario, and being willing to engage, relate to different aspects of the DCE survey and participant. Future research to establish how each of these components relates to the processes of rational decision making by DCE participants and how they can best be measured is required

Fig. 2 Model of participant understanding in the context of DCEs. *DCE* discrete choice experiment



and may allow identification of methods or strategies to improve understanding by targeting specific components [30, 120].

The methods to measure or assess understanding identified in this review, i.e. qualitative pretesting, rationality tests, and self-reported debriefing questions, appear to be focused primarily on understanding the scenario or understanding the concept of choosing, with considerable intersection between the two. These (and other) strategies may be mapped to one or more of the components of understanding in Fig. 2, allowing a greater appreciation for the implications of failing these assessments. Ongoing research into the impact of various design features, such as task or risk presentation formats on the different components of understanding, and how they are best assessed is also required [111, 121], along with the value of including psychological instruments [122] and econometric techniques [24]. There remains little literature around the use of debriefing questions within DCEs [25], and an opportunity to develop and test a standardised set of questions that can address the different components of understanding may be useful.

4.4 Strengths and Limitations

This is a broad scoping review of how understanding is conceptualised within the health DCE literature. While this allows a general overview of the topic and is particularly useful as a first examination of the topic, this review has several limitations.

The broad conceptualisation of understanding taken in searching and extracting the literature made it difficult to balance sensitivity and specificity within the search strategy. Given the frequency of ‘understanding patient preferences for ...’ statements in the stated preference literature,

the search strategy needed to be relatively generic, leading to the need for high levels of manual review and hence potential bias and errors in the study selection and data extraction. There may also be pockets of literature that address the concept of understanding but which use different terminology to describe the concept, and thus were missed in our search.

We also likely did not capture all the applied DCEs that include some aspect of understanding within our search strategy, representing a possible selection bias. This is consistent with the scoping review methodology, where comprehensiveness is not necessarily the goal, and still demonstrates that understanding is often considered and included in applied choice experiments despite having relatively little research around how it is conceptualised or what it means. In addition, we know that many people who include tests of understanding within their DCE do not always analyse or report their results, therefore this is probably an underestimate of the scope of understanding within applied DCEs. We recommend authors to report their approach to and results of assessing understanding, at least in an appendix to their manuscript.

In addition, we did not look at DCE research outside the field of health. There could certainly be methodological papers that have addressed the concept of understanding within DCEs in other topic areas. If this is the case, it is interesting that there has been relatively little carry through to the health context, perhaps reflecting the unique characteristics of choices made about health and healthcare compared with other, less personal or emotional topics.

Finally, using the scoping review methodology means we did not include a formal evaluation of the quality of the evidence, and, by the nature of the study, we are unable to provide a synthesised result or answer to a specific

question [123]. Rather, we provided an overview of the available literature and how it conceptualises understanding [123].

5 Conclusion

While the concept of understanding appears to be an important component of DCE studies in health, relatively few studies specifically examine the concept within their DCEs, or consider how understanding can be measured or improved. Within the health DCE literature, understanding is typically defined in two ways: understanding of the general choice scenario, such as the medical terms or levels of risk, or understanding the concept of making a choice. However, the impact on understanding of participant willingness to engage with the choice task is a gap in the literature.

The most common methods to measure or assess participant understanding in the health DCE literature are qualitative pretesting, rationality tests within the DCE, and self-reported debriefing questions. Regardless of the method used to measure understanding, the reporting of these methods, their results, and the implications for analysis and interpretation are not consistent, and there remains significant scope for further research into each of these methods and how they relate to and address participant understanding.

For those conducting or using the results of health DCEs, we suggest careful consideration of which aspects of participant understanding are most important in the specific choice context, a critical selection and justification of methods to assess understanding, and transparent reporting of these methods and results.

Acknowledgements The authors would like to acknowledge the contributions of Brendan Mulhern, who gave feedback on the initial project proposal and helped with the full-text reviews, and Bernadette Carr, the librarian who gave assistance developing and implementing the search strategy.

