International Journal of Behavioral Science
2017, Vol. 12, Issue 2, 19-30

ISSN: 1906-4675

Item Generation and Content Validation of the *Hajj* Crowd Behavior (HBC) Scale

Shukran Abdul Rahman[1], Nor Diana Mohd Mahudin[2], Zarina Mat Saad[3], Jasni Sulong[4], Zulkarnain Ahmad Hatta[4], Intan Hashimah Mohd Hashim[4], and Noraida Abdul Ghani[4]

Our main objective is the development and initial validation of a new scale to measure crowd behavior of individuals who performed *Hajj*, i.e., a holy, religious pilgrimage to Mecca that Muslims take during the Islamic month of *Dzul-Hijjah.* A scale that measures this construct is important because an understanding of the nature, domains, and roles of crowd behavior could facilitate the effective management and safety of *Hajj* pilgrims or *hujjaj*. Using the logical or rational approach, the scale development process was conducted in three phases. First, the construct of *Hajj* crowd behavior was identified and conceptualized via a literature review (Phase I). Next, an initial pool of 93 items covering three domains of crowd behavior (i.e., behavioral, affective, and cognitive) was generated from a series of semi-structured interviews with 23 *hujjaj* (Phase II). The content validity of these items was then ascertained by the agreement of subject matter experts (n = 15), who indicated whether an item is essential in measuring a particular domain of the *Hajj* crowd behavior construct. This expert review revealed that the majority of items were essential, relevant, and clear (i.e., Content Validity Ratio: CVR > .49), which resulted in 52 items in the final item pool. This revised scale is now suitable to be used in the test try-out phase and has the potential to inform strategies and design of crowd management. Further work is needed to assess its reliability and validity as well as to reduce the number of the items.

Keywords: *Hajj* crowd, conceptualization, scale development, item generation, content validity

Crowd behavior is a complex phenomenon that involves interaction between individuals within the crowd and the crowd environment. Many people join crowds in mass gatherings such as festivals, sporting events, malls, concerts, or religious congregations; hence, exposing them to dangers such as suffocation, stampedes, crush-related incidents, and heat-related illnesses (Still, 2014; Wijermans, 2011; Zeitz, Tan, & Zeitz, 2009). *Hajj*, which is a holy, religious pilgrimage to Mecca that Muslims take during the Islamic month of Dzul-Hijjah, is an example of such gatherings. This annual religious ritual has been deemed as the largest recurring mass gathering in the world, with more than 10 million pilgrims attending it (Al-Tawfiq et al., 2016). The number of *hujjaj* has gradually increased such that the Government of Saudi Arabia has to limit the number of visa issuance. In the year 2016, only 1,862,909 *hujjaj* have been granted the visa and permit to perform the ritual, which is in marked contrast with 3.1 million in 1998 (General Authority of Statistics, 2017).

Performing *Hajj* requires the pilgrims or *hujjaj* to be in specific locations to offer their worship at Kaa'ba and other prescribed sites. The presence of such a huge volume of *hujjaj* in these locations at the same time has caused various issues, including psychological and social ones among the pilgrims. For example, the psychological states of the pilgrims may lead to

[1] Associate Professor, International Islamic University Malaysia.
[2] Assistant Professor, International Islamic University Malaysia.
[3] Senior Lecturer, Universiti Utara Malaysia
[4] Associate Professor, Universiti Sains Malaysia.

certain behaviors (or absence of behaviors) that might be detrimental to other *hujjajs*, or worse.

Performing *Hajj* requires the pilgrims or *hujjaj* to be in specific locations to offer their worship at Kaa'ba and other prescribed sites. The presence of such a huge volume of *hujjaj* in these locations at the same time has caused various issues, including psychological and social ones among the pilgrims. For example, the psychological states of the pilgrims may lead to certain behaviors (or absence of behaviors) that might be detrimental to other *hujjajs*, or worse - could implicate negative, life-threatening consequences such as injury and mortality from crushing and asphyxiation. Therefore, understanding the important factors influencing crowd behavior among *hujjaj* as well as their psychological and social states is of paramount importance to develop appropriate crowd management framework and strategies, especially in dealing with emergency situations (Zeitz, Tan, & Zeitz, 2009).

