FEATURE ARTICLE

Revision and Psychometric Evaluation of the Diabetes Knowledge Questionnaire for People With Type 2 Diabetes

Julie A. Zuñiga,[1] Ya-Ching Huang,[2] So Hyeon Bang,[1] Heather Cuevas,[1] Tara Hutson,[1] Elizabeth M. Heitkemper,[1] Emma Cho,[1] and Alexandra A. García[1]

[1]School of Nursing, University of Texas at Austin, Austin, TX; [2]School of Nursing, Texas A&M University, College Station, TX

OBJECTIVE | Diabetes knowledge is associated with health, including lower A1C levels. The Diabetes Knowledge Questionnaire (DKQ-24), developed 30 years ago for Mexican Americans with type 2 diabetes and since used with diverse samples in many countries, contains outdated items that no longer accurately assess current knowledge needed for diabetes self-management. We revised the DKQ-24 and tested psychometric properties of the DKQ-Revised (DKQ-R) with a diverse sample.

METHODS | We conducted a five-phase instrumentation study as follows: *1*) DKQ-24 items were revised to reflect current diabetes care standards; *2*) the Delphi method was used to evaluate the DKQ-R's content validity ($n = 5$ experts); *3*) cognitive interviews were conducted with people with type 2 diabetes ($n = 5$) to assess their interpretations of DKQ-R items; *4*) cross-sectional administration of the DKQ-R to adults with type 2 diabetes was carried out to assess internal consistency reliability and convergent validity; and *5*) an item analysis was conducted using discrimination index and point biserial analysis.

RESULTS | After receiving the experts' feedback and conducting the cognitive interviews, 39 items were administered to 258 participants with type 2 diabetes (42.2% women; 29.1% Latino, 42.6% Asian, mean age 55.7 years). To select the final items, we considered the item discrimination index, as well as item-to-total correlations, content area, and participant feedback. The final 22-item DKQ-R uses the same yes/no/I don't know response format as the DKQ-24. The DKQ-R is strongly correlated with the DKQ-24 ($r = 0.71$, $P < 0.01$) and is weakly correlated with diabetes numeracy ($r = 0.23$, $P < 0.01$), indicating adequate convergent validity; a Kuder-Richardson-20 coefficient of 0.77 indicated good reliability.

CONCLUSION | The DKQ-R is a reliable and valid updated measure of diabetes knowledge for diverse populations with type 2 diabetes.

In the United States, an estimated 37.3 million people have diabetes, representing 11.3% of the population (1), and diabetes prevalence is expected to increase in the coming decades (2). To avoid diabetes-related complications (e.g., nephropathy, retinopathy, stroke, myocardial infarction, neuropathy, and amputations), patients must keep their blood glucose levels within a therapeutic range. Doing so requires multiple and complex self-care behaviors such as following a healthy eating plan, engaging in sufficient physical activity, being aware of symptoms, performing routine foot care, monitoring glucose levels, and taking medications (3–5).

Effective diabetes care starts with patients having sufficient knowledge about diabetes and typical diabetes management regimens (6). Although other qualities and expertise are needed for effective diabetes self-management (e.g., self-efficacy), having knowledge about diabetes is essential (6,7) and is associated with improved health outcomes, including lower A1C levels (6,7).

One often-used and well-known measure of patients' diabetes knowledge is the Diabetes Knowledge Questionnaire (DKQ-24) that was developed to measure Mexican American patients' knowledge about diabetes and its management in the Starr County Diabetes Education Study, a randomized controlled trial of a community-based educational intervention (8).

Unlike other diabetes knowledge tests such as the Michigan Diabetes Knowledge Test (9), the DKQ-24 does not include questions about the use of insulin as medication because

Corresponding author: Julie A. Zuñiga, jzuniga@nursing.utexas.edu

https://doi.org/10.2337/ds22-0079

most people with type 2 diabetes are not prescribed insulin to control their glucose levels and would not have received education about insulin dosing or side effects. Therefore, questions about insulin as a medication would not accurately reflect the majority of patients' diabetes knowledge, nor would they identify gaps in education that need to be addressed.

The DKQ-24 differs from other knowledge tests in its use of a true/false response format with yes/no answer options rather than multiple-choice options, which can be difficult to answer for people with low literacy levels (9,10). Thus, the DKQ-24 is more accessible to people with low literacy. The DKQ-24 items consist of a short statement to which the patient responds "yes," "no," or "I don't know." Items are checked against an answer key and recoded as correct or incorrect. "I don't know" is scored as incorrect. Correct responses are tallied, and a summed score is produced that can be converted to a percentage of correct responses (11). The DKQ-24 was shortened from 64 to 24 items in 1998 after psychometric testing (11). Since then, the DKQ-24 has been cited more than 440 times, translated into more than 30 languages, and used by clinicians and researchers throughout the world. As of now, it is still being used by researchers. The DKQ-24 has been a valid measure of changes in patients' knowledge about diabetes before and after receiving diabetes self-management education or as a point-in-time assessment of patients' diabetes knowledge (11).