Declarations

Funding During part of this project, Alison Pearce was supported by a University of Technology Sydney Chancellor's Postdoctoral Research Fellowship and the University of Technology Sydney International Researcher Development Scheme. Mark Harrison is supported by a Michael Smith Foundation for Health Research Scholar Award 2017 (#16813), and holds the UBC Professorship in Sustainable Health Care, which, between 2014 and 2017, was funded by Amgen Canada, Astra-Zeneca Canada, Eli Lilly Canada, GlaxoSmithKline, Merck Canada, Novartis Pharmaceuticals Canada, Pfizer Canada, Boehringer Ingelheim (Canada), Hoffman-La Roche, LifeScan Canada, and Lundbeck Canada. The Health Economics Research Unit (HERU) receives funding from the Chief Scientist Office (CSO) of the Scottish Government Health and Social Care Directorates.

Conflict of interest Alison Pearce, Mark Harrison, Verity Watson, Deborah J. Street, Kirsten Howard, Nick Bansback and Stirling Bryan declare they have no conflicts of interest.

Ethical approval Ethics approval and/or informed consent were not required for this study.

Registration As a scoping study, registration of this review was not possible in PROSPERO.

Data availability statement The full data extraction dataset is available from the authors on request.

Author's contribution Initial idea for the study: AP. Feedback on the initial proposal: AP, VW, NB, SB, KH, MH. Confirmation of the inclusion criteria: AP, VW. Full-text reviews completed: AP, NB, MH, KH, DS. Draft manuscript: AP. Critical revision of the manuscript: AP, SB, MH, DS, VW, NB, KH. All authors confirm they have approved the final version of the manuscript to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Alison Pearce will act as the overall guarantor.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial 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-nc/4.0/>.

References

1. Soekhai V, de Bekker-Grob EW, Ellis AR, Vass CM. Discrete choice experiments in health economics: past, present and future. *Pharmacoeconomics*. 2019;37:201–26.
2. Harrison M, Rigby D, Vass C, Flynn T, Louviere J, Payne K. Risk as an attribute in discrete choice experiments: a systematic review of the literature. *Patient*. 2014;7:151–70.
3. Lipkus IM. Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. *Med Decis Mak*. 2007;27:696–713.
4. Bryan S, Dolan P. Discrete choice experiments in health economics: for better or for worse? *Eur J Health Econ*. 2004;5:199–202.
5. Pearce A, Street D, Karikiotis D, McCaffery K, Viney R. Do people with poor health literacy report greater difficulty with discrete choice experiments? 41st Annual AHES Conference: 24–25 September 2019; Melbourne.
6. Araña JE, León CJ, Hanemann MW. Emotions and decision rules in discrete choice experiments for valuing health care programmes for the elderly. *J Health Econ*. 2008;27:753–69.
7. Sever I, Verbič M, Klarić SE. Cost attribute in health care DCEs: just adding another attribute or a trigger of change in the stated preferences? *J Choice Model*. 2019;32:100135.