At present, however, our current understanding of the underlying components or domains of crowd behavior remains unclear as there is little scientific work about the conceptualization and measurement of religious massed crowds, such as their perceptions, emotions, cognitions, and behavioral actions. The issue of measurement is particularly important when considering the management of crowds at potentially hazardous situations such as *Hajj*. Religious-based mass gatherings are known to be associated with various risks (Al-Tawfiq et al., 2016; Still, 2014), and if these risks are not adequately regulated and mitigated, they may lead to unpredictable behaviors in the crowd. Moreover, the behavior and movement during these gatherings are also relatively unique in that they involve a set of fixed rituals that need to be performed at specific hours and days in assigned locations and they are motivated by different goals in their activities, which, in turn, influence how the pilgrims respond to certain situations (AlGadhi & Mahmassani, 1991). Consequently, researchers have argued that efforts should be put to develop an accurate and reliable means to measure and predict crowd behavior (Hutton et al., 2011), particularly in religious mass gathering settings such as *Hajj*.

The development of a reliable and valid instrument to measure *Hajj* crowd behavior is important for two reasons. First, a new scale that produces a reliable and accurate measurement of *Hajj* crowd behavior could provide empirical evidence for conceptualization and the measurement of the construct; hence improving our understanding of how people behave, feel, and think in highly crowded situations. And second, information about the nature of *Hajj* crowd behavior and its components may contribute to informing strategies and design of crowd management. For example, the crowd management team could train *hujjaj* on the ways to relate with others or manage own and others' emotion in the crowd situation. The team could also design a guide for evacuation during emergencies. In other words, the training programs or interventions would be informed by the data on *hujjaj* behaviors, affective states, and cognitions in crowds. This, in turn, could improve the overall operation and coordination of Hajj management.

To address this call for a reliable and valid instrument for *Hajj* crowd behavior, we developed a new scale to measure this construct, offering another way of conceptualizing it from a psychosocial point of view. As this scale is still in the early stage of development, only the first three processes (i.e., scale conceptualization, item generation, and content validation) are reported. The need, importance, and approaches of the third process (i.e., content validation), require more explanation; therefore, the next section discusses this process.

The Importance of Establishing Content Validity

Researchers in behavioral sciences should adopt data collection tool that enables them to gather true information about the variable of interest (Wynd, Schmidt, & Schaefer, 2003). The tool should possess sound psychometric properties, and one of them is content validity. In general, content validity or logical validity refers to a degree of experts' agreement on whether a sample of behaviors has adequately represented a particular domain that accounts for the construct in question (Patrick et al., 2011; Hinkin, 1995; Bollen, 1989). This is an essential aspect of the validity concept because it concerns whether a scale has an adequate content coverage in relation to the items sampled to measure the construct.

Although some researchers have argued that the validity of a test or scale cannot be determined by estimating the content validity alone, the evidence of content validity still serves as important foundation in ensuring the construct validity of a measure (Sireci, 1998). The Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, and National Council of Measurement in Education, 2014) stated that content validity provides evidence on the extent to which a scale or test contains items that are appropriate for use to measure the intended construct. It can also clarify the presence of dimensions and elements of a construct as defined by theories (Polit, Beck, & Owen, 2007; Sekaran & Bougie, 2011).

There are two approaches in validating the appropriateness of items in a scale and the adequacy of a scale in covering the domains of content that account for the construct. The qualitative content validity requires the scale developer to get the expertise of judges who have knowledge on the subject matter. These experts' judgment can be analyzed by subjecting their views about the scale through quantitative analysis (Wynd, Schmidt, & Schaefer, 2003). This process is determined by getting a panel of subject matter experts (SMEs) who have expertise in the field of the study (Rubio et al., 2003) to judge on the contents of the scale and the essentiality of items in appropriately measuring a domain of the construct (Effendi, Matore, & Khairani, 2015; Johnson & Christensen, 2012). The SMEs will also provide constructive feedback on the quality of the overall scale (Effendi, Matore, & Khairani, 2015). In other words, the SMEs assist the scale developer to judiciously include items that are a sample of the domains in the construct and exclude items which are judged to be outside the domains.

A scale has a high content validity when there is a strong agreement among judges who rate the essentiality of items in measuring a domain that accounts for a construct (Donald, 2003). The content validity of a scale will be low if the scale contains items which are not essential in measuring a particular domain of the construct. A scale that lacks content validity at the item development phase has a higher likelihood of having a low relationship with the theoretical framework of the construct (Hinkin, 1995).

