Enlightening Tourism.

A Pathmaking Journal

Universidad de Huelva

Enlightening Tourism. A Pathmaking Journal
Universidad de Huelva

CONTENT VALIDITY OF RESEARCH INSTRUMENTS: ASSESSING DOMESTIC ECOTOURISM IN PROTECTED AREAS

Mohammad Ridhwan Nordin
Politeknik Muadzam Shah (Malaysia)
ridhwannordin2016@gmail.com

Salamiah A. Jamal
Universiti Teknologi MARA Cawangan Selangor (Malaysia)
drsalamiah@uitm.edu.my

Nor Asmalina Mohd Anuar
Universiti Teknologi MARA Cawangan Selangor (Malaysia)
noras288@uitm.edu.my

ABSTRACT

The assessment of content validity is a critical and complex step in the instrument development process, frequently used to measure complex constructs in social science research. Nevertheless, there is scarce documentation in the field of tourism and hospitality on how content validity should be analysed. This study aimed to present the assessment of content validity of an instrument developed to assess domestic ecotourism in protected areas. A 54-item survey was designed based on previous literature. The content validity assessment was conducted by recruiting a panel of experts to evaluate the instrument's elements and rate them based on their relevance to provide a content validity index for each item.

KEYWORDS

Instruments development; Content validity index; Modified Kappa; Ecotourism; Protected areas.

ECONLIT KEYS
I31; Z32; Z33.

1. INTRODUCTION

Research instruments are the most widely used data collection tools, especially in social science research, because of their many advantages. Among the benefits include collecting data from a vast population in a limited time and at a lower cost, offering convenience to respondents, providing anonymity, ensuring lack of interviewer bias, and enabling standardisation of questions. However, the disadvantages of a questionnaire are poor data quality due to incomplete and inaccurate questions, wording problems, and a flawed development process. These problems are critical and can be avoided or mitigated (Gillham, 2008).

To ensure the quality of an instrument, using a previously validated questionnaire can be useful. This will save time and resources in the development process and testing of its validity and reliability. However, there can be situations whereby a new questionnaire is needed (Boynton, 2004), or an adapted and modified questionnaire to suit the new context, setting, industry, population, or culture (Bahkia et al., 2020; Eom & Lu, 2019) may become necessary. Thus, a validity assessment must be conducted to validate the new or modified instrument, mainly if the original instrument was developed in the population of different cultures and industries from the present study (Ehido et al., 2020 & Ghorbanzadeh et al., 2019). Even when the item or scale adopted is unaltered, it is still necessary to validate them to assess their relevance to different countries or regions (Elangovan et al., 2021).

Whenever a new scale or questionnaire needs to be developed, adopted, adapted and modified, a structured validation method will essentially help to produce a quality instrument. However, novice researchers are often confused when selecting and conducting proper validity checks to test their research instrument (questionnaire/ survey), especially among undergraduate or postgraduate researchers, who are considered novices in this area. Most of the time, these novice researchers will address the question and seek the answer on research design in social networking research

groups, such as Research Gate, Doctorate Support Group (DSG) on Facebook, or other related research groups in Telegram and WhatsApp applications.

There are many approaches to validating the "goodness" of the instruments developed. Validity assessments are grouped under three main forms: content validity, criterion-related validity, and construct validity (Sekaran et al., 2016). Amongst them, content validity is a precondition for the other types of validity, thus, it should be given the utmost importance in the instrument validation process (Zamanzadeh et al., 2014). Content validity gives information on the representativeness and clarity of items and assists the improvement of an instrument by obtaining expert panel recommendations (Polit & Beck, 2006; Polit et al., 2007). In addition, it provides preliminary evidence on the construct validity of an instrument.

Practices of content validation have gained significant attention in the tourism literature and studies on the instrument or scale development have been researched in various settings such as food tourism (Huang & Choi, 2019), medical tourism (Sag et al., 2022), rural tourism (Pratama & Wulandari, 2019), ecotourism (Beall & Boley, 2021), memorable tourism experience (Kim & Ritchie, 2014), tourist engagement (Chen et al., 2019; Choe & Kim, 2019; Xie et al., 2020) and anime tourism motivation (Liu et al., 2020). Unfortunately, most of those studies reported an unrigorous validation process, limited to the judgemental method, describing content validity and how the content validity process was conducted qualitatively.

Some studies abovementioned have not comprehensively revealed the criteria of the experts' appointment, how they analysed the content review with statistical methods, and how this empirical result should be reported and documented. Moreover, some of these studies only focused on a scale refinement or purification (Churchill, 1979) that used exploratory and confirmatory analysis (EFA & CFA) in validating the instruments, which abandoned the sequential preliminary procedures, starting from content validation (Almanasreh et al., 2019), that are crucial before the factor analysis is conducted (Elangovan & Sundaravel, 2021). In addition, Koc & Ayyildiz (2021) in their review of 253 scale development studies in the Hospitality and Tourism field found that 58% of 150 studies failed to report the content validity. This deficiency is coinciding with Gursoy et al. (2015), as some of the authors did not explain the process in detail or some might consider it unimportant.

Thus, the purpose of this study was to report on the item development and initial validity of a new survey to assess ecotourists' perceptions of destination image, perceived value, satisfaction and quality of life, especially in the ecotourism context claimed by Koc & Ayyildiz (2021) as remain lacking in instruments and scales. This study mainly focuses on developing and applying complex constructs in tourism research that frequently go through expert validation as part of the content validity process. It is an assessment among professionals or field experts to ensure that the instrument is measuring what it is supposed to measure (Gursoy et al, 2015).

This study employed a two-stage approach (Lynn, 1986). To enlighten the readers on the practicality of conducting the process, guidelines are provided consisting of five steps that are easy to follow and beneficial to novice researchers. The content validity evidence is represented by a content validity index (CVI) (Davis, 1992; Polit & Beck, 2006; Polit et al., 2007; & Lynn, 1986), for instance, several recent studies (Gregori-Giralt et al., 2021; Mason et al., 2020; Shahsavari et al., 2020; Tay et al., 2021) have established content validity using CVI to support the validity of the assessment tool. Based on the evidence, this study describes the best practices and steps to quantify the content validity of an assessment tool using CVI and modified kappa (k*).

The remainder of this article is structured as follows. Section 2 discusses various literature studies on the fundamentals of content validation and the stages involved in content validation. Section 3 presents the methodology by introducing the practical steps in conducting the content validation process. Section 4 provides the results of the study. The section ends with a discussion of the findings, implications, and possible future research avenues. Section 5 concludes the paper with a note on the study's contribution. The article concludes with three rather simple but important implications for future research.

2. LITERATURE REVIEW

The strength of a research study design is strongly dependent on how precisely the identified variables are measured; known as the validity (Kelley, 1999). Validity denotes the extent to which specific items on a tool accurately assess the concept of being measured in a research study. Validity ensures that the questions being asked would

allow valid inferences to be made. There are many approaches to the validation of instruments which involve a sequence as one is a precondition to another. The three types of validity considered for this tourism and hospitality research are (1) content validity; (2) criterion-related validity; and (3) construct validity (Creswell, 2014; DeVellis, 2016; Sekaran et al., 2016). Nonetheless, this study only focused on content validity and although criterion and construct validity are considered important, information on measuring the content validity is necessary as a prerequisite for all assessment instruments (Beckstead, 2009; Polit & Beck, 2006; Shrotryia & Dhanda, 2019).

In the literature, Lennon (1956) was the first person who defined content validity as "the extent to which a subject's responses to the items of a test may be considered to be a representative sample of individual responses to a real or hypothetical universe of situations which together constitute the area of concern to the person interpreting the test". In 1971, Cronbach presented his definition of content validity as the extent to which the items on instruments are sampled adequately from the specified domain of content. Since then, many scholars also phrased content validity as the degree to which the measured items represent a proper sample of the theoretical content domain of a construct (Nunnally & Bernstein, 1994; Polit & Beck, 2006).

In general, content validity is the process of evaluating an instrument to verify that it contains all of the necessary items while excluding those that are unimportant to a certain construct area (Lewis et al., 1995; Boudreau et al., 2001). Content validity is assessed early in an instrument's development by a panel of experts who score each element's relevance and representativeness to the content domain. Thus, the validity process will determine the instrument's ability to accurately reflect or quantify the concept of being evaluated (Grove et al., 2013).