8. Johnson FR, Mohamed AF, Özdemir S, Marshall DA, Phillips KA. How does cost matter in health-care discrete-choice experiments? *Health Econ.* 2011;20:323–30.
9. Spinks J, Mortimer D. Lost in the crowd? Using eye-tracking to investigate the effect of complexity on attribute non-attendance in discrete choice experiments. *BMC Med Inform Decis Mak.* 2015;16:14.
10. Bech M, Kjaer T, Lauridsen J. Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment. *Health Econ.* 2011;20:273–86.
11. Caussade S, de Ortúzar DJ, Rizzi LI, Hensher DA. Assessing the influence of design dimensions on stated choice experiment estimates. *Transp Res Part B Methodol.* 2005;39:621–40.
12. Kjaer T. A review of the discrete choice experiment—with emphasis on its application in health care. University of Southern Denmark; 2005. Report No.: 2005:1. <https://www.sdu.dk/~media/52E4A6B76FF340C3900EB41CAB67D9EA.ashx>
13. Kanninen BJ. Optimal design for multinomial choice experiments. *J Mark Res.* 2002;39:214–27.
14. Dillman DA, Sinclair MD, Clark JR. Effects of questionnaire length, respondent-friendly design and a difficult question on response rates for occupant-addressed census mail surveys. *Public Opin Q.* 1993;57:289–304.
15. Bless H, Bohner G, Hild T, Schwarz N. Asking difficult questions: task complexity increases the impact of response alternatives. *Eur J Soc Psychol.* 1992;22:309–12.
16. Luce MF, Payne JW, Bettman JR. Emotional trade-off difficulty and choice. *J Mark Res.* 1999;36:143–59.
17. Rezaei A, Patterson Z. Detecting, non-transitive, inconsistent responses in discrete choice experiments. Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation; 2015 Jul. Report No.: CIRRELT-2015-30. <https://pdfs.semanticscholar.org/79a9/7abe48f331f2742eace4dd1d91b1df42ae59.pdf>. Accessed 26 May 2020.
18. Scott A. Identifying and analysing dominant preferences in discrete choice experiments: an application in health care. *J Econ Psychol.* 2002;23:383–98.
19. Egleston BL, Miller SM, Meropol NJ. The impact of misclassification due to survey response fatigue on estimation and identifiability of treatment effects. *Stat Med.* 2011;30:3560–72.
20. Lenzner T. Are readability formulas valid tools for assessing survey question difficulty? *Sociol Methods Res.* 2014;43:677–98.
21. Janssen EM, Marshall DA, Hauber AB, Bridges JFP. Improving the quality of discrete-choice experiments in health: how can we assess validity and reliability? *Expert Rev Pharmacoecon Outcomes Res Taylor & Francis.* 2017;17:531–42.
22. Coast J, Al-Janabi H, Sutton EJ, Horrocks SA, Vosper AJ, Swancutt DR, et al. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. *Health Econ.* 2012;21:730–41.
23. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making. *Pharmacoeconomics.* 2008;26:661–77.
24. Johnson FR, Yang J-C, Reed SD. The internal validity of discrete choice experiment data: a testing tool for quantitative assessments. *Value Health.* 2019;22:157–60.
25. Pearce AM, Mulhern BJ, Watson V, Viney RC. How are debriefing questions used in health discrete choice experiments? An Online Survey. *Value Health.* 2020;23:289–93.
26. Ryan M, Bate A. Testing the assumptions of rationality, continuity and symmetry when applying discrete choice experiments in health care. *Appl Econ Lett.* 2001;8:59–63.
27. Ryan M, Watson V, Entwistle V. Rationalising the 'irrational': a think aloud study of discrete choice experiment responses. *Health Econ.* 2009;18:321–36.
28. Hole A, Kolstad J, Gyrd-Hansen D. Inferred vs stated attribute non-attendance in choice experiments: a study of doctors' prescription behaviour. Sheffield Economic Research Paper Series. Sheffield: Department of Economics, University of Sheffield; 2012.
29. 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.
30. Rigby D, Vass C, Payne K. Opening the 'Black Box': an overview of methods to investigate the decision-making process in choice-based surveys. *Patient.* 2020;13:31–41.
31. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med.* 2018;169:467.
32. Peters M, Godfrey C, McInerney P, Baldini Soares C, Khalil H, Parker D. Chapter 11: Scoping reviews. In: Aromataris E, Munn Z (eds). *Joanna Briggs Institute Reviewer's Manual.* The Joanna Briggs Institute; 2017. <https://wiki.joannabriggs.org/display/MANUAL/Chapter+11%3A+Scoping+reviews>. Accessed 7 Jan 2020.
33. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol.* 2018;18:143.
34. Flynn TN. Valuing citizen and patient preferences in health: recent developments in three types of best–worst scaling. *Expert Rev Pharmacoecon Outcomes Res.* 2010;10:259–67.
35. Clark MD, Szczepura A, Gumber A, Howard K, Moro D, Morton RL. Measuring trade-offs in nephrology: a systematic review of discrete choice experiments and conjoint analysis studies. *Nephrol Dial Transplant.* 2018;33:348–55.
36. Webb EJD, Meads D, Eskyte I, King N, Dracup N, Chataway J, et al. A systematic review of discrete-choice experiments and conjoint analysis studies in people with multiple sclerosis. *Patient.* 2018;11:391–402.
37. Vass C, Gray E, Payne K. Discrete choice experiments of pharmacy services: a systematic review. *Int J Clin Pharm.* 2016;38:620–30.
38. Brown TM, Pashos CL, Joshi AV, Lee WC. The perspective of patients with haemophilia with inhibitors and their care givers: preferences for treatment characteristics. *Haemophilia.* 2011;17:476–82.
39. Lee WC, Joshi AV, Woolford S, Sumner M, Brown M, Hadker N, et al. Physicians' preferences towards coagulation factor concentrates in the treatment of Haemophilia with inhibitors: a discrete choice experiment. *Haemophilia.* 2008;14:454–65.
40. Janssen EM, Segal JB, Bridges JFP. A framework for instrument development of a choice experiment: an application to type 2 diabetes. *Patient.* 2016;9:465–79.
41. van Dijk JD, Groothuis-Oudshoorn CGM, Marshall DA, Ijzerman MJ. An empirical comparison of discrete choice experiment and best-worst scaling to estimate stakeholders' risk tolerance for hip replacement surgery. *Value Health.* 2016;19:316–22.
42. 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 Hum Genet.* 2013;21:1202–8.
43. Mühlbacher AC, Nubling M. Analysis of physicians' perspectives versus patients' preferences: direct assessment and discrete choice experiments in the therapy of multiple myeloma. *Eur J Health Econ.* 2011;12:193–203.
44. Hauber AB, Nguyen H, Posner J, Kalsekar I, Ruggles J. A discrete-choice experiment to quantify patient preferences for frequency of glucagon-like peptide-1 receptor agonist injections in the treatment of type 2 diabetes. *Curr Med Res Opin.* 2016;32:251–62.