Methods and Results

The logical or rational approach (Cohen, Swerdlik, & Sturman, 2013) is used to develop the *Hajj* Crowd Behavior scale. This approach requires a review of existing scales and their psychometric properties, identification of relevant concepts or domains, generating items for the new scale, and validating the new scale (Clark & Watson, 1995). It also

considers the inclusion of items that are logically or rationally appear to correspond with the domains specified in the constructs. In this study, the concepts or domains of *Hajj* crowd behavior were conceptualized and identified by a literature review (Phase I). Next, item generation was conducted through semi-structured interviews with 23 *hujjaj* (Phase II). The items generated were then presented to 15 subject matter experts (SMEs) for content validation exercise (Phase III). Each of these phases is elaborated in the subsequent sections. Figure 1 summarizes the procedures followed in these phases.

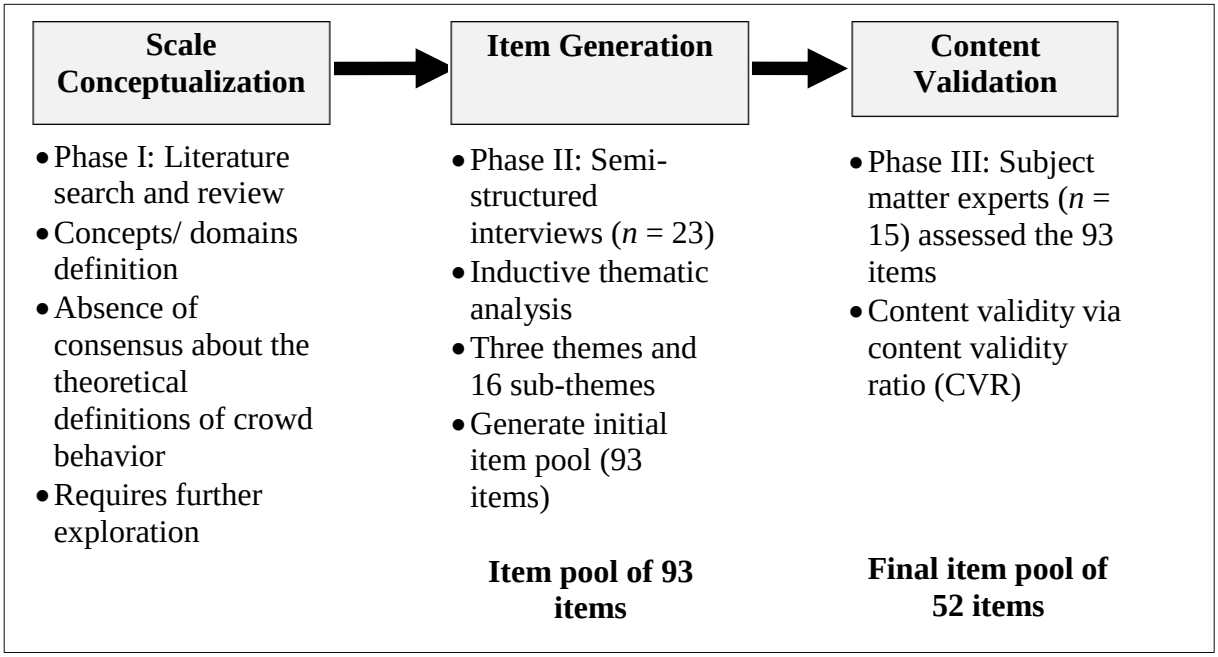

Figure 1. Steps undertaken to develop items for the new *H*ajj Crowd Behavior scale

Phase I: Scale Conceptualization

The first phase of the study involved a literature review that aimed at understanding how crowd behavior has been defined in the literature and how many components or domains it has. We also sought to review studies that have attempted to explore and define the specific construct in question (i.e., *Hajj* crowd behavior) and identifying its domains. Using sources such as existing crowd behavior scales, scholarly journals, books, government reports, and working paper series, our review of the literature indicates that there is no agreed upon definition of *crowd behavior* and its domains. Within the limited literature from the behavioral and psychological studies, we found that while some definitions focused on the number of individuals and the purpose of the crowd (e.g., Zeitz, Tan, & Zeitz, 2009), others concentrated on the social relationship among individuals in the crowd (e.g., Forsyth, 2014) or the idea of individuals having a collective mind, which is characterized by the presence of a common patterns of feelings, thinking, and actions (e.g., Wijermans, 2011; Zarboutis & Marmaras, 2005). In addition, while studies such as those by Hutton et al. (2011) and Pines and Maslach (1993) conceptualized crowd behavior as a function of crowd type and crowd mood, no such scale has been formed that captures its domains and attributes in a rigorous and psychometrically appropriate manner.