In the 30 years since the DKQ-24 was developed, there have been many advancements in knowledge about diabetes and strategies for its management. Thus, some items on the DKQ-24 are no longer appropriate indicators of patients' knowledge about diabetes or diabetes care. For example, the item, "The best way to check my diabetes is by testing my urine" (8,11), for which the correct response was "no," reflects the standard of care in the early 1990s, before fingerstick blood glucose monitoring became common (12). Furthermore, the DKQ-24 contains several items that refer to a person with diabetes as a "diabetic" (8,11), which is not consistent with today's person-centered professional standards. These standards recognize that a person has a disease but is not only a diseased person (13). Moreover, some of the items in the DKQ-24 are "double-barreled," asking about two concepts within a single item, which is potentially confusing and could result in the item being scored as incorrect if the respondent knew part but not all of the content (14). Some DKQ-24 items are worded as double negatives, such as, "Eating sugar and other sweet foods is not bad for diabetics," which requires the respondent to think, "No, it is bad." Double negatives are confusing to both participants and response recorders, who might write the response as "no" without taking into account that the respondent actually meant to answer "yes" (14).

Although the DKQ-24 is reliable, valid, easy to administer, appropriate for people with low literacy levels, and widely used, its items needed to be reformatted and updated to reflect current standards. This article describes the steps taken to update this crucial instrument and test psychometric properties of the revised version. The goal of this instrumentation study was to retain successful elements of the DKQ-24, while creating a reliable, valid, and succinct measure of diabetes knowledge needed for effective self-management that would be appropriate for patients from diverse racial and ethnic backgrounds and with low literacy.

Research Design and Methods

This instrumentation study was conducted in five phases: *1*) revision of the DKQ-24 items to reflect current diabetes care standards and best practices for writing items; *2*) use of the Delphi process to evaluate content validity of the revised questionnaire (DKQ-R) by five experts; *3*) cognitive interviews with five participants with type 2 diabetes to assess their interpretation of DKQ-R items, leading to further revisions; *4*) cross-sectional administration of the DKQ-R with a sample of racially and ethnically diverse adults with type 2 diabetes to assess internal consistency reliability and convergent validity; and *5*) item analysis using discrimination index and point biserial analysis to select the best items. The methods reflect coronavirus disease 2019 (COVID-19) protocols and restrictions in place for patient and research staff safety.

Phase 1: Revision of DKQ-24 Items

The research team, composed of clinicians with expertise in diabetes care and education and researchers with expertise in survey development, reviewed and discussed the DKQ-24 items, coded them by topic, noted their relevance to diabetes care, and identified problems with item wording. The DKQ-24 included items about basic diabetes pathophysiology, which was considered good to know but unnecessary for effective diabetes self-management; thus, those items were eliminated. Gaps in self-management topic areas were identified and new questions were generated by research team members. Attention was taken to make items with as few words and syllables as possible to lower the reading level, consistent with guidelines for respondents with low literacy (15). The research team deliberated on each newly generated or reworded item. At the end of phase 1, the newly developed DKQ-R contained 62 items, many purposefully redundant, to be tested in subsequent phases so that, ultimately, the final DKQ-R would contain a relatively equal number of true/false items covering all diabetes self-management behaviors.

Phase 2: Evaluation of DKQ-R Validity

A Delphi study is a systematic process for a panel of experts to achieve consensus (16). Five experts in diabetes care (a senior nurse researcher, a registered nurse, an advanced practice clinical nurse specialist, a certified diabetes care and education specialist, and an endocrinologist) from different institutions who were not part of the research team were invited to provide feedback on DKQ-R items. Experts were offered $100 for each round of review. The experts provided feedback in two rounds of the Delphi process, with refinements made after each round. The experts agreed that the 39 revised items covered the breadth and depth of diabetes knowledge required for patients' diabetes self-management.