45. Youssef E, Cooper V, Miners A, Llewellyn C, Pollard A, Lagarde M, et al. Understanding HIV-positive patients' preferences for healthcare services: a protocol for a discrete choice experiment. *BMJ Open*. 2016;6:e008549.
46. Wong SF, Norman R, Dunning TL, Ashley DM, Lorgelly PK. A protocol for a discrete choice experiment: understanding preferences of patients with cancer towards their cancer care across metropolitan and rural regions in Australia. *BMJ Open*. 2014;4:e006661.
47. Qin L, Chen S, Flood E, Shaunik A, Romero B, de la Cruz M, et al. Glucagon-like peptide-1 receptor agonist treatment attributes important to injection-experienced patients with type 2 diabetes mellitus: a preference study in Germany and the United Kingdom. *Diabetes Ther*. 2017;8:335–53.
48. Hauber B, Caloyeras J, Posner J, Brommage D, Belozeroff V, Cooper K. Hemodialysis patients' preferences for the management of secondary hyperparathyroidism. *BMC Nephrol*. 2017;18:254.
49. Marshall DA, Deal K, Conner-Spady B, Bohm E, Hawker G, Loucks L, et al. How do patients trade-off surgeon choice and waiting times for total joint replacement: a discrete choice experiment. *Osteoarthr Cartil*. 2018;26:522–30.
50. Mohamed AF, Gonzalez JM, Fairchild A. Patient benefit-risk tradeoffs for radioactive iodine-refractory differentiated thyroid cancer treatments. *J Thyroid Res*. 2015;2015:438235.
51. Tada Y, Ishii K, Kimura J, Hanada K, Kawaguchi I. Patient preference for biologic treatments of psoriasis in Japan. *J Dermatol*. 2019;46:466–77.
52. de Freitas HM, Ito T, Hadi M, Al-Jassar G, Henry-Szatkowski M, Nafees B, et al. Patient preferences for metastatic hormone-sensitive prostate cancer treatments: a discrete choice experiment among men in three European Countries. *Adv Ther*. 2019;36:318–32.
53. Bridges JFP, Mohamed AF, Finnern HW, Woehl A, Hauber AB. Patients' preferences for treatment outcomes for advanced non-small cell lung cancer: a conjoint analysis. *Lung Cancer*. 2012;77:224–31.
54. Mansfield C, Srinivas S, Chen C, Hauber AB, Hariharan S, Matczak E, et al. The effect of information on preferences for treatments of metastatic renal cell carcinoma. *Curr Med Res Opin*. 2016;32:1827–38.
55. Schmidt K, Damm K, Vogel A, Golpon H, Manns MP, Welte T, et al. Therapy preferences of patients with lung and colon cancer: a discrete choice experiment. *Patient Prefer Adherence*. 2017;11:1647–56.
56. Muhlbacher AC, Junker U, Juhnke C, Stemmler E, Kohlmann T, Leverkus F, et al. Chronic pain patients' treatment preferences: a discrete-choice experiment. *Eur J Health Econ*. 2015;16:613–28.
57. Bottomley C, Lloyd A, Bennett G, Adlard N. A discrete choice experiment to determine UK patient preference for attributes of disease modifying treatments in Multiple Sclerosis. *J Med Econ*. 2017;20:863–70.
58. Veldwijk J, Lambooi MS, de Bekker-Grob EW, Smit HA, de Wit GA. The effect of including an opt-out option in discrete choice experiments. *PLoS One*. 2014;9:e111805.
59. Heringa M, Floor-Schreuderling A, Wouters H, De Smet PAGM, Bouvy ML. Preferences of patients and pharmacists with regard to the management of drug-drug interactions: a choice-based conjoint analysis. *Drug Saf*. 2018;41:179–89.
60. Vennedey V, Danner M, Evers S, Fauser S, Dirksen C, Stock S, et al. Using qualitative research to facilitate the interpretation of quantitative results from a discrete choice experiment: insights from a survey in elderly ophthalmologic patients. *Patient Prefer Adherence*. 2016;10:993–1002.