Furthermore, in the few existing scales of crowd behavior found in the review, we detect limited evidence for a comprehensive measure of crowd behavior in religious settings,

particularly *Hajj*. In other words, inspection of the literature led us to no specific scale that is used to measure *Hajj* crowd behaviors. Most of the scales were developed for use in settings different from the ones investigated in research on *Hajj*, such as sporting events, concerts, or malls. Given the research is to be conducted among *hujjaj*, it is imperative that the scale developed appropriately measures *Hajj* crowd behavior. Since the literature review revealed little information on the conceptualization and the domains of *Hajj* crowd behavior, the second phase of the study was carried out to address these issues.

Phase II: Item Generation

In this phase, a series of semi-structured interviews was conducted to gauge the views of the *hujjaj* on various aspects of crowd behaviors when they performed their *Hajj*. Similar to the aims in Phase I, the second phase sought to identify the domains of *Hajj* crowd behavior and outline the format of its scale. A total of 23 participants from Malaysia (Males = 10, Females = 13; age ranged from 14-50 years old) were interviewed to elicit their experiences in crowds at different *Hajj* sites in *Mecca* and *Madinah*. All interview sessions were audio-taped and transcribed verbatim. To protect their anonymity, each participant was assigned with a unique code identifier.

The semi-structured interview questions were developed using information from the literature review and consisted of broad questions about experiences in crowds while performing four *Hajj* activities (i.e., circumambulation of the Kaa'ba (*tawaf*), the passing between the hills of Safa and Marwa (*sa'i*), the vigil at 'Arafat (*Wuquf 'Arafat*), and the stoning of the '*devil*' at Mina. Using an inductive, bottom-up thematic analysis approach (Braun & Clarke, 2006), data were extracted into a spreadsheet and key themes were derived. Throughout this process, the four steps of conducting thematic analysis, i.e., data familiarization, initial coding, themes development, and theme definition and labelling (Howitt, 2013) were followed.

A total of three themes emerged. The first theme is the behavior of the crowd, which we refer to as the outward manifestation of observable behaviors perceived in a crowd. Within this theme, seven sub-themes were identified, namely (1) aggressive behaviors, (2) coping and helping others to cope behaviors, (3) defensive behaviors, (4) avoidance, (5) protective behaviors, (6) tolerance, and (7) hazardous acts manifested in a crowd. The second theme concerns the feelings or affective states of the *hujjaj*, which we define as one's expression of a subjectively experienced feeling or emotional state in a crowd. Five sub-themes, i.e., (1) positive, (2) negative, (3) positive comfortable, (4) negative comfortable, and (5) positive spiritual feelings or emotions in a crowd were also identified. Finally, the third theme includes thoughts or cognitions about the crowd. As a psychological construct, the cognitive aspect of behavior refers to what people think may be causal or related in some way to their behavior. Within this theme, four sub-themes emerged, i.e., (1) spiritual aspects of *Hajj*, (2) how *Hajj* should be managed, (3) negative thoughts of others, and (4) thoughts of safety in the crowd.

These themes and their sub-themes became the basis for the construction of item pool for the scale to measure *Hajj* crowd behavior. A blueprint of the scale (see Table 1) was developed to outline these themes and sub-themes, which are now called domains and sub-domains. A total of 93 items (i.e., Behavior = 34 items; Affective = 30 items; and Cognitive = 29 items) with five response options (1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree) was generated through this procedure. Meanwhile, Figure 2

illustrates the proposed framework for the construct. Following this framework, we define crowd behavior as the *outward manifested behaviors, emotion, and cognitive experiences of the individual crowd members* (in this case, the *hujjaj*).

Each domain is to be interpreted independently of the other domains. The total score for each domain is calculated by adding the responses to the items and the higher the scores, the higher the prevailing crowd behavior. For example, a higher sum of scores in the Behavior domain indicates a greater perceived behavior manifested in the crowd. Similarly, a higher sum of scores for Affective domain reflects a greater strength of the respective emotions as experienced in the crowd. Also, a higher total score of the Cognitive domain indicates the more cognitive aspect of crowd behavior the respondents reported. Having generated the items and conceptualized the construct, the *Hajj* Crowd Behavior scale is ready to be tested in the next phase of scale development, i.e., content validation.

Table 1

Blueprint of the Hajj Crowd Behavior Scale based on Phase I

Number	Domain	Definition	Sub-Domain	Number of Items
1	Behavior *Total items: 34*	The extent to which the *hujjaj* perceive their own (1) aggressive behaviors, (2) coping and helping others to cope behaviors, (3) defensive behaviors, (4) avoidance, (5) protective behaviors, (6) tolerance, and (7) hazardous acts manifested in a crowd.	Aggressive	12
			Coping and helping others to cope	5
			Defensive	5
			Avoidance	3
			Protective	2
			Tolerance	4
			Hazardous acts	3
2	Affective *Total items: 30*	The extent to which the *hujjaj* report their own experience of (1) positive, (2) negative, (3) positive comfortable, (4) negative comfortable, and (5) positive spiritual feelings or emotions in a crowd.	Positive feelings or emotions	7
			Negative feelings or emotions	7
			Positive comfortable feelings or emotions	7
			Negative comfortable feelings or emotions	3
			Positive spiritual feelings or emotions	6
3	Cognitive *Total items: 29*	One's thoughts, focusing mostly on what an individual think about (1) the spiritual aspects of *Hajj*, (2) how *Hajj* should be managed, (3) negative thoughts of others, and (4) thoughts of safety in the crowd.	Thoughts on spiritual aspects of *Hajj*	7
			Thoughts on how *Hajj* should be managed	10
			Negative thoughts of others	6
			Thoughts of safety in the crowd	6

In this phase, an expert committee consisted of 15 subject matter experts (SMEs) from two institutions that host local and international students were involved. These SMEs fulfilled

the following criteria: (1) familiar with *Hajj*-related matters; (2) have more than five years of academic experience of learning and teaching in Islamic studies and/or Psychology and/or Sociology at tertiary level education; and (3) have dealt with individuals of different cultural background. The SMEs' discipline areas helped them to judge the essentiality and appropriateness of the items. More specifically, the SMEs from Islamic studies would be familiar with *Hajj* rituals while the SMEs from Sociology and Psychology would have the knowledge of human behaviors, cognitive process, and emotions in social settings. Ten out of 15 SMEs have also performed *Hajj*.

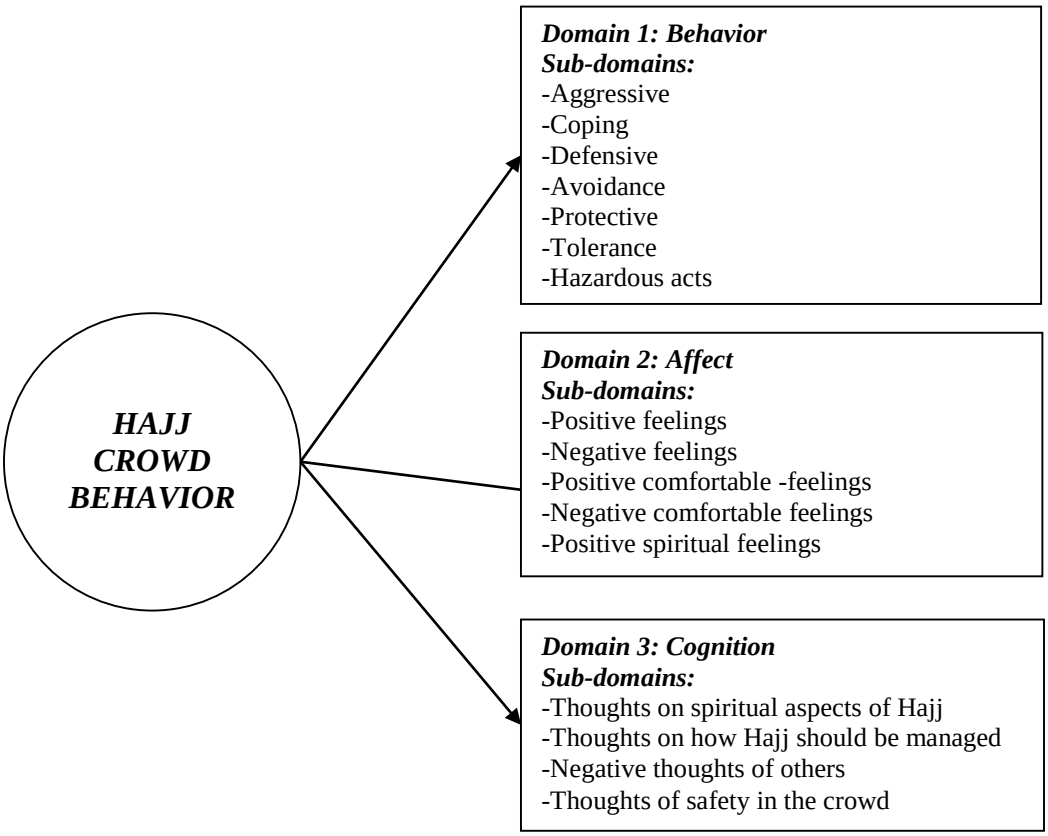

Figure 2. Proposed framework for the initial conceptualization of *Hajj* crowd behavior

The content validation exercise followed three steps. First, a Content Validity Ratio (CVR) form, which contains a pool of items that are arranged by construct, domains, and sub-domains were developed. This form provides information about the research, the need for the scale in the data collection, the structure of the scale, the constructs they measure, the corresponding domains and subdomains, and the number of items for each subdomain. The form also has sections on the definitions of the construct and its corresponding domains and subdomains, as well as instructions on the ways to respond to the items. The definitions served as an important reference for the SMEs to decide on the essentiality of each item and guided them to identify what was representative of the domains. There is also an instruction for the SMEs to write any comments on the items and overall scale.

The second step required the SMEs to provide their responses in the CVR form. In particular, they were asked to judge whether each item is *essential*, *useful but not essential*, or *not necessary* (Cohen, Swerdlik, & Sturman, 2013). They also have to indicate whether (1) the items are comprehensive with respect to the construct being measured; (2) the coverage of

the scale was adequate; (3) the statements used were appropriate in terms of word choice and suitability of the statements in relation to culture, gender, and language; and (4) whether there were other essential items on crowd during *Hajj* that were not included in the scale. The SMEs were also asked to provide additional and alternative items that they deemed necessary.

In the third step, the responses from the SMEs were analyzed to obtain the content Validity Ratio (CVR) of all items. This ratio was established by counting the number of SMEs giving *essential* rate to the items and subjecting the count of SMEs with *essential* judgment to the CVR calculation using Lawshe's (1975) formula: CVR= (*ne* - *N*/2)/ (*N*/2). *ne* refers to the number of panelists who rate the item as essential, while *N* refers to the total number of SMEs participating in the content validation exercise. Items are regarded as having some content validity if more than half of the panelists indicated them as *essential* and using this formula with the total number of 15 SMEs, the cut-off of 0.49 is required to determine whether an item should be retained or discarded. The results are presented in Table 2.

Table 2

Content Validity Ratio (CVR) Results

No	Domain	Sub-Domain	No. of Items	Rejected	Accepted		
					No. of Items	Items	CVR
1.	Behavior	Aggressive	12	10	2	I retaliated when being pushed	0.73
						I stepped on others	0.60
		Coping and helping others to cope	5	2	3	I advise people to be patient	0.73
						I help others to complete their rituals	0.60
						I read zikir to calm anxiousness	0.73
		Defensive	5	3	2	I brace dependents from being pushed	0.87
						I brace myself from being pushed	1.00
		Avoidance	3	3	0	-	-
		Protective	2	0	2	I safeguard own self.	0.87
						I safeguard my dependents.	0.73
		Tolerance	4	2	2	I respect others	0.73
						I tolerate while moving in crowd	0.87
		Hazardous acts	3	3	0	-	-
2	Affective	Positive feelings or emotions	7	2	5	I feel excited	1.00
						I am impressed	0.60
						I feel good	0.60
						I feel happy	0.73
						I feel inspired	1.00
		Negative feelings or emotions	7	5	2	I feel angry	0.73
						I feel comfortable	0.87
		Positive comfortable feelings or emotions	7	1	6	I feel easy	0.60
						I feel relaxed	0.73
						I feel calm	1.00
						I feel fresh	0.73
						I feel energized	0.87
						I feel able	0.60
		Negative comfortable feelings or emotions	3	1	2	I feel confused	0.60
						I feel scared	0.60

Table 2 *(Continued)*

No	Domain	Sub-Domain	No. of Items	Rejected	Accepted		
					No. of Items	Items	CVR
		Positive spiritual feelings or emotions	6	1	5	I feel patient	0.87
						I feel blessed	1.00
						I feel grateful	1.00
						I tried to be patient	0.60
						I put trust on Allah	0.87
3	Cognition	Thoughts on spiritual aspects of *Hajj*	7	2	5	I think the pilgrims should deal adversity with faith	0.73
						I think, the more difficult the test, the higher the rewards from Allah	0.73
						I think the pilgrims should remain patient	0.87
						I think that pilgrims should know the purpose of the pilgrimage	0.60
						I think about the difficulties faced by the Prophet	0.73
		Thoughts on how *Hajj* should be managed	10	1	9	I think a special wheelchair lane is necessary during Tawaf	0.60
						I think the authorities should regulate the cleanliness of toilets	0.73
						I think the authorities should provide more toilet facilities	0.87
						I think the pilgrims should plan ahead to meet at specific locations whenever they are separated	0.73
						I think the crowd should be adequately dealt with by the authorities	0.87
						I think there should be enough signs and directions for pilgrims	0.87
						I think there should be more active officers in crowd control	0.87
						I think more workers are needed in the event of emergency	0.87

Table 2 *(Continued)*

No	Domain	Sub-Domain	No. of Items	Rejected	Accepted		
					No. of Items	Items	CVR
						I think all countries should organize a better Hajj training programs	0.60
		Negative thoughts of others	6	4	2	I think others do not have discipline	0.60
						I think others do not bother about hygiene	0.60
		Thoughts of safety in the crowd	6	1	5	I worry about injuries	0.60
						I worry about stampede	0.87
						I worry about safety	0.73
						I worry about being separated from the loved ones	0.73
						I worry about the disease being spread	0.87

Discussion

This study describes the scale conceptualization, item generation, and content validation of a new measure to examine the self-reported crowd behavior of *hujjaj* when they performed the *Hajj* ritual, i.e., the *Hajj* Crowd Behavior scale (HCB). As a result of the literature review, semi-structured interviews, and subject matter experts' review, the *Hajj* Crowd Behavior scale now contains 52 items with three domains: Behavior, Affective, and Cognitive. The content validation exercise, in particular, has provided important information on the content validity of the HCB scale. To reiterate, 52 out of 93 items showed satisfactory CVR values, i.e., above 0.49. More specifically, a total of 23 items from the Behavior domain with CVR values below the cut-off value were discarded while 11 items above the cut-off value were retained. Ten items were discarded and 20 items were retained for the Affective domain, and finally, for the Cognitive domain, nine items were discarded and 21 items were retained.

The items included in this new scale covers a broad range of individual, social, and interpersonal factors that are relevant to understanding crowd behavior in religious settings. A few of these factors are currently not included in the existing measures of crowd behavior. For example, positive spiritual feelings or emotions are relatively unique to religious settings. In addition, thoughts on spiritual aspects of *Hajj* are also a new contribution to the existing crowd behavior measures. Another interesting finding is that all items in two sub-domains of Behavior, i.e., Avoidance and Hazardous Acts are not endorsed by the SMEs. It is likely that these items were deemed unnecessary because of their similarity to other items in the scale. In future studies, we will consider three options: (1) merge the similar items; (2) reword the items; or (3) discard the items.

It is important to note that this scale is still in development and requires further studies to assess its psychometric properties. Therefore, the next step in the development of this scale is to examine the reliability and validity and to reduce the number of the items. Nevertheless,

based on the findings, we believe that the *Hajj* Crowd Behavior scale is a research instrument that has potential to measure behaviors of a crowd. The content validity testing confirmed the validity and clarity of the scale developed, with evidence for the relevance of the items based on the considerable degree of agreement between opinions of the SMEs. The findings also suggest that the developed scale is suitable for studying crowd behavior due to its appropriate validity, simplicity, and functionality, particularly in *Hajj* setting. Furthermore, the current study represents the first attempt to establish a scale to measure crowd behavior among *hujjaj*. More assessment and deeper understanding of *Hajj* crowd behavior may have potential applications to improve crowd management and safety. To further refine the scale, it is recommended for future researchers to recognize the cultural and spiritual aspects of targeted participants in order to interpret the results accurately.