Phase 3: Cognitive Interviews

After the study was approved by two university institutional review boards, the DKQ-R was administered in one-on-one cognitive interviews to five adults with type 2 diabetes recruited via convenience sampling from a community-based clinic. Cognitive interview participants received $25 for their participation. Cognitive interviews are recommended by the National Center for Health Statistics as a way of exploring how respondents interpret test questions and derive responses (17). Thus, data from cognitive interviews consists of respondents' narrative statements about what they thought of the item, notes about whether the respondents seemed to struggle with the item wording or content, and participants' responses to the item itself. Two authors (J.A.Z. and Y.-C.H.) used a semistructured interview guide to conduct the interviews via computer videoconferencing. The researchers read each question to participants and asked them to say what the question meant in their own words or to explain what they understood about the question. The interviewees were also asked to describe how they came up with their answers to the questions. After five interviews, participants' feedback was consistent, reaching saturation. After compiling responses into a spreadsheet, the research team revised some items for clarity and eliminated several items based on redundancy or participants' negative experiences with the items. At the end of phase 3, the DKQ-R consisted of 39 items covering the breadth of self-care behaviors.

Phase 4: Assessment of Reliability and Validity in a Diverse Sample

We conducted a cross-sectional online survey to administer the DKQ-R and evaluate its psychometric properties. Eligible participants were adults ≥18 years of age who were diagnosed with type 2 diabetes and were able to read, comprehend, and respond to questions in English, Spanish, or Mandarin Chinese.

The study was advertised through Facebook; the University of California, San Francisco, Collaborative Approach for Asian Americans, Native Hawaiians, and Pacific Islanders Research and Education research participant registry; churches in two adjacent cities; and a research recruitment company. People interested in the study contacted the researchers (J.A.Z. and Y.-C.H.) by e-mail, and the researchers provided information about the study's purpose and screened potential participants against the inclusion criteria. This type of recruitment can result in bias against individuals with low computer literacy. However, at the time of data collection, we were unable to recruit participants from clinics because of COVID-19 pandemic restrictions, which did not allow research team members in clinic waiting rooms.

Recruited individuals who were identified as eligible for the study were sent a link to an online survey. After logging on, respondents completed screening questions that confirmed their eligibility, documented informed consent, and proceeded with the survey. To prevent fraudulent responses, researchers allowed only one survey for each Internet protocol address and examined geolocations, time to complete the survey, and e-mail addresses to ensure that they were logical. Responses from Internet robots (bots) were identified by their similar e-mail addresses and geolocation information and extremely fast completion times. Survey responses that followed such patterns were excluded. Additionally, the survey included attention check questions with obvious answers to minimize bot responses. For example, a question might ask, "Please select 'I don't know' as the answer for this question." Human respondents were expected to answer the question correctly, and an incorrect answer might result in exclusion from analysis. Participants completed the 20-minute survey in English, Spanish, or Mandarin Chinese and received a $25 electronic gift card for completing it.

Instruments

The DKQ-R and other study measures were translated from English to Mandarin Chinese and back-translated from Chinese to English by two separate translators who were native speakers and had graduate degrees in health care fields. The second author (Y.-C.H.) reviewed the final translations to compare the original English to the back-translated English versions to ensure that the items were conceptually equivalent. The instruments were translated into Spanish by a professional, certified, native-speaking translator who preserved conceptual equivalence (18). The first and senior

authors (J.A.Z. and A.A.G.) reviewed the Spanish version for equivalence to the English version.

Demographic and clinical data included participants' self-reported age, sex, race/ethnicity, marital status, education level, time since type 2 diabetes diagnosis, prescribed diabetes medications, and diagnosis with any of the following nine medical conditions commonly experienced by people with diabetes and older adults: hypertension, heart disease, stroke, cancer, arthritis, hepatitis, kidney disease, asthma, and chronic obstructive pulmonary disease (19).

The DKQ-24 was administered for comparison with the DKQ-R. The DKQ-24 comprises 24 questions about basic diabetes knowledge, with response choices of "yes," "no," and "I don't know." Responses are checked against an answer key and scored as correct or incorrect. Total scores range from 0 to 24. Higher scores indicate more diabetes knowledge (11). The DKQ-R consisted of 39 questions after phase 3. The response options matched the original DKQ-24. With more items, the sum of correct scores could range from 0 to 39.

A five-item version of the Diabetes Numeracy Test (the DNT-5) was administered so that a correlation could be calculated to assess for convergent validity with the DKQ-R. Diabetes numeracy refers to arithmetic skills such as addition, subtraction, multiplication, division, fractions, and decimals, as well as numerical hierarchy, inference skills, and multistep calculations that patients use in diabetes self-management (20). The DNT-5's items were selected from the longer 15-item DNT-15 (20) and are formulated as word problems related to diabetes treatment. A sample question is, "1/2 cup of potatoes counts as 1 carbohydrate choice. How many choices does 2 cups of potatoes count as?" (20). Because the majority of patients with type 2 diabetes are treated with oral medications rather than insulin (21), we substituted an item about insulin dose calculation with one about oral diabetes medication. Each correctly answered question received a score of 1; total scores range from 0 to 5. Higher total scores indicate greater diabetes-related numeracy (21).