61. Naik-Panvelkar P, Armour C, Rose J, Saini B. Patients' value of asthma services in Australian pharmacies: the way ahead for asthma care. *J Asthma*. 2012a;49:310–6.
62. Gregor JC, Williamson M, Dajnowiec D, Sattin B, Sabot E, Salh B. Inflammatory bowel disease patients prioritize mucosal healing, symptom control, and pain when choosing therapies: results of a prospective cross-sectional willingness-to-pay study. *Patient Prefer Adherence*. 2018;12:505–13.
63. Marshall DA, Johnson FR, Phillips KA, Marshall JK, Thabane L, Kulin NA. Measuring patient preferences for colorectal cancer screening using a choice-format survey. *Value Health*. 2007;10:415–30.
64. Ivanova J, Hess LM, Garcia-Horton V, Graham S, Liu X, Zhu Y, et al. Patient and oncologist preferences for the treatment of adults with advanced soft tissue sarcoma: a discrete choice experiment. *Patient*. 2019;12(4):393–404.
65. Zanolini A, Chipungu J, Vinikoor MJ, Bosomprah S, Mafwenko M, Holmes CB, et al. HIV self-testing in Lusaka province, Zambia: acceptability, comprehension of testing instructions, and individual preferences for self-test kit distribution in a population-based sample of adolescents and adults. *AIDS Res Hum Retrovir*. 2018;34:254–60.
66. Mansfield C, Sikirica MV, Pugh A, Poulos CM, Unmuesig V, Morano R, et al. Patient preferences for attributes of type 2 diabetes mellitus medications in Germany and Spain: an online discrete-choice experiment survey. *Diabetes Ther*. 2017;8:1365–78.
67. Liu N, Finkelstein SR, Kruk ME, Rosenthal D. When waiting to see a doctor is less irritating: understanding patient preferences and choice behavior in appointment scheduling. *Manage Sci*. 2018;64:1975–96.
68. Lloyd A, Nafees B, Barnett AH, Heller S, Ploug UJ, Lammert M, et al. Willingness to pay for improvements in chronic long-acting insulin therapy in individuals with type 1 or type 2 diabetes mellitus. *Clin Ther*. 2011;33:1258–67.
69. Damen THC, de Bekker-Grob EW, Mureau MAM, Menke-Pluijmers MB, Seynaeve C, Hofer SOP, et al. Patients' preferences for breast reconstruction: a discrete choice experiment. *J Plast Reconstr Aesthet Surg*. 2011;64:75–83.
70. Byun J-H, Kwon S-H, Lee J-E, Cheon J-E, Jang E-J, Lee E-K. Comparison of benefit-risk preferences of patients and physicians regarding cyclooxygenase-2 inhibitors using discrete choice experiments. *Patient Prefer Adherence*. 2016;10:641–50.
71. Jan S, Mooney G, Ryan M, Bruggemann K, Alexander K. The use of conjoint analysis to elicit community preferences in public health research: a case study of hospital services in South Australia. *Aust N Z J Public Health*. 2000;24:64–70.
72. Tinelli M, Ozolins M, Bath-Hextall F, Williams HC. What determines patient preferences for treating low risk basal cell carcinoma when comparing surgery vs imiquimod? A discrete choice experiment survey from the SINS trial. *BMC Dermatol*. 2012;12:19.
73. de Vries ST, de Vries FM, Dekker T, Haaijer-Ruskamp FM, de Zeeuw D, Ranchor AV, et al. The role of patients' age on their preferences for choosing additional blood pressure-lowering drugs: a discrete choice experiment in patients with diabetes. *PLoS One*. 2015;10:e0139755.
74. Kistler CE, Hess TM, Howard K, Pignone MP, Crutchfield TM, Hawley ST, et al. Older adults' preferences for colorectal cancer-screening test attributes and test choice. *Patient Prefer Adherence*. 2015;9:1005–16.
75. Hofman R, de Bekker-Grob EW, Raat H, Helmerhorst TJ, van Ballegoijen M, Korffage IJ. Parents' preferences for vaccinating daughters against human papillomavirus in the Netherlands: a discrete choice experiment. *BMC Public Health*. 2014;14:454.