Researchers could gain significant information by conducting a systematic and comprehensive content validation. However, developing a valid measure of an instrument is not as simple as it seems. There is a myriad of proposed processes for conducting content validity. Sireci (1998) classified the process of content validity based on judgemental and statistical (quantifying). Other scholars such as Shrotryia et al. (2019) and Zamanzadeh et al. (2014) categorised the process as a two-stage process involving development and judgement (quantifying). On the other hand, Amos et al. (2022) followed prior researchers such as Almanasreh et al. (2019), who

recommended a three-stage process involving the development stage, judgement and quantifying, and then revision and reconstruction.

Based on the comparison of numerous recommended processes, this study concluded that all the studies mentioned above are based on Lynn (1986), who introduced two stages of the content validity process consisting of the development and judgement-quantification stages. Thus, utilising a single-stage method (either development or judgement) to evaluate content validity is frequently insufficient and may result in a low-quality construct (Almanasreh et al., 2019). Hence, we concurred with Lynn (1986) that these two stages are considered essential for the instrument validity process. However, a third stage is also advisable to be conducted after the judgement and quantification stages.

2.1) DEVELOPMENT STAGE

The first stage (development phase) involved a thorough literature search (Liu et al., 2020) on an existing scale of the constructs and dimensions intended to measure with consultation from the experts (Chen et al., 2019). The thorough literature search process is guided based on conceptual and operational definitions, the attributes and characteristics of the desired construct, dimensions, components and its boundaries (Zamanzadeh et al., 2014). Even though all studies intend to adapt previously validated instruments or adapt and modify them, it is necessary to determine the conceptual and operational definitions of the construct to be measured. Both definitions should be defined clearly and serve as the cornerstone to the subsequent steps before developers sample the suitable items (Beck, 1999). During this process, the experts play a role in checking the appropriateness of the selected construct and dimension (Elangovan & Sundaravel, 2021).

It is always recommended for novice researchers to opt for the systematic literature searching process (Shaffril et al., 2020) when selecting the research and academic articles relating to the intended constructs, dimensions, and existing instruments. It is believed that this method is the most explicit, reproducible, and comprehensive. Thus, by employing this approach, researchers could summarise and critically interpret prior studies on specific intended constructs and dimensions, explore new and emerging

trends, identify the dearth, and help assess potential shortcomings for future research, if any (Major & Savin-Baden, 2010; Pickering & Byrne, 2014).

This method begins with the formulation of research questions and acts as guidance on the overall review process (de Menezes & Helliher, 2011). This is followed by clarification on the theoretical or conceptual context of the subject area. Then, experts or the supervisory team on the area of study are consulted in assisting for keyword identification suitability (Vada et al., 2020) developed from the research questions (Okoli, 2015), scoping from prior studies, keywords suggested by database engines or relying on online thesaurus (Shaffril et al., 2020). Next, the search is run on different selected leading and supporting databases on the main and enriched keywords by advanced searching techniques. Several leading databases that are recommended are Scopus, Web of Science, and Science Direct (Gusenbauer, 2019; Haddaway et al., 2015; Shaffril et al., 2020; Vada et al., 2020)

Specific inclusion and exclusion criteria to screen the articles need to be applied. As Okoli (2015) argued, it is almost impossible for a researcher to review a plethora of published articles, thus applying specific criteria is a more realistic idea. The selection criteria are specific to the context of the study, time frame, empirical data and published in a journal, appropriate methodology and English version articles to avoid confusion on understanding. After this screening process, the teams could assist the researcher in conducting a quality appraisal of the retrieved articles sorting the articles that meet the objective of the study. Now, the researcher is eligible to generate pool items from the high-quality existing instruments.

Then, input from the thorough selection of generated pool items is translated to generate proposed items of validation that need to be done by experts. At this stage, the generated pool items from existing instruments are refined and organised in a suitable format and sequenced so that the finalised items are transformed into a structured validation assessment form. DeVellis (2016) suggested that during this refinement process on generated items, there are several tasks that researchers should consider for the inclusion criteria of validation items. The selected items should accurately reflect the instrument's purpose in measuring the dimensions and the underlying latent variables. Items should be in logical sequence and researchers should avoid lengthy items, as this would increase complexity and reduce clarity. Moreover,

researchers need to consider the reading difficulty level, exclusion or inclusion of positively or negatively worded items, redundancy, and avoid double-barrelled, confusing and misleading items. A good item should be unambiguous and unbiased to the specific groups such as gender, minority or linguistic differences (Elangovan & Sundaravel, 2021) and should represent the domain of interest (Thoyre et al., 2014).

Researchers are required to confirm that the instrument items and overall instruments have content validity by a specific number of experts. There are myriad suggestions for determining the number of experts, and it is always considered subjective and inconsistent (Lam et al., 2018). The essence of determining the number of experts is to decrease the probability of chance agreement if a high number of experts are involved (Shrotryia & Dhanda, 2019; Zamanzadeh et al., 2014). Thus, the smaller the better in controlling chance agreement. The renowned recommendation is a minimum of three experts (Lynn, 1986). However, the maximum number of experts has not been specified. The final decision on the number of experts' appointments, depends on the complex nature of the study, the desired level of expertise, and the range of knowledge to represent the areas of study (Grant & Davis, 1997).

Even though several studies, especially in a tourism context (Huang & Choi, 2019; Hong et al., 2020; Sag et al., 2022), informed that their instruments were validated by a panel of experts, the details on the selection criteria of experts were often unreported. Grant and Davis (1997) provided guidelines for selecting these individuals. An expert should have a well-defined criterion such as qualifications, experiences, history of publications on the pertinent area of interest, reviewing work, and relevant training on the subject matter. After assembling an expert panel, the researchers can now collect and analyse their quantitative and qualitative perspectives. It is always suggested that the researchers be present during the process to facilitate the validation (Taherdoost, 2016). However, this notion seems unidealistic and impossible to get experts from one geographical location. Moreover, the ongoing Covid-19 pandemic has limited the mobility of people and increased the perceived risk of gathering in an event. Thus, submission of the assessment form through an online platform to the selected experts is more efficient and less risky.

2.2) JUDGEMENT-QUANTIFYING STAGE

This stage is instigated after researchers receive feedback from the experts; thus, the quantitative approach can be conducted. After the experts return the form, the instrument is statistically analysed by a content validity index (CVI) and Kappa statistic. For each item, the individual content validity index (I-CVI) is calculated by the number of experts who provide a ranking of 3 or 4, divided by the total number of experts (Davis, 1996). The I-CVI for relevance and clarity is calculated using the same formula. The scale level content validity based on average (S- CVI)/AVE is also computed for both relevance and clarity for each domain (Lynn, 1986; Polit & Beck, 2006). This is calculated based on the average of the I-CVI scores for all items on the scale or the average of proportion relevance judged by all experts. The acceptable CVI value for each item ranged from 1.00 to 0.71 (Davis, 1996), whereas generally the S-CVI/AVE should be greater than 0.80 (Polit & Beck, 2006). An I-CVI of less than 0.70 is indicative of the indicator being either irrelevant or unclear which then requires changes to be made by way of revision, rewording to improve clarity, and possibly eliminating the indicator (Thoyre et al., 2014).

Even though CVI is widely used in conducting the content validation process (Shrotriya & Dhanda, 2019), researchers need to consider an inflated value to occur because of the possibility of a chance agreement. Thus, Kappa statistics are suggested to give a more quantifiable understanding of expert evaluation. The Kappa statistic is a consensus indicator of inter-expert agreement that is used in conjunction with the CVI to confirm that the expert agreement is not due to chance (Zamanzadeh et al., 2014). Polit et al. (2007) introduced a modified Kappa (k*) as a promising method to evaluate the content validity beyond the degree of chance agreement. The modified Kappa (k*) can be calculated using the value of the probability of chance agreement (Pc) and the computed I-CVI. The threshold of modified Kappa (k*) is divided into three criteria: fair (0.40-0.59), good (0.60-0.74) and excellent (>0.74).

Thus, in conclusion, content validity involves a two-stage process, item development, and the judgement-quantifying stage. It is recommended to quantify the content validity based on CVI and modified Kappa (k*) coefficient. Nevertheless, CVI is the most frequently used by researchers because of the simplicity of the calculation

method. Therefore, this study would recommend researchers use both measurement techniques to increase confidence in the validity of the developed instruments.

3. METHODOLOGY

This study underpinned the two-stage process suggested by Lynn (1986) and other recent studies (Amos et al., 2022; Shamsudheen & Chowdhury, 2022; Huang & Choi, 2019; Taheri et al., 2014). This chapter provides a set of guidelines that researchers can use in conducting the content validity process. A five-step process is introduced as detailed in Table 1 below.

Stages		Steps
Development	1	Instrument development.
	2	Selecting panel of experts.
	3	Preparing experts participation
Judgement and Quantifying	4	Conducting content validation
	5	Analysing data.

Table 1: Five Steps in Content Validity Process.
Source: Authors.

Step 1: Instrument Development

A good research instrument will result in an appropriate data collection process, and upon successful implementation, it will lead to replicable research and enable firm conclusions to be drawn (Creswell, 2014). The research instrument in this study was designed and developed according to previous literature and sourced from several instruments. An extensive literature review was conducted to determine the dimensions for destination image, perceived value, satisfaction, and quality of life. As mentioned earlier, this study exerted a systematic literature searching process (Shaffril et al., 2020) involving the identification of keywords and database, screening to select the research articles based on the determining of inclusion and exclusion criteria, and then eligibility to select the research articles relating to the intended study context and focus, constructs and dimensions.

In the identification process, this study used keywords based on the research questions (Okoli, 2015), keywords from previous studies, thesaurus and suggestions

from database search engines (Shaffril et al., 2020). Then, this study enriched and developed a full search string (Table 2) from the list of keywords. This full search string was used on three main databases, which are Science Direct, Scopus and Web of Science (WoS) search engine. However, this study also uses google scholar as an additional database for searching articles that are not listed in those three main databases. All the specific constructs/dimensions search strings are linked with the Boolean operator (OR) and (AND).

Constructs	Databases	Search String
Destination Image	Science Direct	“destination image” OR “tourism destination image” OR "tourism image"
	Scopus	"destination image" OR "tourism destination image" OR "tourism image"
	WoS	“destination image” OR “tourism destination image” OR "tourism image"
Perceived Value	Science Direct	"tourism perceived value" OR "tourist perceived value" OR "customer value" OR "consumer value" OR "service value"
	Scopus	"perceived value" OR "customer value" OR "consumer value" OR "service value" AND "tourism" OR "tourist*"
	WoS	“customer perceived value” OR “perceived value” OR "tourist perceived value"
Satisfaction	Science Direct	“tourist satisfaction” OR "ecotourist satisfaction" OR "tourism satisfaction" OR "destination satisfaction" OR "vacation satisfaction" OR "trip satisfaction"
	Scopus	(tourist OR tourism OR vacation OR destination AND satisfaction)
	WoS	“tourist satisfaction” OR "ecotourist satisfaction" OR "tourism satisfaction" OR "destination satisfaction" OR "vacation satisfaction" OR "trip satisfaction"
Quality of Life	Science Direct	"tourism quality of life" OR "tourists quality of life" OR "touristification on quality of life" OR "visitor quality of life" OR "well-being perception” OR "leisure life satisfaction"
	Scopus	("tourism AND quality of life” OR “tourists AND quality of life” OR "touristification on quality of life" OR "visitor AND quality of life" OR "well-being perception" OR "leisure life satisfaction")
	WoS	"tourism quality of life" OR "tourists quality of life" OR "touristification on quality of life" OR "visitor quality of life" OR "well-being perception” OR "leisure life satisfaction"

Table 2: Search strings.
Source: Authors.

Next, from the search results, the articles went through a screening process against the selection criteria. Based on the inclusion and exclusion criteria (Table 3), this study screened the relevant articles that matched those criteria. After the relevant articles were sorted, this study looked into the title, abstracts, and/or keywords that contained at least the search term as well as the result and instruments reported in the eligibility process to finalise the exact usable article for the instrument development.

Criterion	Inclusion	Exclusion
Document Type	Article journal (empirical paper) conference proceeding	Chapter in book, book series
Publication Stage	Final	In press
Source Type	Journal, Conference proceeding	Book, Book series
Timeline	2005-2021	<2005
Language	English	Non-English
Relevancy	Related to the keywords	Unrelated to the keywords
Subject Area	Tourism, Ecotourism, National Park, Protected Park	Others

Table 3: Inclusion and exclusion criteria.
Source: Authors.

Based on the literature search from the three databases, this study obtained 2,256 articles on destination image, 5,420 articles on perceived value, 2,419 articles on satisfaction and 469 articles on quality of life (see Table 4). Next, the author conducted a screening process based on inclusion and exclusion criteria and this study excludes a total of 3,500 articles. Next step, the authors exported all data from the three databases to Microsoft Excel to manually remove duplicated documents. There are a total of 1,445 duplicated articles, which were removed, and the remaining were then screened against quality assessment criteria.

This assessment involves sorting the articles based on low, moderate and high-level quality (Petticrew & Robert, 2006). This study only selected 176 articles that meet high-level quality articles that are mutually agreed upon by the supervisory team. The team examined the titles and abstracts as well as the methodology section. Articles which found poor reporting, and a lack of replicable for generating pool instruments have been

excluded. The specific sources selected in Table 4 are based on high-level quality articles that meet the criteria of (i) study area (tourism) (ii) quantitative method, (iii) comprehensive instrument reporting and (iv) revealing the set of instruments used in the study (v) instruments suitable to adapt for the study.

Constructs	Databases	Title-Abs-Key	Inclusion	Duplicate	Quality Assessment
Destination Image	Science Direct	393	289	628 (removed)	942 (removed)
	Scopus	1018	608		
	WoS	845	642		
	Total	2256	1632	1004	62
Perceived Value	Science Direct	1734	436	105 (removed)	897 (removed)
	Scopus	1047	496		
	WoS	2639	147		
	Total	5420	1079	974	77
Satisfaction	Science Direct	243	180	698 (removed)	11 (removed)
	Scopus	1678	252		
	WoS	498	299		
	Total	2419	731	33	22
Quality of Life	Science Direct	183	6	14 (removed)	29 (removed)
	Scopus	193	31		
	WoS	93	21		
	Total	469	58	44	15

Table 4: Systematic literature search result.
Source: Authors.

The study opted for a structured questionnaire as the research instrument. The primary sources for the items in the instrument had been identified, adapted, and modified from past studies (Ahn & Back, 2019; Aliman et al., 2016; Castellanos-Verdugo et al., 2016; Gallarza et al., 2017; Glyptou, K., 2020; Han et al., 2020; Kim et al., 2020; Kim & Thapa, 2017; Kim et al., 2015; Li et al., 2021; Rasoolimanesh et al., 2016; Rasoolimanesh et al., 2021; Prebensen et al., 2012; Prebensen et al., 2013; Wang et al., 2018; William & Soutar, 2009). Table 5 highlights the summary of research instruments and its sources.

Construct	Dimension	Source	Databases	No. of Item
Destination Image	Destination attributes	Aliman et al., 2016 Glyptou, 2020 Li et al., 2021	Google Scholar Scopus/WoS SD/Scopus/Wos	10
Perceived value	Functional	William & Soutar, 2009 Rasoolimanesh et al., 2016	SD/Scopus/WoS SD/Scopus/WoS	6
	Economic	Prebensen et al., 2012 Wang et al., 2018 Rasoolimanesh et al., 2016	Scopus/WoS SD/Scopus/WoS SD/Scopus/WoS	5
	Novelty	Wang et al., 2018 Kim & Thapa, 2017	SD/Scopus/WoS SD/Scopus/WoS	4
	Health and Wellbeing	Prebensen et al., 2013 Ahn & Back, 2019 Gallarza et al. 2017	Scopus/WoS Scopus/WoS Scopus/WoS	7
Satisfaction		Castellanos-Verdugo et al., 2016	SD/Scopus/WoS	6
Quality of life		Kim et al., 2020 Han et al., 2020 Kim et al., 2015	WoS SD/Scopus/WoS SD/Scopus/WoS	7
Demography items				9

Table 5: Research Instruments.

Step 2: Selecting Panel of Experts

The next stage involved confirming the items by a specific number of experts to assess and ensure the content validity of the instrument. The panel of experts were selected based on well-defined criteria such as qualification, professional experiences, field content expertise, a recent line of research publications on the topic, and availability to participate in the assessment activity within the stipulated time frame (Almanaresh et al., 2019; Gregori-Giralt et al., 2021; Rubio et al., 2003).

The selection of an individual to review and critique an assessment tool (i.e., questionnaire) is usually based on the individual's expertise within the topic to be studied, the construct being developed, and familiarity with the target population on whom the instrument will be used. Considering the recommendations from previous literature, this study followed the advice by Lynn (1986), where there should be a minimum of three experts for good content validity results. Thus, three experts were approached and requested to review the instrument. Furthermore, Polit et al. (2007)

argued that three or four experts will establish a perfect result on the chance agreement as the probability of chance agreement will decrease if a high number of experts are involved (Shrotryia & Dhanda, 2019). The questionnaire was reviewed by three academicians; two from the Faculty of Hotel and Tourism Management, Universiti Teknologi MARA (UiTM), and one from the Faculty of Economics and Business Management, Universiti Sultan Zainal Abidin (UNISZA). These academicians are highly experienced and qualified in the field of tourism, marketing and quality of life. Table 6 shows the list of experts involved in this study.

Designation	Affiliation	Expertise
Professor	Universiti Teknologi MARA (UiTM)	Tourism
Doctorate senior lecturer	Universiti Teknologi MARA (UiTM)	Tourism Marketing
Doctorate senior lecturer	Universiti Sultan Zainal Abidin (UNISZA)	Marketing, Quality of Life

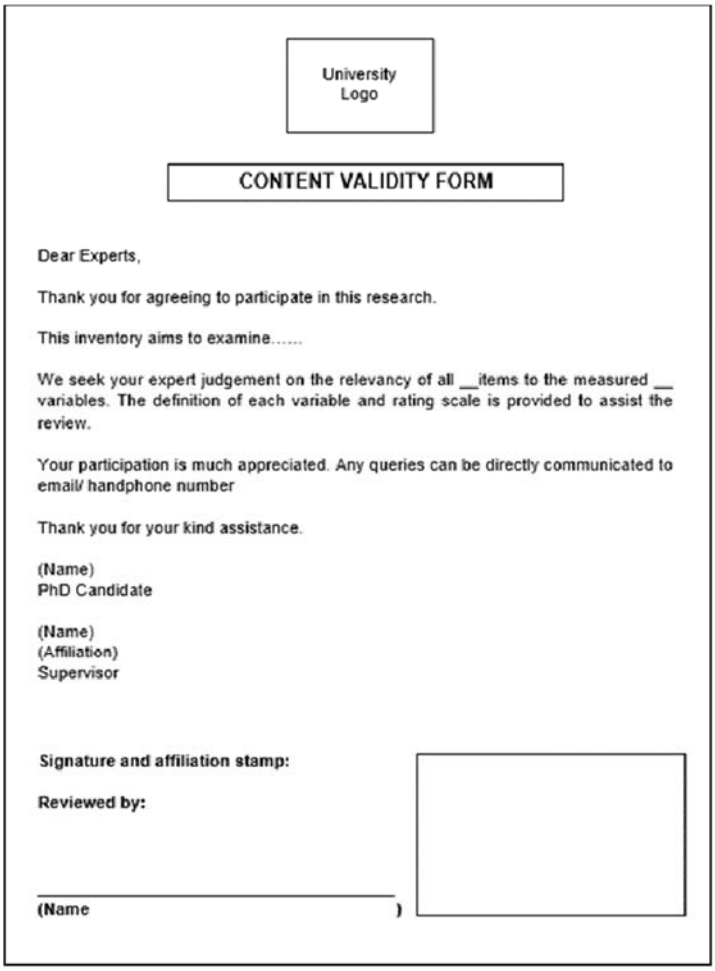
Table 6: List of Panel of Experts Appointed.
Source: Authors.

Step 3: Preparing Experts' Participation

After identifying the potential panel members, the first step is to make a phone call or send an email or a WhatsApp text to introduce yourself, the study, and the invitation to the expert to be appointed as one of the panels of experts to validate the instrument. It is recommended to provide ample time, by at least scheduling one week in advance of the appointment, for the candidates to respond to your request. The second step is when the expert agrees with the appointment, and the content validation form needs to be prepared to ensure that the panel of experts have a clear understanding of the aim and objectives of the task. It is recommended to provide a cover letter when sending a questionnaire for review – please refer to Figure 1 for an example.

The ongoing Covid-19 pandemic has created a challenging environment (Dodds & Hess, 2020), especially in conducting a face-to-face validation process. Online distribution of the instrument is the most effective way and a very feasible adaptation, given the current restrictions (Sy et al., 2020). Moreover, this non-face-to-face approach significantly reduces personal hazard from the pandemic and facilitates cost, time, and resource efficiency. It is suggested that the panel members be given two

weeks to conduct the content validity assessment process or an extended duration of that time. The researcher must understand that the reviewer has to make adjustments to their professional and personal lives due to the significant disruption caused by the Covid-19 pandemic.



University Logo

CONTENT VALIDITY FORM

Dear Experts,

Thank you for agreeing to participate in this research.

This inventory aims to examine.....

We seek your expert judgement on the relevancy of all __items to the measured __ variables. The definition of each variable and rating scale is provided to assist the review.

Your participation is much appreciated. Any queries can be directly communicated to email/ handphone number

Thank you for your kind assistance.

(Name)
PhD Candidate

(Name)
(Affiliation)
Supervisor

Signature and affiliation stamp:

Reviewed by:

_____
(Name)

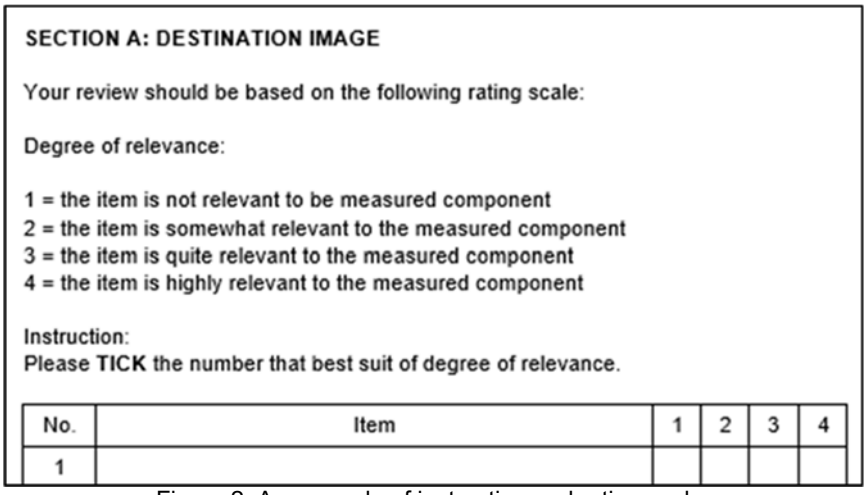
Figure 1: An example of the front page of the content validity form.
Source: Authors.

Step 4: Conducting Content Validation

The CVI is based on expert ratings for each item based on an instrument's content relevance or representativeness, usually on a 4-point Likert scale ranging from 1 (not relevant or not representative) to 4 (highly relevant or representative). The use of a four-point scale adheres to Lynn's (1986) guideline to avoid neutral and ambiguous

points. The number of items with a score of 3 or 4 is divided by the total number of answers to calculate an index of CVI.

The experts are specifically requested to critically review the dimensions and its items before providing a score on each item. They are encouraged to provide written comments to improve the relevance of items to the targeted dimensions. All comments are taken into consideration to refine the dimension and its items if any. The experts are also expected to review the logical sequence of the items, scale suitability, wording, understandability, and ambiguity. Upon completion of the process of reviewing the domain and items, the experts are requested to provide a score for each item, independently, based on the relevant scale (Figure 2). The experts are then required to submit their responses to the researcher once they have completed the assignment of scores to all items.



SECTION A: DESTINATION IMAGE

Your review should be based on the following rating scale:

Degree of relevance:

1 = the item is not relevant to be measured component
2 = the item is somewhat relevant to the measured component
3 = the item is quite relevant to the measured component
4 = the item is highly relevant to the measured component

Instruction:
Please **TICK** the number that best suit of degree of relevance.

No.	Item	1	2	3	4
1					

Figure 2. An example of instruction and rating scale.
Source: Authors.

Step 5: Analysing Data

After obtaining responses from the experts, a quantitative analysis is performed. Souza et al. (2017) argued that there is no consensus approach to examining the content validity of an instrument. However, the most widely utilised method of quantifying the content validity for an instrument is the calculation of the content validity index (CVI). CVI is simple and easy to compute, quickly understood and interpreted, and able to provide content validity for each item as well as the instrument as a whole.

Also, it can be utilised to assess the performance of the experts and most importantly, it allows the instrument developer to decide on whether to retain or exclude items from an instrument.

This article has categorised CVI into two forms: CVI for an item (I-CVI) and CVI for the whole scale (S-CVI). The CVI value is the percentage of judges who agree with an item, where the index value of at least 0.78 and higher than 0.90 is accepted (Polit & Beck, 2006; Polit et al., 2007). An acceptable S-CVI/Ave value according to available guidelines is a minimum of 0.90. The index value of 1.00 for the three experts was considered acceptable for this study. Table 7 summarises the recommended cut-off value based on the number of experts.

Number of experts	Cut-off Value	Source
Two experts	At least 0.80	Davis (1992)
Three to five experts	Should be 1	Polit & Beck (2006), Polit et al., (2007)
At least six experts	At least 0.83	Polit & Beck (2006), Polit et al., (2007)
Six to eight experts	At least 0.83	Lynn (1986)
Nine and above	At least 0.78	Lynn (1986)

Table 7: Recommendation of CVI Cut-off Value Based on Number of Experts.

However, the article has also adopted the modified kappa method to strengthen the agreement among the experts. Polit et al. (2007) recognised the drawback of Cohen's (1960) coefficient kappa (k) and introduced a modified kappa (k*), that can adjust each I-CVI for chance agreement. The kappa statistic represents the proportion of agreement remaining after a chance of agreement is removed. The evaluation for kappa values is categorised as follows: unacceptable (< 0.39), fair (0.40-0.59), good (0.60-0.74), or excellent (= k > 0.74). The formula for obtaining the CVI is shown in Table 8 below.

CVI Indices	Formula
I-CVI = content validity index of an item	I-CVI = number of experts giving a rating of 3 and 4 / total number of experts
S-CVI/ Ave = scale-level content validity index based on the average method	S-CVI/Ave = (sum of I-CVI scores) / (number of items) Or S-CVI/ Ave = (sum of proportion relevance rating) / (number of expert)
UA-CVI = content validity index of an entire instrument based on the	UA-CVI = (sum of UA scores) / (number of item)

universal agreement method	
The modified kappa (k^*)	Step 1: calculate the probability of chance of agreement: $p_c = (N! / (A! (N – A)!)) \times 0.5^N$ Step 2: calculate modified kappa: $k^* = I\text{-}CVI – p_c / 1 - p_c$

Table 8: Types of CVI Indices and Formula.
Source: Yusoff (2019).

The research instrument is inserted below, in the appendix.

4. RESULTS

The content validation of the instruments was conducted by a panel of three experts, as recommended in the guidelines by Lynn (1986). All three appointed experts returned their responses, resulting in a response rate of 100%. Based on the review of the experts, Tables 9 and 10 present the results from the CVI evaluation of the 54 items.

According to the recommendation by Polit et al. (2007), with a standard of .90 or higher for the S-CVI/Ave, the scale would be composed of some items on which there was a complete agreement (I-CVI = 1.00) and a few items on which there was a modest amount of disagreement (i.e., I-CVIs of at least .78). Altogether, 46 items demonstrated excellent content validity (I-CVI ≥ 1, k* ≥ 1), seven items were fair (I-CVI < 0.78, 0.40 ≤ k* ≤ 0.59), and one item was considered unacceptable or poor (I-CVI < 0.78, k* ≤ 0.39). It can be concluded that only one item, I-CVI, did not meet a satisfactory level, while both S-CVI/Ave and S-CVI/UA met satisfactory levels of acceptance, and thus, the scale of the questionnaire achieved a satisfactory level of content validity.

Items	Expert1	Expert2	Expert3	Expert in Agreement	I-CVI	UA
Q1	1	0	0	1	0.3	0
Q2	1	0	1	2	0.7	0
Q3	1	1	1	3	1.0	1
Q4	1	1	1	3	1.0	1
Q5	1	1	1	3	1.0	1
Q6	1	1	1	3	1.0	1
Q7	1	1	1	3	1.0	1
Q8	1	1	1	3	1.0	1
Q9	1	1	1	3	1.0	1

Q10	1	1	1	3	1.0	1
Q11	1	0	1	2	0.7	0
Q12	1	1	0	2	0.7	0
Q13	1	1	1	3	1.0	1
Q14	1	1	1	3	1.0	1
Q15	1	1	1	3	1.0	1
Q16	1	0	1	2	0.7	0
Q17	1	1	1	3	1.0	1
Q18	1	1	1	3	1.0	1
Q19	1	1	1	3	1.0	1
Q20	1	1	1	3	1.0	1
Q21	1	1	1	3	1.0	1
Q22	1	1	1	3	1.0	1
Q23	1	0	1	2	0.7	0
Q24	1	1	1	3	1.0	1
Q25	1	1	1	3	1.0	1
Q26	1	1	1	3	1.0	1
Q27	1	1	1	3	1.0	1
Q28	1	1	1	3	1.0	1
Q29	1	1	1	3	1.0	1
Q30	1	1	1	3	1.0	1
Q31	1	1	1	3	1.0	1
Q32	1	1	1	3	1.0	1
Q33	1	1	1	3	1.0	1
Q34	1	1	1	3	1.0	1
Q35	1	1	1	3	1.0	1
Q36	1	1	1	3	1.0	1
Q37	1	1	1	3	1.0	1
Q38	1	1	1	3	1.0	1
Q39	1	1	1	3	1.0	1
Q40	1	1	1	3	1.0	1
Q41	1	1	1	3	1.0	1
Q42	1	1	1	3	1.0	1
Q43	1	1	1	3	1.0	1
Q44	1	0	1	2	0.7	0
Q45	1	1	1	3	1.0	1
Q46	1	1	1	3	1.0	1
Q47	0	1	1	2	0.7	0
Q48	1	1	1	3	1.0	1
Q49	1	1	1	3	1.0	1
Q50	1	1	1	3	1.0	1
Q51	1	1	1	3	1.0	1
Q52	1	1	1	3	1.0	1
Q53	1	1	1	3	1.0	1
Q54	1	1	1	3	1.0	1
Proportion relevance	0.98	0.98	0.98	S-CVIAve	0.94	
The average proportion of items judged as relevant across the three experts			0.98	S-CVI/UA		0.85

Table 9: Content validity evaluation result.
Source: Authors.
a I-CVI= Number of experts rating the item either 3 or 4/total number of experts.

b S-CVI/Ave= Sum of the I-CVIs (I-CVI1+I-CVI2+I-CVI3+ …+I-CVI)/total number of items.
c S-CVI/UA= Number of items that achieved rating 3 or 4 by all experts/total number of items.
d Ave-proportion of agreement across experts = Proportion of agreement of each expert/total number of experts.
(I-CVI = content validity index of an item, S-CVI/Ave = Average of Content validity index of the entire instrument, S-CVI/ UA = Universal agreement on Content validity index of the entire instrument).

Items	Expert in Agreement	I-CVI	*Pc*	*k**	Evaluation
Q1	1	0.3	.375	-0.12	Unacceptable
Q2	2	0.7	.375	0.52	Fair
Q3	3	1.0	.125	1	Excellent
Q4	3	1.0	.125	1	Excellent
Q5	3	1.0	.125	1	Excellent
Q6	3	1.0	.125	1	Excellent
Q7	3	1.0	.125	1	Excellent
Q8	3	1.0	.125	1	Excellent
Q9	3	1.0	.125	1	Excellent
Q10	3	1.0	.125	1	Excellent
Q11	2	0.7	.375	0.52	Fair
Q12	2	0.7	.375	0.52	Fair
Q13	3	1.0	.125	1	Excellent
Q14	3	1.0	.125	1	Excellent
Q15	3	1.0	.125	1	Excellent
Q16	2	0.7	.375	0.52	Fair
Q17	3	1.0	.125	1	Excellent
Q18	3	1.0	.125	1	Excellent
Q19	3	1.0	.125	1	Excellent
Q20	3	1.0	.125	1	Excellent
Q21	3	1.0	.125	1	Excellent
Q22	3	1.0	.125	1	Excellent
Q23	2	0.7	.375	0.52	Fair
Q24	3	1.0	.125	1	Excellent
Q25	3	1.0	.125	1	Excellent
Q26	3	1.0	.125	1	Excellent
Q27	3	1.0	.125	1	Excellent
Q28	3	1.0	.125	1	Excellent
Q29	3	1.0	.125	1	Excellent
Q30	3	1.0	.125	1	Excellent
Q31	3	1.0	.125	1	Excellent
Q32	3	1.0	.125	1	Excellent
Q33	3	1.0	.125	1	Excellent
Q34	3	1.0	.125	1	Excellent
Q35	3	1.0	.125	1	Excellent
Q36	3	1.0	.125	1	Excellent
Q37	3	1.0	.125	1	Excellent
Q38	3	1.0	.125	1	Excellent
Q39	3	1.0	.125	1	Excellent
Q40	3	1.0	.125	1	Excellent
Q41	3	1.0	.125	1	Excellent
Q42	3	1.0	.125	1	Excellent
Q43	3	1.0	.125	1	Excellent
Q44	2	0.7	.375	0.52	Fair
Q45	3	1.0	.125	1	Excellent
Q46	3	1.0	.125	1	Excellent
Q47	2	0.7	.375	0.52	Fair

Q48	3	1.0	.125	1	Excellent
Q49	3	1.0	.125	1	Excellent
Q50	3	1.0	.125	1	Excellent
Q51	3	1.0	.125	1	Excellent
Q52	3	1.0	.125	1	Excellent
Q53	3	1.0	.125	1	Excellent
Q54	3	1.0	.125	1	Excellent

Table 10: The modified *kappa* content validity result.
Source: Authors.

a I-CVI, item-level content validity index.
b pc (probability of a chance occurrence) was computed using the formula for a binomial random variable, with one specific outcome: pc= [N! / A! (N A)!]*.5N where N ¼ number of experts and A ¼ Number agreeing on good relevance.
c k* ¼ kappa designating agreement on relevance: k* ¼ (I-CVI - pc) / (1 - pc)
d Evaluation criteria for kappa, using guidelines described in Cicchetti and Sparrow (1981) and Fleiss (1981): Poor = *k* of < .39; Fair = *k* of .40 to .59; Good = *k* of .60 –.74; and Excellent = *k* > .74.

5. CONCLUSIONS

Validating an instrument is a never-ending process. Even though there are many types of assessment to test the validity of instruments, the content validation process and CVI's are essential to validate the research instrument and should be treated and reported as equally necessary as other types of validation. The content validity index (CVI) is very robust in that it eliminates incongruity and provides precise interpretation, which helps to construct a more reliable and valid instrument design (Masuwai et al., 2016). Content validity deserves a rigorous assessment process as the obtained information from this process is valuable for the affirmation of the research instrument's quality.

The main contribution of this study was in elucidating the method to develop and validate instruments via CVI. Previous related studies have not adequately reported their methods to establish the instrument’s content validity. For instance, Beall and Boley (2021) claimed using content validity but failed to report the comprehensive process of instrument development in the context of ecotourism. Meanwhile, Pratama and Wulandari (2019) attempted to develop a measurement scale of tourism readiness and reported scale validity through Exploratory Factor Analysis (EFA) without reporting content validity assessment.

A meticulous process of content expert validation can avoid the problems at a later stage. However, novice researchers are at a disadvantage in operationalising the instrument validation process. The challenges include how to communicate the

information, collect feedback and analyse the data. This article is an attempt to design a standard format for the expert validation of a research instrument. Through a diligent literature review, the five-step process of expert review for validation has been introduced as a guide for future researchers, especially in the ecotourism setting.

It is recommended that the format is designed with a cover page inviting the experts, explaining their role, and introducing them to the research and the instrument. Information regarding the scale and the list of the scale items should be provided in the subsequent pages. Demographic questions are also included for validation. The expert review format will provide standard communication and feedback strategies between the researcher and the expert reviewers to help develop relevant, concise, and coherent research instruments. The expert will validate and make suggestions on the relevance, accuracy and the inclusion or deletion of an item, as well as minimise the effect of cultural sensitivity, bias or structural aspects such as language issues including double-barrelled, negative, confusing or leading questions.

Based on the result, it was confirmed that the questionnaire containing 54 items designed to assess domestic ecotourists in a protected area setting has high validity. As the three content experts had affirmed, the questionnaire exhibited excellent content validity, indicating the relevance of the various items in determining the actual study data. The overall content validity index of the instrument demonstrated a high I-CVI value of 1.0 and an S-CVI value of more than 0.90. Only one item was considered for deletion, as the result obtained was unacceptable. Thus, an instrument developer should decide whether to retain, modify, omit or add new items to the instrument, based on the results of the assessment. Moreover, the adoption of the modified kappa (k*) approach would enable novice researchers to strengthen the congruence chance agreement for each item.

It can be concluded that this instrument has established adequate and acceptable results. Therefore, prospective novice researchers may adopt the step-by-step procedure discussed in this article to develop the instrument's content validity for their studies. Future research can progress with multiple testing of the instruments for reliability and other forms of validity for better applicability.

Declarations of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is supported by the Ministry of Higher Education Scholarship, Department of Polytechnic Education and Community College (KPM.BBP.700-57/19/22 (2).

References

Ahn, J.; Back, K.J. Cruise brand experience: Functional and wellness value creation in the tourism business. *International Journal of Contemporary Hospitality Management,* Vol. 31, No. 5, 2019, pp. 2205-2223.

Aliman, N.K.; Hashim, S.M.; Wahid, S.D.M.; Harudin, S. Tourist expectation, perceived quality and destination image: Effects on perceived value and satisfaction of tourists visiting Langkawi Island, Malaysia. *Asian Journal of Business and Management,* Vol. 2, No. 3, 2014, pp. 212-222.

Almanasreh, E.; Moles, R.; Chen, T.F. Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, Vol. 15, No. 2, 2019, pp. 214-221.

Amos, D.; Boakye-Agyeman, N.A. Content validation of value-adding indicators for corporate real estate management: insights from a developing country. *Journal of Corporate Real Estate*, Vol. ahead-of-print No. ahead-of-print, 2022, pp. ahead-of-print.

Bahkia, A.S.; Awang, Z.; Rahlin. N.A.; Zulkifli, A.R. The Importance of Supportive Leadership in the Sewerage Operation Industry: A case of Indah Water Konsortium Private (IWK). *Humanities & Social Science Reviews*, Vol. 8, No. 3, 2020, pp. 149-162.

Beall, J.; Boley, B.B. An ecotourist by whose standards? Developing and testing the Ecotourist Identification Scale (EIS). *Journal of Ecotourism*, Vol. 21, No.2, 2021, pp. 99-120.

Beck, C.T. Content validity exercises for nursing students. *Journal of Nurse Education*, Vol. 38, No. 3, 1999, pp. 133-135.

Beckstead, J.W. Content validity is nought. *International Journal of Nursing Studies*, Vol. 46, No. 9, 2009, pp.1274-1283.

Boudreau, M.C.; Gefen, D.; Straub, D. W. Validation in information systems research: A state-of-the-art assessment. *MIS Quarterly*, Vol. 25, No. 1, 2001, pp.1-16.

Boynton, P.M.; Greenhalgh, T. Selecting, designing, and developing your questionnaire. *BMJ*, Vol. 328, No. 7451, 2004, pp. 1312-1315.

Castellanos-Verdugo, M.; Vega-Vázquez, M.; Oviedo-García, M.Á.; Orgaz-Agüera, F. The relevance of psychological factors in the ecotourist experience satisfaction through ecotourist site perceived value. *Journal of Cleaner Production,* Vol. 124, 2016, pp. 226-235.

Chen, G.; Zhao, L.; Huang, S. (2020). Backpacker identity: Scale development and validation. *Journal of Travel Research*, Vol. 59, No. 2, 2019, pp. 1-14.

Choe, J.Y.J.; Kim, S.S. Development and validation of a multidimensional tourist's local food consumption value (TLFCV) scale. *International Journal of Hospitality Management*, Vol. 77, 2019, pp. 245-259.

Churchill, G.A. Jr. A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, Vol.16, No.1, 1979, pp. 64-73.

Cohen, J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, Vol. 20, 1960, pp. 37-46.

Creswell, J.W. *Qualitative, quantitative and mixed methods approach*. Thousand Oaks, CA: Sage. 2014.

Cronbach, L.J.; Test Validation. In: Thorndike, R.L. (Ed.). *Educational Measurement*. Washington, DC: American Council on Education, 1971, pp. 443–507.

Davis, L.L. Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, Vol. 5, No. 4, 1992, pp. 194-197.

de Menezes, L. M.; Helliher, C. Flexible working and performance: A systematic review of the evidence for a business case. *International Journal of Management Reviews*, Vol. 13, No. 4, 2011, pp. 452–474.

DeVellis, R.F. *Scale Development: Theory and Applications*. Thousand Oaks, CA: Sage, 2012.

Dodds, S.; Hess, A.C. Adapting research methodology during COVID-19: lessons for transformative service research. *Journal of Service Management*, Vol. 32, No. 2, 2020, pp. 203-217.

Ehido, A.; Awang, Z.; Halim, B.A.; Ibeabuchi, C. Developing items for measuring the quality of work life among Malaysian academics: An exploratory factor analysis procedure. *Humanities & Social Sciences Reviews*, Vol. 8, No. 3, 2020, pp. 1295-1309.

Elangovan, N.; Sundaravel, E. Method of preparing a document for survey instrument validation by experts. *MethodsX*, Vol. 8, 2021, pp. 101326.

Eom, H.J.; Lu, Z.L. Establishing the measurement invariance in measures of consumers' perceived value. *Asia Pacific Journal of Marketing and Logistics*, Vol. 32, No. 5, 2019, pp. 1055-1069.

Gallarza, M.G.; Arteaga, F.; Del Chiappa, G.; Gil-Saura, I.; Holbrook, M.B.A. Multidimensional service-value scale based on Holbrook's typology of customer value. *Journal of Service Management,* Vol. 28, No. 4, 2017, pp. 724-762.

Gillham, B. *Developing a Questionnaire: Real World Research.* London: A&C Black, 2008.

Glyptou, K. Destination Image Co-creation in Times of Sustained Crisis. *Tourism Planning & Development*, Vol. 18, No. 2, 2020, pp. 166-188.

Gregori-Giralt, E.; Menéndez-Varela, J.L. The content aspect of validity in a rubric-based assessment system for course syllabuses. *Studies in Educational Evaluation*, Vol. 68, 2021, pp. 100971.

Grove, S. K.; Burns, N.; Gray, J. *The practice of nursing research: Appraisal, synthesis, and generation of evidence*. St. Louis, MO: Elsevier Health Sciences, 2013.

Gulliksen, H. Intrinsic validity. *American Psychologist*, Vol. 5, 1950, pp. 511–517.

Grant J.S.; Davis L.L. Selection and use of content experts for instrument development. *Research in Nursing & Health*, Vol. 20, No. 3, 1997, pp 269–274.

Gursoy, D.; Uysal, M.; Sirakaya-Turk, E.; Ekinci, Y.; Baloglu, S. *Handbook of Scales in Tourism and Hospitality Research*. London: CABI Publishing, 2015.

Gusenbauer, M. Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics*, Vol. 118, No.1, 2019, pp.177-214.

Haddaway, N.R.; Collins, A.M.; Coughlin, D.; Kirk, S. The role of Google Scholar in evidence reviews and its applicability to grey literature searching. *PloS one*, Vol.10, No. 9, 2015, pp. e0138237.

Han, H.; Yun, J.; Hyun, S.S. Nature based solutions and customer retention strategy: Eliciting customer well-being experiences and self-rated mental health. *International Journal of Hospitality Management*, Vol. 86, 2020, pp. 102446.

Huang, S.; Choi, H.S.C. Developing and validating a multidimensional tourist engagement scale (TES). *The Service Industries Journal*, Vol. 39, No. 7-8, 2019, pp. 469-497.

Kelley, D.L. *Measurement made accessible: A research approach using qualitative, quantitative and quality improvement methods*. Thousand Oaks, CA: Sage, 1999.

Kim, M.; Thapa, B. Perceived value and flow experience: Application in a nature-based tourism context. *Journal of Destination Marketing & Management*, Vol. 8, 2017, pp. 373-384.

Kim, H.; Kim, Y. G.; Woo, E. Examining the impacts of touristification on quality of life (QOL): The application of the bottom-up spill over theory. *The Service Industries Journal*, 2020, pp. 1-16.

Kim, J.H.; Ritchie, J.B. Cross-cultural validation of a memorable tourism experience scale (MTES). *Journal of Travel Research*, Vol. 53, No. 3, 2014, pp. 323-335.

Kim, H.; Woo, E.; Uysal, M. Tourism experience and quality of life among elderly tourists. *Tourism Management*, Vol. 46, 2015, pp. 465-476.

Koc, E.; Ayyildiz, A.Y. An overview of tourism and hospitality scales: Discussion and recommendations. *Journal of Hospitality and Tourism Insights*, Vol. ahead-of-print, No. ahead-of-print, 2021, pp. ahead-of-print.

Lam, K.W.; Hassan, A.; Sulaiman, T.; Kamarudin, N. Evaluating the Face and Content Validity of an Instructional Technology Competency Instrument for University Lecturers in Malaysia. *International Journal of Academic Research in Business and Social Sciences*, Vol. 8, No. 5, 2018, pp.363–381.

Lennon R.T. Assumptions underlying the use of content validity. *Educational and Psychological Measurement*, Vol. 16, No. 3, 1956, pp. 294–304.

Lewis, B.R.; Snyder, C.A.; Rainer Jr, R.K. An empirical assessment of the information resource management construct. *Journal of Management Information Systems*, Vol. 12, No. 1, 1995, pp. 199-223.

Li, T.T.; Liu, F.; Soutar, G.N. Experiences, post-trip destination image, satisfaction and loyalty: A study in an ecotourism context. *Journal of Destination Marketing & Management*, Vol. 19, 2021, pp. 100547.

Liu, S.; Lai, D.; Huang, S.; Li, Z. Scale development and validation of anime tourism motivations. *Current Issues in Tourism*, Vol. 24, No. 20, 2020, pp. 2939-2954.

Lynn, M.R. Determination and quantification of content validity. *Nursing Research*, Vol. 35, No. 6, 1986, pp. 382-386.

MacKenzie, S.B.; Podsakoff, P.M.; Podsakoff, N.P. Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, Vol. 35, No. 2, 2011, pp. 293-334.

Major, C.H.; Savin-Baden, M. *An introduction to qualitative research synthesis*. London: Routledge, 2010.

Mason, J.; Classen, S.; Wersal, J.; Sisiopiku, V.P. Establishing face and content validity of a survey to assess users' perceptions of automated vehicles. *Transportation Research Record*, Vol. 2674, No. 9, 2020, pp. 538-547.

Masuwai, A.M.; Saad, N.S. Evaluating the face and content validity of a Teaching and Learning Guiding Principles Instrument (TLGPI): A perspective study of Malaysian teacher educators. *Geografia*, Vol. 12, No. 3, 2016, pp. 11-21.

Mohamed Shaffril, H.A.; Samsuddin, S.F., Abu Samah, A. The ABC of systematic literature review: The basic methodological guidance for beginners. *Quality & Quantity*, Vol. 55, No. 4, 2020, pp.1319-1346.

Nunnally, J.C.; Bernstein, I.H. Validity. *Psychometric Theory*, Vol. 3, No. 1, 1994, pp. 99-132.

Okoli, C. A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems*, Vol. 37, No. 1, 2015, pp. 43.

Pickering, C.; Byrne, J. The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. *Higher Education Research & Development*, Vol. 33, No. 3, 2014, pp. 534–548.

Polit, D.F.; Beck, C.T. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, Vol. 29, No. 5, 2006, pp. 489-497.

Polit, D.F.; Beck, C.T.; Owen, S.V. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, Vol. 30, No. 4, 2007, pp. 459-467.

Pratama, A.B.; Wulandari, C. Villager readiness in rural tourism: a scale development using exploratory factor analysis. *International Journal of Tourism Policy*, Vol. 9, No. 3, 2019, pp. 251-261.

Prebensen, N.K.; Woo, E.; Chen, J.S.; Uysal, M. Motivation and Involvement as Antecedents of the Perceived Value of the Destination Experience. *Journal of Travel Research*, Vol. 52, No. 2, 2012, pp. 253–264.

Prebensen, N.K.; Woo, E.; Uysal, M. Experience value: Antecedents and consequences. *Current Issues in Tourism*, Vol. 17, No. 10, 2013, pp. 910–928.

Rasoolimanesh, S.M.; Dahalan, N.; Jaafar, M. Tourists' perceived value and satisfaction in a community-based homestay in the Lenggong valley world heritage site. *Journal of Hospitality and Tourism Management*, Vol. 2, 2016, pp. 72–81.

Rubio, D.M.; Berg-Weger, M.; Tebb, S.S.; Lee, E.S., Rauch, S. Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research*, Vol. 27, No. 2, 2003, pp. 94-104.

Sağ, I.; Zengul, F.D.; Avcı, K. Determining public's attitude toward medical tourism development: a scale proposal. *International Journal of Pharmaceutical and Healthcare Marketing*, Vol. 16, No. 2, 2022, pp. 243-259.

Sekaran, U.; Bougie, R. *Research methods for business: A skill building approach*. West Sussex: John Wiley & Sons, 2016.

Shaffril, H.A.M.; Ahmad, N.; Samsuddin, S.F.; Samah, A.A.; Hamdan, M. E. Systematic literature review on adaptation towards climate change impacts among indigenous

people in the Asia Pacific regions. *Journal of Cleaner Production*, Vol. 258, 2020, pp.120595.

Shahsavari, H.; Bakhshi, F.; Khosravi, A.; Najafi, Z.; Ghorbani, A. Perceived professional preparedness of senior nursing students' questionnaire: Development and psychometric evaluation. *Nurse Education Today*, Vol. 93, 2020, pp. 104533.

Shamsudheen, S.V.; Chowdhury, M.A.F. Salesforce product knowledge in Islamic financial institutions; development and validation of measurement scale. *Journal of Islamic Marketing,* Vol. 13, No. 5, 2022, pp. 1033-1049.

Shrotryia, V.K.; Dhanda, U. Content validity of assessment instrument for employee engagement. *Sage Open*, Vol. 9. No.1, 2019, pp. 1-7

Sireci, S.G. The construct of content validity. *Social Indicators Research*, Vol. 45, No. 1, 1998, pp. 83-117.

Souza, A.C.D.; Alexandre, N.M.C.; Guirardello, E.D.B. Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Serviços de Saúde,* Vol. 26, 2017, pp. 649-659.

Sy, M.; O'Leary, N.; Nagraj, S.; El-Awaisi, A.; O'Carroll, V.; Xyrichis, A. Doing interprofessional research in the Covid-19 era: A discussion paper. *Journal of Interprofessional Care*, Vol. 34, No. 5, 2020, pp. 600-606.

Taherdoost, H. Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. How to test the validation of a questionnaire/survey in a research. *International Journal of Academic Research in Management.* Vol. 5, No. 3, 2016, pp. 28-36.

Taheri, B.; Jafari, A.; O'Gorman, K. Keeping your audience: Presenting a visitor engagement scale. *Tourism Management*, Vol. 42, 2014, pp. 321-329.

Tay, L.C.; Tan, F.Y.; Yahya, K.K.; Rasli, A. Validation of corporate environmental citizenship measurement in Malaysia. *European Business Review*, Vol. 33, No. 2, 2021, pp. 345-360.

Thoyre, S.M.; Pados, B.F.; Park, J.; Estrem, H.; Hodges, E.A.; McComish, C.; Murdoch, K. Development and content validation of the pediatric eating assessment tool (Pedi-EAT), *American Journal of Speech-Language Pathology*, Vol. 23, No. 1, 2014, pp. 46-59.

Vada, S.; Prentice, C.; Scott, N.; Hsiao, A. Positive psychology and tourist well-being: A systematic literature review. *Tourism Management Perspectives*, Vol. 33, 2020, pp. 100631.

Wang, W.; Chen, J.S., Prebensen, N.K. Market analysis of value-minded tourists: Nature-based tourism in the Arctic. *Journal of Destination Marketing & Management*. Vol. 8, 2018, pp. 82-89.

Wang, Y.; Li, X. R. Developing and validating a scale of host territoriality in peer-to-peer accommodation. *Tourism Management*, Vol. 88, 2022, pp. 104425.

Williams, P.; Soutar, G. Value, satisfaction and behavioural intentions in an adventure tourism context. *Annals of Tourism Research*, Vol. 36, No. 3, 2009, pp. 413-438.

Xie, C.; Zhang, J.; Morrison, A.M. Developing a scale to measure tourist perceived safety. *Journal of Travel Research*, Vol. 60, No. 6, 2021, pp. 1232-1251.

Yusoff, M.S.B. ABC of content validation and content validity index calculation. *Education in Medicine Journal*, Vol. 11, No. 2, 2019, pp. 49-54.

Zamanzadeh, V.; Rassouli, M.; Abbaszadeh, A.; Majd, H.A.; Nikanfar, A.; Ghahramanian, A. Details of content validity and objectifying it in instrument development. *Nursing Practice Today*, Vol. 1, No. 3, 2014, pp. 163-171.

Appendix: Research Instrument

Construct	Item Statement	Source
Destination Image	I perceived this attraction to be well organized	Aliman et. al, 2016
	I perceived this attraction as easily accessible	Aliman et. al, 2016
	I perceived this attraction has adequate amenities	Aliman et. al, 2016
	I perceived this attraction as a safe destination	Glyptou, K. 2020
	I perceived this attraction as a friendly community	Glyptou, K. 2020
	I perceived this attraction has a peaceful atmosphere	Glyptou, K. 2020
	I perceived this attraction as an exciting destination	Li, Liu & Soutar 2021
	I perceived this attraction as attractive beautiful nature	Li, Liu & Soutar 2021
	I perceived this attraction can make people relax	Li, Liu & Soutar 2021
Perceived Value	This attraction is safe	
	This attraction is well organized	William & Soutar, 2009
	This attraction has an acceptable standard of quality	William & Soutar, 2009
	This attraction provides sufficient basic amenities	Rasoolimanesh et. al., 2016
	This attraction is easily accessible	
	The local community was hospitable	
	The entrance fee to this attraction is reasonable	Prebensen et. al. (2012)
	It is worthy for the money spent on activities in this attraction	Wang, Chen & Prebensen (2018)
	Visiting this attraction provides a great value for the money spent	
	Local foods and beverages in this attraction were reasonably priced	
	The handicraft sold were worth buying	Rasoolimanesh et. al., 2016
	This attraction makes me feel adventurous	Wang, Chen & Prebensen (2018)
	This attraction satisfies my curiosity	
	This attraction provides authentic experiences	
	This attraction provides a lot of things to do	
	The activity that I participate in this attraction makes me relax	Kim & Thapa, 2017
	The activity that I participate in this attraction makes me happy	Prebensen, Woo & Uysal, 2013
	The activity that I participated in this attraction relieved my tension	Ahn & Back, 2019

	The activity that I participate in this attraction boosts my energy The activity that I participate in this attraction benefits my physical health The activity that I participate in this attraction benefits my mental wellbeing The activity that I participate in this attraction helps me to forget my work-related activities for a moment	Gallarza et. al. 2017
Satisfaction	I am glad I decided to visit this attraction I feel happy having decided to visit this attraction This attraction has exceeded my expectations I think I did the right thing in choosing to visit this attraction Overall, I am happy with this visit Overall, I am satisfied with this visit	Castellanos-Verdugo et.al., 2016
Quality of Life	Overall, I am satisfied with my life Overall, I have achieved the most important things in my life Overall, my life is close to my ideal Overall, I have the best health conditions Overall, I feel good about my wellbeing	Kim, Kim & Woo, 2020
	Overall, I felt happy upon completion and participated in the activities	Han, Yu & Sean, 2020
	Overall, I have enriched my quality of life	Kim, Woo & Uysal, 2015

Source: Authors compilation.

Article info: Received 27/05/22. Accepted 18/08/22. Refereed anonymously.