Phase 5: Item Analysis

All data analyses were performed using IBM SPSS Statistics, v. 26, software. Two-tailed P values ≤ 0.05 determined significance. In addition to descriptive analyses, Pearson correlations were conducted to assess convergent and construct validity among the DKQ-24, DKQ-R, and DNT-5 total scores. The Kuder-Richardson Formula 20 coefficient was used to assess internal consistency reliability for the DKQ-24 and DKQ-R (22). Discrimination indices (ranging from $+1$ to -1) indicate how much a correct response on an item correlates with scores on the overall scale. Items with negative or zero discrimination scores do not contribute to the overall test score. Positive item discriminations are desirable, but a discrimination index close to 1.0 implies that the item may be duplicative of other items on the test (23). We used Truman Kelley's "27% of sample" group size rule to compute the discrimination index as the number of correct responses on the item in the low-scoring 27% of the sample subtracted from the number of correct responses on the item in the high-scoring 27% of the sample, dividing the difference by the sample size. Discrimination indices of ≥ 0.4 are regarded as highly discriminating and those <0.2 as minimally discriminating (24). We examined point biserial correlations, which are measures of item-to-total score correlation. Similar to the discrimination index, item-to-total correlations are indications of item difficulty. Scoring correctly on an individual item should correspond to the total score; higher correlations (point biserial correlation >0.25) signify that people who responded correctly to the item scored higher on the total questionnaire (25–27). We also considered the percentage of correct responses on each item as a measure of difficulty (23), item content area, and participants' subjective feedback. A Flesch-Kincaid reading ease score was calculated for the English version to test readability by assessing the number of syllables for words and the number of words per sentence. The readability score equates to school grade levels 1–12 (28). The Fernández Huerta calculator, adapted from the Flesch-Kincaid formula, was used to assess readability of the Spanish version. Scores can range from 0 to 100, with higher scores indicating greater reading ease; scores were converted to school grade reading levels for comparison with the English version (29).

Results

The sample included 258 participants from diverse racial and ethnic groups (23% non-Hispanic White, nearly 4% Black, 29% Latino, nearly 43% Asian, and 1% Native Hawaiian or Native American) (Table 1). About two-thirds were foreign-born. More than half took the survey in English, one-fourth in Spanish, and one-fifth in Mandarin Chinese. Participants were, on average, 56 years of age, and more than half were women and married or partnered with a significant other. Most participants were highly educated and had both medical and dental insurance. Three-fourths of the participants reported that their income met their family needs. The average length of participants' diabetes duration was 11 years. Nearly all participants took oral antidiabetic medication, about one-third were prescribed insulin, and one-fourth used noninsulin injectable medications. Most participants had ≥ 1 (mean 1.48 ± 1.26) comorbid chronic conditions, with hypertension (60.5%), arthritis (28.5%), and kidney

TABLE 1 Descriptive Characteristics of Sample (N = 258)

Variables	Mean ± SD or %	Range
Ethnicity/race		
Non-Hispanic White	23.3	–
Black	3.9	–
Latino	29.1	–
Asian	42.6	–
Native American or Native Hawaiian	1.2	–
Foreign born	64.0	–
Age, years	55.71 ± 11.55	22–83
Female sex	57.8	–
High school or greater education	93.4	–
Married or with a significant other	62.0	–
Income met family need	76.3	–
With health insurance	89.8	–
With dental insurance	64.2	–
Duration of diabetes, years	11.04 ± 9.26	0–42
Diabetes treatment methods		
Oral medication	93.7	–
Insulin	34.1	–
Noninsulin injectable medication	24.5	–
Number of comorbidities	1.48 ± 1.26	0–5

problems (15.5%) having the highest prevalence. The mean score for the DKQ-24 was 17.7 ± 3.5 (73.8%). The mean score for the DNT-5 was 3.7 ± 1.3.

The Flesch-Kincaid reading ease score indicated a sixth-grade reading level for the English version (28). The Fernández, Huerta calculator score was 76.01, indicating a somewhat easy reading level (about a seventh-grade level) (29). The tally of correct scores on the DKQ-R ranged from 40.7 to 95%. We analyzed DKQ-R items in each language and retained items that discriminated well in all three languages. We selected one or two items that were most discriminating among the several for each topic.

Every point biserial correlation was in the expected direction and >0.25. The range of discrimination indices was 0.13–0.67. Two items had a discrimination index <0.2: items related to the need for an annual eye exam and those about the acceptability of eating sugar-free foods, which scored 0.13 and 0.16, respectively. Those items were retained because of their importance to diabetes self-management. Diabetes educators and clinicians need to identify patients who lack foundational knowledge about preventing diabetes-related complications. Seventeen other items had low discrimination indices (<0.20) across the three language versions. Those 17 items were removed, leaving a final total of 22 items. Fifteen items in the final DKQ-R had a discrimination index >0.40, indicating very good discrimination (Table 2) (24).

The Kuder-Richardson Formula 20 coefficient for the 22-item DKQ-R was 0.77, and reliability coefficients were similar among English, Spanish, and Chinese versions, at 0.80, 0.72, and 0.75, respectively. The DKQ-R was strongly correlated with scores on the DKQ-24 (r = 0.71, P <0.001) and positively correlated with scores on the DNT-5 (r = 0.23, P <0.001) and with education level (r = 0.19, P <0.01) (Table 3).

The range of total correct DKQ-R scores was 40.7–95.0% (Table 2). The three items with the most correct responses (i.e., least difficult items) were about the need to have an annual eye exam (95% correct), the ability of high glucose levels to damage the kidneys (94.2%), and the amount of fiber in white bread compared with whole-wheat bread (90.7%). The three items with the most incorrect responses (i.e., most difficult items) were about the symptoms of high blood glucose (59.3% scored incorrectly), the meaning of an A1C level (46.5% incorrect), and whether it is safe for people with diabetes to wear flip-flop shoes (42.3% incorrect).

Discussion

The purpose of this five-phase instrumentation study was to revise the DKQ-24 such that items were consistent with current diabetes self-management standards, item development followed best practices, and the revised questionnaire was reliable and valid with a racially and ethnically diverse sample. After a methodical process of developing and pilot-testing items, we translated the DKQ-R into three of the most common languages used in the United States (30) and administered it in an online survey to 258 adults with type 2 diabetes. From the pool of 62 possible items, 22 items that cover high-priority content for diabetes self-management as recommended by the American Diabetes Association (ADA) (5) were selected based on adequate response discrimination and item-total correlations.

The DKQ-24 was developed for Spanish- or English-speaking Mexican Americans with type 2 diabetes living along the U.S.-Mexico border (8). The DKQ-R was developed to meet the need for a reliable and valid diabetes knowledge questionnaire for a more diverse patient populaton. The DKQ-R's 22 items discriminated well and were internally consistent in English, Spanish, and Mandarin Chinese. Some items were included in the DKQ-R even if the majority of participants answered them correctly because the content was deemed important for all people with diabetes.

DKQ-R items are written with as few words and syllables as possible to reduce their reading difficulty. The DKQ-R attained a sixth-grade reading level, as indicated by its Flesch-Kincaid score. The grade level would have been

TABLE 2 Items, Answer Key, Correct and Incorrect Responses, Discrimination Index, and Item-to-Total Correlations on DKQ-R ($N = 258$)

DKQ-R Item	Answer Key	Correct Responses, %	Incorrect Responses, %	Discrimination Index	Item-to-Total Correlation*
The body needs insulin to use the sugar that a person eats.	Yes	77.9	22.1	0.37	0.37
When someone has diabetes, their children are more likely to get diabetes.	Yes	76.0	24.0	0.34	0.34
A fasting blood glucose level of 100 is too low.	No	88.8	11.2	0.31	0.50
People with high glucose have problems with blood circulation in their legs.	Yes	89.9	10.1	0.31	0.43
People with high glucose can have damage to the kidneys.	Yes	94.2	5.9	0.21	0.51
Shaking and sweating are signs of high blood glucose.	No	40.7	59.3	0.59	0.41
It is safe for people with diabetes to wear flip flops.	No	57.8	42.3	0.47	0.40
Vaccines are highly recommended for people with diabetes.	Yes	74.8	25.2	0.43	0.40
People with diabetes should see the dentist twice a year.	Yes	75.6	24.5	0.40	0.36
Frequent urination and thirst are signs of low blood glucose.	No	62.8	37.2	0.67	0.56
A person with diabetes can eat as much as they want if it is sugar-free.	No	86.8	13.2	0.16	0.27
A1C measures a 3-month average of blood glucose levels.	Yes	82.2	17.9	0.36	0.46
After eating lunch, a glucose level of 200 is too low.	No	88.0	12.0	0.31	0.49
Stress lowers blood glucose levels.	No	81.0	19.0	0.39	0.50
Smoking will raise the chances of diabetes complications.	Yes	81.0	19.0	0.43	0.49
People with diabetes should increase the amount of trans-fat in their diets.	No	69.4	30.6	0.49	0.44
An A1C of 120 is good	No	53.5	46.5	0.59	0.44
Someone with diabetes is more likely to get heart disease.	Yes	88.8	11.3	0.30	0.46
Feeling angry can raise blood glucose levels.	Yes	65.9	34.1	0.47	0.41
Cooking with lard is healthier than cooking with vegetable oil.	No	81.0	19.0	0.31	0.39
People with diabetes should have their eyes checked every year.	Yes	95.0	5.1	0.13	0.36
White bread has more fiber than whole-wheat bread.	No	90.7	9.3	0.23	0.36

*Each item was significantly correlated to the total DKQ-R ($P < 0.001$).

lower given the numbers of syllables and words per sentence (28) except for the frequent use of the four-syllable word "diabetes." Had a single-syllable synonym for "diabetes" been available in English, the Flesch-Kincaid score would have indicated a fifth-grade reading level. The Fernández Huerta calculation indicated a similar level of difficulty for the Spanish-language version. The DKQ-R can be read aloud, either in one-on-one interviews or using read-aloud software, for people with lower reading levels.

Higher scores on the DKQ-R indicate more diabetes knowledge. Cutoff scores representing adequate diabetes knowledge have not been set. The range of correct scores on the DKQ-24 with Mexican Americans in the Starr County Diabetes Education study was 14–96% (11). The range of correct scores for the DKQ-R in this study was less variable (40.7–95%). Future studies could compare the change in scores before and after diabetes education and intervention.

TABLE 3 Correlations Between the DKQ-R, DKQ-24, DNT-5, and Education Level

	DKQ-R	DKQ-24	DNT-5
DKQ-R	1.0		
DKQ-24	0.71†	1.0	
DNT-5	0.23†	0.23†	1.0
Education level	0.19*	0.20*	0.48

*P <0.01. †P <0.001.

The number of correct responses on the DKQ-R compared with the DKQ-24 is likely higher because our sample was well educated; nearly all participants had at least a high school education. Education level is correlated with DKQ-R scores. It is possible that we recruited a sample with high levels of education because the survey was administered online, and people with more education may have more computer literacy or resources to access online surveys. With a less literate sample, the DKQ-R could be administered in person, eliminating the need for computer access (11).

It is to be expected that samples with less formal education would be less likely to answer knowledge test items correctly because people with more formal education are more experienced in completing knowledge tests and may have acquired information about diabetes informally in the education setting. It is likely that scores on the DKQ-R would also correlate with having recent diabetes education or frequent contact with diabetes care providers; however, we did not collect those data.

Almost all of the items discriminated those who scored well from those who did not. Discrimination indices were >0.20 for all but two items, and nine items had discrimination indices >0.40, indicating that they were highly discriminating items (24). All of the items were well correlated with total scores, another indicator of discrimination. Three items were retained in the DKQ-R despite having a discrimination index <0.2 or being answered correctly by >94% of the participants. These items were retained because they represent essential content for diabetes self-management and prevention of diabetes-related complications (5). It is possible that people with less formal education or less diabetes education would score lower on these items.

Nearly all participants in our study correctly answered items about the need to have an annual eye exam, the ability of high glucose levels to damage the kidneys, and the relative amount of fiber in white bread compared with whole-wheat bread. It is possible that most people with diabetes would be familiar with this content, having learned it from their health care providers or through informal channels such as mass media (e.g., magazines, television, or social media campaigns) or from family and friends.

Our participants were much less likely to respond correctly to questions about symptoms of high blood glucose, whether an A1C level of 120 was good, and whether it is safe for people with diabetes to wear flip-flop shoes. We purposely wrote the item about an A1C level of 120 because people often mistake A1C for blood glucose levels. Responding that an A1C level of 120 is good shows a clear knowledge deficit regarding the difference between A1C and blood glucose levels. Of the four A1C questions tested, this item was the most discriminating. Although participants answered these four items incorrectly, they may have previously discussed these topics with health care providers without retaining the information. Furthermore, content about A1C levels or about types of footwear is less likely to be conveyed in mass media messages. According to our item developers and Delphi process experts, these are areas of confusion for people with diabetes.

Another consideration regarding which items received the most correct or incorrect responses is the format of the response options. The DKQ-R response choices are "yes," "no," and "I don't know." Respondents' answers are compared with an answer key before the items are recoded for analysis as correct or incorrect. The response "I don't know" is always coded as incorrect. Including "I don't know" as a response choice lessens the chance of choosing the correct response by chance and is a strength of the DKQ-R (23,25,31). By design, the number of items for which the correct response is "yes" is equal to the number for which the response is "no." The correct response for the three items most often answered incorrectly was "no," perhaps because some respondents may be biased against choosing "no" as a response (32). This may be a limitation of the response choice format compared with a multiple-choice format. However, multiple choice tests are more burdensome to test-takers and especially to those with lower literacy levels (23).

Based on the improvements in the item wording and content and the evidence for its reliability and validity, clinicians and researchers should adopt the DKQ-R and discontinue using the DKQ-24. Clinicians can use the DKQ-R to quickly assess patients' knowledge and identify gaps to be addressed as part of diabetes education. The 22-item DKQ-R takes participants <20 minutes to complete, and scoring it is simple. The questions are in line with diabetes self-management behaviors outlined by the ADA (5). Clinicians could use either missed items or total score to tailor education or use the instrument as a pre-/post-test to assess diabetes education classes.

Despite the research team's stringent standards to maintain rigor in the study, there are some possible limitations to the findings. Collection of online survey data are associated with a risk of interference by bots or by one person taking the survey multiple times, presumably for the $25 electronic gift card incentive. A protocol was created to ensure that inaccurate data were excluded uniformly, and this process included confirmation of inclusion of participants by two research team members. Consensus did not need to be reached; if either research member believed the survey was taken using a fake account, it was excluded from analysis. As part of the protocol, participants' responses were checked for uniqueness. For example, one bot wrote the wrong answer with the same incorrect spelling across all of its responses, and all of those surveys were excluded. Although the process was rigorous, it is possible that some participants did not answer truthfully (e.g., that they did not actually have diabetes) or did not answer mindfully. Additionally, using online methods to recruit participants and collect data resulted in the enrollment of a relatively very well-educated sample. It is possible that participants with lower education and income levels would perform differently on the DKQ-R. Therefore, future research should test the DKQ-R in samples with lower education and income levels and shorter durations of diabetes.

The DKQ-24 was strongly correlated with A1C, which added additional evidence for its validity based on the assumption that people with more knowledge of diabetes would have better glycemic control (11). However, in this study, we were unable to collect blood samples because of COVID-19 pandemic restrictions and the need for socially distant online recruitment and data collection. Thus, evidence for validity would be strengthened through future research demonstrating a relationship between DKQ-R score and A1C. Additional validity evidence could be gathered by using the DKQ-R as a pre-/post-test measure of diabetes education. If scores change significantly in the expected direction compared with people randomly assigned to a waitlist or control group, DKQ-R administrators could have more confidence in its usefulness.

Conclusion

The DKQ-R is an updated, reliable, and valid measurement of diabetes knowledge available in three languages for use with adults of different races and ethnicities who have type 2 diabetes. The revised instrument and answer key are shown in Table 2 and can be used without an access fee. Clinicians and researchers should begin to use the DKQ-R in place of the 1998 DKQ-24 to quickly assess patients' diabetes knowledge and identify gaps to be addressed with diabetes self-management education.

ACKNOWLEDGMENTS

The authors acknowledge the previous authors of the DKQ-24, especially Dr. Sharon Brown and Evangalina Villagomez.

FUNDING

Funding for this project was provided by the Texas State University Research Enhancement Program and a University of Texas at Austin's Vice President of Research Creativity Grant.

DUALITY OF INTEREST

No potential conflicts of interest relevant to this article were reported.

AUTHOR CONTRIBUTIONS

J.A.Z. contributed to the development of the instrument, wrote the manuscript, and researched data. Y.-C.H. contributed to the development of the instrument, conducted the analysis, wrote the manuscript, and researched data. S.B. conducted the analysis, researched data, contributed to discussion, and reviewed/edited the manuscript. H.C., T.H., E.H., E.C., and A.A.G. contributed to the development of the instrument, contributed to discussion, and reviewed/edited the manuscript. J.A.Z. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

1. Centers for Disease Control and Prevention. *National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States*. Available from https://www.cdc.gov/diabetes/data/statistics-report/index.html. Accessed 12 September 2022
2. Gregg EW, Boyle JP, Thompson TJ, Barker LE, Albright AL, Williamson DF. Modeling the impact of prevention policies on future diabetes prevalence in the United States: 2010–2030. Popul Health Metr 2013;11:18
3. Harding JL, Pavkov ME, Magliano DJ, Shaw JE, Gregg EW. Global trends in diabetes complications: a review of current evidence. Diabetologia 2019;62:3–16
4. American Diabetes Association. *Standards of Medical Care in Diabetes—2021* abridged for primary care providers. Clin Diabetes 2021;39:14–43
5. Davis J, Fischl AH, Beck J, et al. 2022 National standards for diabetes self-management education and support. Sci Diabetes Self Manag Care 2022;48:44–59
6. Marciano L, Camerini A-L, Schulz PJ. The role of health literacy in diabetes knowledge, self-care, and glycemic control: a meta-analysis. J Gen Intern Med 2019;34:1007–1017
7. Brown SA, Becker BJ, García AA, Brown A, Ramírez G. Model-driven meta-analyses for informing health care: a diabetes meta-analysis as an exemplar. West J Nurs Res 2015;37:517–535
8. Villagomez ET. *Health Beliefs, Knowledge, and Metabolic Control in Diabetic Mexican American Adults*. Houston, TX, University of Texas Health Science Center at Houston, 1989
9. Fitzgerald JT, Funnell MM, Anderson RM, Nwankwo R, Stansfield RB, Piatt GA. Validation of the Revised Brief Diabetes Knowledge Test (DKT2). Diabetes Educ 2016;42:178–187

10. Dunn SM, Bryson JM, Hoskins PL, Alford JB, Handelsman DJ, Turtle JR. Development of the diabetes knowledge (DKN) scales: forms DKNA, DKNB, and DKNC. Diabetes Care 1984;7:36–41
11. Garcia AA, Villagomez ET, Brown SA, Kouzekanani K, Hanis CL. The Starr County Diabetes Education Study: development of the Spanish-language diabetes knowledge questionnaire. Diabetes Care 2001;24:16–21
12. McMillin JM. Chapter 141: Blood glucose. In *Clinical Methods: The History, Physical, and Laboratory Examinations*. 3rd ed. Walker HK, Hall WD, Hurst JW, Eds. Boston, MA, Butterworths, 1990. Available from https://www.ncbi.nlm.nih.gov/books/NBK248/#top. Accessed 11 May 2023
13. Dickinson JK, Guzman SJ, Maryniuk MD, et al. The use of language in diabetes care and education. Diabetes Care 2017;40:1790–1799
14. DeVellis RF, Thorpe CT. *Scale Development: Theory and Applications*. Newbury Park, CA, Sage, 2021
15. Baur C, Prue C. The CDC Clear Communication Index is a new evidence-based tool to prepare and review health information. Health Promot Pract 2014;15:629–637
16. Lynn MR, Layman EL, Englebardt SP. Nursing administration research priorities: a national Delphi study. J Nurs Adm 1998;28:7–11
17. Centers for Disease Control and Prevention. Cognitive interviewing. Available from https://www.cdc.gov/nchs/ccqder/evaluation/CognitiveInterviewing.htm. Accessed 5 October 2022
18. Jones PS, Lee JW, Phillips LR, Zhang XE, Jaceldo KB. An adaptation of Brislin's translation model for cross-cultural research. Nurs Res 2001;50:300–304
19. Federal Interagency Forum on Aging-Related Statistics. *Older Americans 2020: Key Indicators of Well-Being*. Washington, D.C., U.S. Government Printing Office, 2020
20. Huizinga MM, Elasy TA, Wallston KA, et al. Development and validation of the Diabetes Numeracy Test (DNT). BMC Health Serv Res 2008;8:96
21. Saydah SH. Medication use and self-care practices in persons with diabetes. In *Diabetes in America*. 3rd ed. Cowie CC, Casagrande SS, Manke A, et al., Eds. Bethesda, MD, National Institute of Diabetes and Digestive and Kidney Diseases, 2018. Available from https://www.ncbi.nlm.nih.gov/books/NBK567996. Accessed 11 May 2023
22. Kuder GF, Richardson MW. The theory of the estimation of test reliability. Psychometrika 1937;2:151–160
23. Kurpius SE, Stafford ME. *Testing and Measurement: A User-Friendly Guide*. Thousand Oaks, CA, Sage, 2006
24. Ebel RL. *Essentials of Educational Measurement*. 3rd ed. Englewood Cliffs, NJ, Prentice Hall, 1979
25. Salkind NJ, Ed. *Encyclopedia of Research Design*. Vol. 1. Los Angeles, CA, Sage, 2010
26. Attali Y, Fraenkel T. The point-biserial as a discrimination index for distractors in multiple-choice items: deficiencies in usage and an alternative. J Educ Meas 2000;37:77–86
27. Hingorjo MR, Jaleel F. Analysis of one-best MCQs: the difficulty index, discrimination index and distractor efficiency. J Pak Med Assoc 2012;62:142–147
28. Flesch R. A new readability yardstick. J Appl Psychol 1948;32:221–233
29. Fernández Huerta J. Simple readability measurements. Consigna 1959;214:29–32 [In Spanish]
30. U.S. Census Bureau. Detailed languages spoken at home and ability to speak English for the population 5 years and over: 2009–2013. Available from https://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html. Accessed 12 September 2022
31. Fowler FJ Jr. *Improving Survey Questions: Design and Evaluation*. Thousand Oaks, CA, Sage, 1995
32. Cronbach LJ. Studies of acquiescence as a factor in the true-false test. J Educ Psychol 1942;33:401–415