76. Hofman R, de Bekker-Grob EW, Richardus JH, de Koning HJ, van Ballegooijen M, Korfage IJ. Have preferences of girls changed almost 3 years after the much debated start of the HPV vaccination program in the Netherlands? A Discrete Choice Experiment. *PLoS One*. 2014;9:e104772.
77. Laba T-L, Brien J, Jan S. Understanding rational non-adherence to medications. A discrete choice experiment in a community sample in Australia. *BMC Fam Pract*. 2012;13:61.
78. Hol L, de Bekker-Grob EW, van Dam L, Donkers B, Kuipers EJ, Habbema JDF, et al. Preferences for colorectal cancer screening strategies: a discrete choice experiment. *Br J Cancer*. 2010;102:972–80.
79. Becker F, Anokye N, de Bekker-Grob EW, Higgins A, Relton C, Strong M, et al. Women's preferences for alternative financial incentive schemes for breastfeeding: a discrete choice experiment. *PLoS One*. 2018;13:e0194231.
80. Buchanan J, Wordsworth S, Schuh A. Patients' preferences for genomic diagnostic testing in chronic lymphocytic leukaemia: a discrete choice experiment. *Patient*. 2016;9:525–36.
81. de Bekker-Grob EW, Rose JM, Donkers B, Essink-Bot M-L, Bangma CH, Steyerberg EW. Men's preferences for prostate cancer screening: a discrete choice experiment. *Br J Cancer*. 2013;108:533–41.
82. Adam D, Keller T, Mühlbacher A, Hinse M, Icke K, Teut M, et al. The value of treatment processes in Germany: a discrete choice experiment on patient preferences in complementary and conventional medicine. *Patient*. 2019;12:349–60.
83. Whitaker KL, Ghanouni A, Zhou Y, Lyrtzopoulos G, Morris S. Patients' preferences for GP consultation for perceived cancer risk in primary care: a discrete choice experiment. *Br J Gen Pract*. 2017;67:e388–95.
84. Johnson FR, Kanninen B, Bingham M, Özdemir S. Experimental design for stated-choice studies. In: Kanninen BJ, editor. *Valuing environmental amenities using stated choice studies: a common sense approach to theory and practice*. Dordrecht: Springer; 2007. p. 159–202. https://doi.org/10.1007/1-4020-5313-4_7
85. Naik-Panvelkar P, Armour C, Rose JM, Saini B. Patient preferences for community pharmacy asthma services: a discrete choice experiment. *Pharmacoeconomics*. 2012b;30:961–76.
86. Mühlbacher AC, Bethge S. Reduce mortality risk above all else: a discrete-choice experiment in acute coronary syndrome patients. *Pharmacoeconomics*. 2015;33:71–81.
87. Naunheim MR, Naunheim ML, Rathi VK, Franco RA, Shrimel MG, Song PC. Patient preferences in subglottic stenosis treatment: a discrete choice experiment. *Otolaryngol Head Neck Surg*. 2018;158:520–6.
88. Cernauskas V, Angeli F, Jaiswal AK, Pavlova M. Underlying determinants of health provider choice in urban slums: results from a discrete choice experiment in Ahmedabad. *India BMC Health Serv Res*. 2018;18:473.
89. Fifer S, Rose J, Hamrosi KK, Swain D. Valuing injection frequency and other attributes of type 2 diabetes treatments in Australia: a discrete choice experiment. *BMC Health Services Res*. 2018;18:675.
90. Naunheim MR, Rathi VK, Naunheim ML, Alkire BC, Lam AC, Song PC, et al. What do patients want from otolaryngologists? A discrete choice experiment. *Otolaryngol Head Neck Surg*. 2017;157:618–24.
91. Lokkerbol J, Geomini A, van Voorthuysen J, van Straten A, Tiemens B, Smit F, et al. A discrete-choice experiment to assess treatment modality preferences of patients with depression. *J Med Econ*. 2019;22:178–86.