References

AlGadhi, S. A. H., & Mahmassani, H. (1991). Simulation of crowd behavior and movement: Fundamental relations and application. *Transportation Research Record, 1320*(1320), 260-268.

Al-Tawfiq, J. A., Gautret, P., Benkouiten, S., & Memish, Z. A. (2016). Mass gatherings and the spread of respiratory infections. Lessons from the *Hajj*. *Annals of the American Thoracic Society, 13*(6), 759-765.

American Educational Research Association, American Psychological Association, and National Council of Measurement in Education. (2014). *The Standards for Educational and Psychological Testing.* Washington, D.C.: American Psychological Association.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: John Wiley & Sons.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319.

Cohen, R. J., Swerdlik, M. E., & Sturman, E. (2013). Psychological Testing and Assessment: An Introduction to Tests and Measurement (8th Ed.). New York: NY McGraw-Hill Higher.

Donald, M. G. (2003). Handbook of self and identity. *Educational and Psychological Measurement, 51,* 755-765.

Effendi, M., Matore, E. M., & Khairani, A. Z. (2015). Assessing content validity of IKBAR among field experts in Polytechnics. *Australian Journal of Basic Applied Science, 7*, 255-257.

Forsyth, D. R. (2014). Group dynamics (6th ed.). Belmont, CA: Wadsworth Cengage Learning.

General Authority of Statistics. (2017). *The Number of Pilgrims for the Year 1995*. Retrieved June 14, 2017, from https://www.stats.gov.sa/en/28

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*(5), 967-988.

Hinkin, T. R., Tracey, J. B., & Enz, C. A. (1997). *Scale Construction: Developing Reliable and Valid Measurement Instruments.* Retrieved February 17, 2017, from Cornell University, School of Hotel Administration site: http://scholarship.sha.cornell.edu /articles/613

Howitt, D. (2013). *Introduction to Qualitative Methods in Psychology*. United Kingdom: Pearson.

Hutton, A., Zeitz, K., Brown, S., & Arbon, P. (2011). Assessing the psychosocial elements of crowds at mass gatherings. *Prehospital and Disaster Medicine, 26*(6), 414-421.

Johnson, B., & Christensen, L. (2012). *Educational Research Quantitative, Qualitative, and Mixed Approaches.* Alabama: SAGE Publications Inc.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*, 563–575.

Patrick, D. L., Burke, L. B., Gwaltney, C. J., Leidy, N. K., Martin, M. L., Molsen, E., & Ring, L. (2011). Content validity - Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2- assessing respondent understanding. *Value in Health, 14*(8), 978-988.

Pines, A., & Maslach, C. (1993). *Experiencing Social Psychology*. New York: McGraw-Hill.

Polit, D. F., Beck, C. T. & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing and Health, 30*, 459–467.

Rubio, D. M., Berg-Weger, M., S., Tebb, S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research, 27*(2), 94-105.

Still, G. K. (2014). *Introduction to Crowd Science*. UK: CRC Press.

Sekaran, U., & Bougie, R. (2011). *Research Methods for Business: A Skill Building Approach* (5th Ed.). New Delhi: John Wiley & Sons.

Sireci, S. G. (1998). The construct of content validity. *Social Indicators Research, 45*(1), 83–117.

Wijermans, F. E. H. (2011). *Understanding crowd behavior: Simulating situated individuals*. (Doctoral Dissertation). University of Groningen, SOM research school. ISBN: 978-90-367-4839-1.

Wynd, C. A., Schmidt, B., & Schaefer. M. A. (2003). Two quantitative approaches for estimating content validity western. *Journal of Nursing Research, 25*(5), 508-518.

Zarboutis, N., & Marmaras, N. (2005). Investigating crowd behavior during emergency evacuations using agent-based modelling. In *Proceedings of EAM*, 17-19.

Zeitz, K. M., Tan, H. M., & Zeitz, C. J., (2009). Crowd behavior at mass gatherings: A literature review. *Prehospital and Disaster Medicine, 24*(1), 32-38.