92. Whitty JA, Stewart S, Carrington MJ, Calderone A, Marwick T, Horowitz JD, et al. Patient preferences and willingness-to-pay for a home or clinic based program of chronic heart failure management: findings from the Which trial? *PLoS One*. 2013;8:e58347.
93. Meads DM, O'Dwyer JL, Hulme CT, Chintakayala P, Vinall-Collier K, Bennett MI. Patient preferences for pain management in advanced cancer: results from a discrete choice experiment. *Patient*. 2017;10:643–51.
94. Kløjgaard ME, Bech M, Sjøgaard R. Designing a stated choice experiment: the value of a qualitative process. *J Choice Model*. 2012;5:1–18.
95. Katz DA, Stewart KR, Paez M, Vander Weg MW, Grant KM, Hamlin C, et al. Development of a discrete choice experiment (DCE) questionnaire to understand veterans' preferences for tobacco treatment in primary care. *Patient*. 2018;11:649–63.
96. Abdel-All M, Angell B, Jan S, Praveen D, Joshi R. The development of an Android platform to undertake a discrete choice experiment in a low resource setting. *Arch Public Health*. 2019;77:20.
97. Helter TM, Boehler CEH. Developing attributes for discrete choice experiments in health: a systematic literature review and case study of alcohol misuse interventions. *J Subst Use*. 2016;21:662–8.
98. Coast J, Horrocks S. Developing attributes and levels for discrete choice experiments using qualitative methods. *J Health Serv Res Policy*. 2007;12:25–30.
99. McGrady ME, Prosser LA, Thompson AN, Pai ALH. Application of a discrete choice experiment to assess adherence-related motivation among adolescents and young adults with cancer. *J Pediatr Psychol*. 2018;43:172–84.
100. Kenny P, Hall J, Viney R, Haas M. Do participants understand a stated preference survey? A qualitative approach to assessing validity. *Int J Technol Assess Health Care*. 2003;19:664–81.
101. San Miguel F, Ryan M, Amaya-Amaya M. "Irrational" stated preferences: a quantitative and qualitative investigation. *Health Econ*. 2005;14:307–22.
102. Veldwijk J, Determann D, Lambooy MS, van Til JA, Korfage IJ, de Bekker-Grob EW, et al. Exploring how individuals complete the choice tasks in a discrete choice experiment: an interview study. *BMC Med Res Methodol*. 2016;16:45.
103. Cheraghi-Sohi S, Bower P, Mead N, McDonald R, Whalley D, Roland M. Making sense of patient priorities: applying discrete choice methods in primary care using "think aloud" technique. *Fam Pract*. 2007;24:276–82.
104. Veldwijk J, Lambooy MS, van Til JA, Groothuis-Oudshoorn CGM, Smit HA, de Wit GA. Words or graphics to present a discrete choice experiment: does it matter? *Patient Educ Couns*. 2015;98:1376–84.
105. Vass C, Rigby D, Payne K. "I Was Trying to Do the Maths": exploring the impact of risk communication in discrete choice experiments. *Patient*. 2019;12:113–23.
106. Kenny P, Goodall S, Street DJ, Greene J. Choosing a doctor: does presentation format affect the way consumers use health care performance information? *Patient*. 2017;10:739–51.
107. Maddala T, Phillips KA, Johnson FR. An experiment on simplifying conjoint analysis designs for measuring preferences. *Health Econ*. 2003;12:1035–47.
108. Ryan M, Bate A. Testing the assumptions of rationality, continuity and symmetry when applying discrete choice experiments in health care. *Appl Econ Lett*. 2010;8:59–63.
109. Skedgel CD, Wailoo AJ, Akehurst RL. Choosing vs. allocating: discrete choice experiments and constant-sum paired comparisons for the elicitation of societal preferences. *Health Expect*. 2015;18:1227–40.
110. Dolnicar S. Asking good survey questions. *J Travel Res*. 2013;52:551–74.

111. Lim SL, Yang J-C, Ehrisman J, Havrilesky LJ, Reed SD. Are videos or text better for describing attributes in stated-preference surveys? *Patient*. 2020;13(4):401–8.
112. Conrad F, Blair J. From Impressions to Data: Increasing the Objectivity of Cognitive Interviews. Office of Survey Methods Research, US Bureau of Labor Statistics; 1996. <https://www.bls.gov/osmr/research-papers/1996/st960080.htm>. Accessed 5 May 2020.
113. Jacquemet N, James AG, Luchini S, Shogren JF. Social Psychology and environmental economics: a new look at ex ante corrections of biased preference evaluation. *Environ Resour Econ*. 2011;48:413–33.
114. Edwards PJ, Roberts I, Clarke MJ, DiGuseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database Syst Rev*. 2009;3:MR000008.
115. Atkinson TM, Schwartz CE, Goldstein L, Garcia I, Storfer DF, Li Y, et al. Perceptions of response burden associated with completion of patient-reported outcome assessments in oncology. *Value Health*. 2019;22:225–30.
116. Watson V, Becker F, de Bekker-Grob E. Discrete choice experiment response rates: a meta-analysis. *Health Econ*. 2017;26:810–7.
117. Mattmann M, Logar I, Brouwer R. Choice certainty, consistency, and monotonicity in discrete choice experiments. *J Environ Econ Policy*. 2019;8:109–27.
118. Harris L, Brown G. Mixing interview and questionnaire methods: practical problems in aligning data. *Pract Assess Res Eval*. 2010;15:1–19.
119. Lancsar E, Louviere J. Deleting ‘irrational’ responses from discrete choice experiments: a case of investigating or imposing preferences? *Health Econ*. 2006;15:797–811.
120. Jacquemet N, Luchini S, Shogren JF, Watson V. How to Improve Response Consistency in Discrete Choice Experiments? An Induced Values Investigation. Bordeaux: FAERE; 2016. pp. 1–17. https://faere.fr/pub/Conf2016/Luchini_dce_FAERE2016.pdf. Accessed 17 Apr 2018.
121. Di Santostefano R, Reed S, Yang J-C, Levitan B, Johnson FR. Do people understand benefits and risks? Even ineffective treatments for delaying Alzheimer’s disease are preferable to “doing nothing” for many older adults. *Pharmacoepidemiol Drug Saf*. 2018;27:511.
122. Russo S, Jongerius C, Faccio F, Pizzoli SFM, Pinto CA, Veldwijk J, et al. Understanding patients’ preferences: a systematic review of psychological instruments used in patients’ preference and decision studies. *Value Health*. 2019;22:491–501.
123. Sucharew H, Sucharew H. Methods for research evidence synthesis: the scoping review approach. *J Hosp Med*. 2019;14:416–8.

Affiliations

Alison Pearce¹  · Mark Harrison²  · Verity Watson³  · Deborah J. Street⁴  · Kirsten Howard¹  · Nick Bansback⁵  · Stirling Bryan⁵ 

¹ Sydney School of Public Health, University of Sydney, Sydney, NSW, Australia

² Centre for Health Evaluation and Outcome Sciences, University of British Columbia, Vancouver, BC, Canada

³ Health Economics Research Unit, University of Aberdeen, Aberdeen, Scotland

⁴ Centre for Health Economics Research and Evaluation, University of Technology Sydney, Sydney, NSW, Australia

⁵ School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada