

# Measuring Dementia Knowledge in German: Validation and Comparison of the Dementia Knowledge Assessment Scale, the Knowledge in Dementia Scale, and the Dementia Knowledge Assessment Tool 2

Florian Melchior and Birgit Teichmann\*

*Network Aging Research, Heidelberg University, Heidelberg, Germany*

Accepted 1 May 2023

Pre-press 7 June 2023

## Abstract.

**Background:** Assessing dementia knowledge is critical for developing and improving effective interventions. There are many different tools to assess dementia knowledge, but only one has been validated in German so far.

**Objective:** To validate two tools for assessing dementia knowledge – the Dementia Knowledge Assessment Scale (DKAS-D) and the Knowledge in Dementia Scale (KIDE-D) for the German general population – and compare their psychometric properties with the Dementia Knowledge Assessment Tool 2 (DKAT2-D).

**Methods:** A convenience sample of 272 participants completed online surveys. Analyses included internal consistency, structural validity, construct validity through the known-groups method, retest-reliability with a subgroup of  $n = 88$ , and floor and ceiling effects. This study used the STROBE checklist.

**Results:** Internal consistency was acceptable for DKAT2-D ( $\alpha = 0.780$ ), very good for DKAS-D ( $\alpha = 0.873$ ), and poor for KIDE-D ( $\alpha = 0.506$ ). Construct validity was confirmed for all questionnaires. Retest-reliability was good for DKAT2-D (0.886; 0.825–0.926) and KIDE-D (0.813; 0.714–0.878), while it was great for DKAS-D (0.928; 0.891–0.953). Trends toward ceiling effects were observed for DKAT2-D and KIDE-D but not for DKAS-D. The principal component analysis did not reveal a coherent structure for DKAT2-D or KIDE-D, while the confirmatory factor analysis proposed the removal of 5 items for DKAS-D, resulting in the shortened DKAS20-D, which had nearly identical properties.

**Conclusion:** Both DKAS-D and its shortened version, DKAS20-D, are reliable instruments for evaluating programs intended for the general population, as they were found to be convincing in all aspects.

Keywords: Alzheimer's disease, confirmatory factor analysis, dementia, knowledge, scale psychometrics, reliability, validity

## INTRODUCTION

In Germany, around 330,000 new cases of dementia are counted each year, making it one of the

most common neurodegenerative diseases [1]. However, this number has continued to rise over the past decades and, according to forecasts by the German Alzheimer Society, will grow by another 45,000 new cases each year [1, 2]. The reason for this is the ever-increasing life expectancy, as the prevalence of dementia increases with age. For those under the age of 65, the prevalence is only about 2%, while it rises to

---

\*Correspondence to: Birgit Teichmann, Network Aging Research, Heidelberg University, Bergheimer Strasse 20, 69115 Heidelberg, Germany. Tel.: +49 6221 548124; E-mail: teichmann@nar.uni-heidelberg.de.

over 35% for those over 90, with women being disproportionately more affected [1, 3]. Though the number of dementia patients will increase from 1.53 million in 2022 to around 2.35 million by 2050, accompanied by an increased need for caregivers.

In Germany, about 74% of all people in need of care are cared for at home by their relatives. About two-thirds of informal caregivers provide care without the professional support of outpatient care services [4]. It is mainly women who take on caring responsibilities for their relatives, mostly children and spouses of the person with dementia (PwD) [5, 6], whereby the share of men has grown over the past years [7].

The increase in informal caregiving has driven the interest in psychological consequences for those who provide it, resulting in a better understanding of the factors which need to be considered. The stress experienced by informal caregivers when providing care has been shown to pose a significant risk to their mental health, potentially leading to mental disorders such as depression and anxiety [8–11]. To counteract this, psychosocial interventions for carers are offered to improve the efficiency of the care activity. Over the last decade, two initiatives were established to build a network of institutions providing educational support for people interested in acquiring knowledge about dementia. These initiatives are called “Dementia Friends” in the UK and “Dementia Partners” in Germany. The training material for these interventions typically includes information on understanding dementia and its symptoms, communication strategies for interacting with PwD, and practical tips for managing challenging behaviors. The goal of the training is to equip individuals with the knowledge and skills they need to provide effective care while also taking care of their own mental health. These courses are open to all individuals who want to support others in their community, not just caregivers of people with dementia [12], and there are courses that are specifically tailored to certain occupational groups, such as police officers, retail employees, firefighters, emergency service personnel, and others [13].

With millions of participants and over 98,000 trained individuals in Germany alone [12, 14, 15], the program has expanded internationally in recent years, funded by their respective governments [16, 17].

These interventions often include knowledge acquisition as one of the central elements. In the last few years, a number of studies have been published, which have confirmed the importance of

caregivers’ knowledge about dementia [18, 19], especially when combined with individualized training [20]. Nevertheless, these publications also illustrated the current methodological constraints of inaccurate measurements or no measurements at all for dementia knowledge and the problems with self-developed, non-validated instruments [21, 22], even though there are numerous questionnaires available.

Sullivan et al. (2017) and Spector et al. (2012) compared various dementia knowledge questionnaires [23, 24] and found major differences in the quality of their psychometric properties. The researchers also discovered weaknesses that should be considered in future projects. According to their reviews, several questionnaires are being used outside their target sample without re-establishing their psychometric properties. These are mostly questionnaires for nurses and health professionals that are applied to survey informal caregivers or the general public. Some of the questionnaires also contain items that no longer reflect the current state of research.

The validation and improvement of programs such as the Dementia Partners Program require precise instruments; however, to date, only the Dementia Knowledge Assessment Tool 2 (DKAT2) has been translated into German and validated for the general public [25]. Although the Dementia Knowledge Assessment Scale (DKAS) by Annear et al. (2017) [26] and the Knowledge in Dementia Scale (KIDE) by Elvish et al. (2014) [27] are already available in German, and KIDE has been applied in research [28, 29], both have not been validated yet.

In light of these findings and to identify the scale best suited for evaluating psychosocial interventions, the study aims to validate the German versions of the KIDE and DKAS and compare their psychometric properties with those of DKAT2-D, which we have already validated in a previous study [25].

## MATERIALS AND METHODS

### *Design*

The study design corresponds to an exploratory cross-sectional study, evaluating and comparing the psychometric properties of DKAS-D, KIDE-D, and DKAT2-D. The present study adheres to the EQUATOR guidelines for reporting research using the “Strengthening the Reporting of Observational Studies in Epidemiology” (STROBE) checklist.

## Participants

The data used is based on a convenience sample recruited online from September 2022 to December 2022. Subjects were recruited through newsletters, flyers, publicity, and through social media channels such as WhatsApp, Instagram, and Facebook. There were no inclusion criteria other than the minimum age of 18 years. The sample size was calculated according to Polit's (2012) recommendation of ten subjects per item used [30]. DKAS is the longest questionnaire with 25 items, whereby a sample size of  $N = 250$  is targeted.

The final sample size was  $N = 272$ , and some of the participants completed the questionnaire again after four weeks, resulting in a subgroup of  $n = 88$ .

## Data collection and procedure

The data was collected via a Google Forms questionnaire, which required the participants to consecutively fill out all three scales. Additional questions included sociodemographic data and knowing someone with dementia, caring for a PwD or working with them, and if he or she has participated in any form of dementia training course. At the end of the questionnaire, the participants were asked if they wanted to participate a second time in four weeks, leaving their email addresses to be contacted. Those who agreed to participate a second time were required to create a unique identifier consisting of the letters of their parents' names and the digits of their birthdays, which were used to match the records. The time needed to complete the questionnaire was estimated to be around 10–15 minutes.

There was no missing data, as Google Forms only accepted completed sets of data.

## Instruments

### DKAT2-D

The DKAT2 questionnaire was developed in 2014 by Toye et al. [31], the German version was translated and validated by Teichmann et al. (2022) [25]. It measures how much general knowledge a person has about the dementia disease. It has also been translated and validated in Greek, Spanish, and Portuguese and showed an internal consistency of 0.68–0.83 [25, 32–34] with a retest-reliability of 0.918, 95% CI [0.879, 0.945] [25]. The questionnaire consists of 21 items, each of which proposes a true or false statement about the dementia disease, e.g., "Dementia

occurs because of changes in the brain" and must be answered by either "yes", "no", or "I don't know". Correct answers are rewarded with one point, and incorrect or "I don't know" answers are counted as zero points. The score is calculated by adding up the points from each individual item, so that a score from 0 to 21 can be achieved, with a higher score associated with a better knowledge of the dementia disease. Items 5, 6, 7, 8, 12, 16, 18, and 20 are reverse-coded because they are false statements.

### DKAS-D

The 25-item version of DKAS by Annear et al. (2017) also measures a person's knowledge about dementia and achieved an internal consistency of 0.85 [26] in their validation. The scale is characterized by four subscales: "Causes and Characteristics", "Communication and Behaviours", "Care Considerations", and "Risks and Health Promotion" and contains statements that the participant must answer on a 4-point Likert scale, which gives the following possible answers: "true", "probably true", "probably false", and "false". In case of uncertainty, the participant can also state, "I don't know". If the statement is true, and "true" was given as an answer, two points are awarded, for "probably true" the participant receives one point, and for all other options zero points are awarded for this item. This applies vice versa if the item is a false statement. Therefore, the range for the final score is 0–50, with a higher score indicating a better knowledge about dementia. DKAS is considered by Annear et al. as a direct successor and improved version of DKAT2 [35]. In addition, it has been validated in Japanese [36], Spanish [37], Chinese [38], traditional Chinese [39], and Turkish [40].

The back-translation method was used for the translation process [41]. The English original was first translated into German and then back-translated independently by two native English speakers. The back translations were compared with the original and, in the case of linguistic differences, adjusted, retranslated, and back-translated again. The final German version was then tested for comprehensibility by external nurses of different specialties and professional years.

### KIDE-D

The "Knowledge in Dementia Scale" (KIDE), developed by Elvish et al. in 2014, is a 16-item questionnaire with true and false statements about dementia that participants must answer with either

“agree” or “disagree”. Correct answers are rewarded with one point and false answers with zero points. Therefore, the range for the total score is between 0–16, with a higher score being associated with better knowledge about dementia. Elvish et al. reported a Cronbach’s alpha of 0.66 to 0.72 [27, 42] with no coherent factor structure [42]. KIDE was also translated using the back-translation method [43–45] but is not validated yet either. To our knowledge, KIDE has not yet been translated into any other language.

### *Statistical analysis*

The program IBM SPSS Statistics Version 27 was applied for the descriptive and inferential statistical analysis of the data [46]. For all three questionnaires, the following psychometric properties were examined: internal consistency with Cronbach’s alpha, structural validity with principal component analysis or confirmatory factor analysis, construct validity through the method of known groups, an item analysis, and a test for floor and ceiling effects.

The confirmatory factor analysis (CFA) was performed using the AMOS 27.0 package for SPSS [47].

### *Internal consistency and reliability*

For internal consistency, Cronbach’s alpha was calculated, which expresses the degree of shared variance among items and is helpful in assessing the reliability of single-construct scales with more than 10 items [48–50]. The recommended range for Cronbach’s alpha is 0.70–0.90 [48, 50].

For evaluating the retest-reliability, we compared the data of the total sample ( $N=272$ ) with the subgroup that participated ( $n=88$ ) after a four-week interval. Furthermore, we calculated the interclass correlation coefficient, which characterizes the similarities in the results of the two surveys [51, 52]. The retest reliability was calculated according to Koo and Li (2016) in SPSS using a two-way mixed effects model with the mean of  $k$  measurements and absolute agreement.

### *Structural validity*

A principal component analysis (PCA) with varimax rotation was performed for DKAT2-D to check whether the questionnaire exhibited a factor structure. In the next step, we compared the PCA to our previous validation [25], where we could not identify a comprehensible factor structure.

We examined if DKAS-D retains its structural identity originally proposed by its authors [26] by

performing a CFA with the maximum likelihood estimation procedure. The results of the CFA were reported according to the recommendations of Jackson et al. (2009) and Schreiber et al. (2006) [53, 54]. We additionally examined whether KIDE-D also had an ambiguous factor structure for our sample, as reported by the original authors [42]. The criteria that had to be met for the PCA were the Kaiser-Meyer-Olkin coefficient of  $>0.6$  [55, 56] and a significant Bartlett’s test of sphericity [57].

### *Construct validity*

To investigate construct validity, we used the known-groups method with persons with and without experience with dealing with dementia as well as with those who had either participated in a dementia training course or not. We scored “Experience with dementia” if the person either knew someone with dementia, cared for a person with dementia, or worked with them. The information about attending a dementia course was explicitly asked about in the questionnaire and prioritized if more than one category applied. It was expected that knowledge would be higher among those who had experience with a person with dementia than among those who had not, and that participation in dementia training should also result in significantly higher knowledge about dementia. For that reason, we tested the following hypothesis for each of the three scales: 1) Individuals with prior experience with PwD will score higher on the scales than those without such experience. 2) Individuals who have participated in dementia training will score higher on the scales than those who have not received any training.

To test these hypotheses, we used the Wilcoxon-Mann-Whitney (resp. Mann-Whitney-U) (WMW) test [58]. Because the interpretation of the WMW test depends on whether the two groups tested have the same or different distributions [59, 60], a two-sample Kolmogorov-Smirnov test was performed to verify the variance homogeneity of the groups [61]. In the case of equal distribution, the WMW test makes a statement about the medians of the two groups; in the case of unequal distributions, the WMW test considers the distribution form as a whole and considers the mean ranks of the two groups.

According to Fritz et al. (2012), effect sizes for the WMW tests are calculated with the Pearson-Correlation coefficient [62].

### Power analysis

To ensure that the WMW tests for experience with dementia and participation in a dementia training are not underpowered, we performed a power analysis with G\*Power (version 3.1.9.7) [63, 64] according to Kang (2021) [65]. The following hypotheses were formulated: 1) the group with dementia experience will have a significantly higher median or mean rank in total score than the group without experience, and 2) the group who has participated in a dementia training will have a significantly higher median or mean rank in total score than the group that did not participate. These hypotheses were investigated for all three scales. The corresponding null hypotheses were: 1) there will be no significant difference in median or mean rank for the total score between the two groups, or the group with dementia experience has a lower median or mean rank in total score than the group with no experience, and 2) there will be no significant difference in median or mean rank for the total score between the two groups, or the group that participated in a dementia training will have a lower median or mean rank for the total score than the group that did not participate.

The expected effect size is at least  $d=0.5$  with a power of  $1 - \beta = 0.95$  and significance level  $\alpha = 0.05$ . The allocation ratio was set at 3 because it was found in our earlier study and other studies that included the general population that there tended to be an unequal number of people who had attended dementia training or had experience with PwD, which resulted in a higher number of participants needed [25, 26, 40]. The resulting recommended sample size was  $N = 244$  for both groups combined, which is lower than the applied rule of thumb previously stated.

### Item analysis

For all three questionnaires, an item analysis was conducted in which we examined the difficulty of all items. At the same time, for DKAT2-D and DKAS-D, we also examined the percentage of “don’t know” responses in the entire questionnaire, which we termed “item ignorance”. We conducted the difficulty and ignorance analysis for the total sample as well as separately for those who did and did not participate in dementia training. In addition, we examined the item-total correlation of all items, which indicates the extent to which an individual item’s score is consistent with the overall scale score, which is a valuable statistic for identifying items with weak explanatory power. Furthermore, we investigated inter-item correlation, which indicates how strongly items correlate

with each other. A commonly used criterion for item-total and mean inter-item correlations is a correlation of around 0.2 to 0.4 for the items to contribute significant information to the scale, although a higher correlation does not equate to higher reliability; on the contrary, too high a correlation means that an item becomes redundant and artificially increases reliability [66–68].

### Floor and ceiling effects

Another important criterion that should be considered is whether ceiling or floor effects occur. The floor or ceiling is a value that an observation cannot exceed, such as a perfect score. A ceiling or floor effect is when a variable accumulates at this value and skews its distribution, introducing a bias and producing misleading results in analyses that assume a normal distribution [69, 70]. Thresholds for this effect are not strictly defined, and we considered such an effect to be present if more than 10% of all participants scored a minimum or maximum score on a questionnaire. In addition, we re-examined ceiling effects for the sample that participated in dementia training to determine whether a ceiling effect was sample dependent.

### Ethics

The study was performed according to the ethical standards outlined in the Declaration of Helsinki. Participants participated voluntarily in the study after being informed about the aim of the study and subsequently provided their written consent for participation. The General Data Protection Regulation (GDPR) in a research context [71] was respected through the confidentiality and anonymity of the data. This study was approved by the Ethics Committee of the Faculty for Behavioural and Empirical Cultural Sciences Heidelberg University, Germany (AZ Teich 2022 2/1).

## RESULTS

The final sample consisted of 272 subjects, with an additional subgroup of 88 participants. The socio-demographic data are depicted in Table 1. The average age of the participants was 45.03 years, with most of them being female (71.3%). Most participants hold at least a diploma or master’s degree (32.7%). In terms of experience with dementia, 75% of participants know a PwD, 11% care for a PwD, and 20.6% work with PwD. Among all participants, 28.7% have participated in dementia training.

Table 1  
Participants' characteristics of the total sample and the Subgroup

Characteristics	Full sample (N = 272)		Subgroup <sup>1</sup> (n = 88)	
	n	%	n	%
Age				
mean	45.03		47.39	
SD	19.05		18.84	
Gender				
male	77	28.3%	26	29.5%
female	194	71.3%	62	70.5%
diverse	1	0.4%	0	
Education				
9 years or less	7	2.6%	0	
10 years	16	5.9%	5	5.7%
12-13 years	64	23.5%	20	22.7%
Vocational training	43	15.8%	16	18.2%
Bachelor	31	11.4%	13	14.8%
Master/Diploma	89	32.7%	28	31.8%
PhD	20	7.4%	5	5.7%
Others	2	0.7%	1	1.1%
Occupation				
School student	4	1.4%	0	
Student	58	21.3%	17	19.3%
Unemployed	2	0.7%	0	
Retiree	44	16.2%	17	19.3%
Care profession	28	10.3%	13	14.8%
Therapeutical profession	17	6.3%	4	4.5%
Physician	2	0.7%	1	1.1%
Others	117	43.1%	36	40.1%
Experience with PwD				
I know one or more persons with dementia	204	75.0%	70	79.5%
I care for a person with dementia	30	11.0%	13	14.8%
I work with PwD	56	20.6%	21	23.9%
Participation in a program about dementia				
Yes, I participated	78	28.7%	35	39.8%
I did not participate	194	71.3%	53	60.2%

<sup>1</sup>Subgroup after four weeks for the retest.

### DKAT2-D

The DKAT2-D had a mean value of 14.17 (SD = 3.69) on a range from 0 to 21 points.

#### Internal consistency

Psychometric properties from all the questionnaires are listed in Table 2. DKAT2-D achieved a Cronbach's alpha of  $\alpha = 0.78$  which is considered to be a sufficient value [50], although it is worth mentioning that the DKAT2-D had an unsatisfactory Cronbach's alpha ( $\alpha = 0.536$ ) for those who attended a dementia training. This is presumably because items 1, 13, and 18 were answered 100% correctly, so there was no variance for these items.

#### Structural validity

The Kaiser-Meyer-Olkin (KMO) value for the 21-item scale is 0.739 with a significant Bartlett's test for sphericity  $\chi^2(210) = 965.07, p < 0.001$ , which makes

this dataset suitable for the PCA. However, the dimensionality reduction of the PCA failed to identify any coherent factors.

#### Construct validity

All results for the known-groups method are shown in Table 3, including the test for variance homogeneity and effect size. The Wilcoxon-Mann-Whitney tests were able to show that DKAT2-D was able to discriminate between (1) the groups with or without previous experience with PwD ( $U = 2802.0, z = -5.861, p < 0.001, |r| = 0.36$ ) and (2) the groups that had participated or had not participated in dementia training ( $U = 2626.5, z = -8.447, p < 0.001, |r| = 0.51$ ).

#### Item analysis

In the total sample, 67.5% of the 21 items were answered correctly on average by one participant,

Table 2  
Psychometric properties of the DKAT2-D, DKAS-D, shortened DKAS20-D and KIDE-D

	Cronbach's alpha			Mean item-total correlation	Mean inter-item correlation	Mean item difficulty	Mean item ignorance
	Total sample	NoDT <sup>1</sup>	DT <sup>2</sup>	Total sample	Total sample	Total sample	Total sample
DKAT2-D	0.780	0.743	0.536	0.331	0.141	67.50%	20.10%
DKAS-D	0.873	0.838	0.809	0.433	0.215	70.71%	14.30%
DKAS20-D	0.872	0.831	0.821	0.472	0.259	69.13%	15.75%
KIDE-D	0.506	0.458	0.461	0.159	0.051	77.70%	n. a.

<sup>1</sup>NoDT, has not participated in dementia training; <sup>2</sup>DT, has participated in dementia training.

Table 3  
Results of the Wilcoxon-Mann-Whitney tests for all questionnaires

Group variable	Mean rank		Median		U <sup>1</sup>	z <sup>2</sup>	p <sup>3</sup>	r <sup>4</sup>	K-S test <sup>5</sup>
	Yes (n = 219)	No (n = 53)	Yes	No					
Experience with dementia									
DKAT2-D	<b>150.21</b>	<b>79.87</b>	15	11	2802.0	-5.861	<0.001	-0.36	<b>&lt;0.001</b>
DKAS-D	<b>148.92</b>	<b>85.18</b>	31	23	3083.5	-5.296	<0.001	-0.32	<b>0.026</b>
DKAS20-D	<b>148.68</b>	<b>86.19</b>	24	17	3137.0	-5.193	<0.001	-0.31	<b>0.031</b>
KIDE-D	141.88	114.25	<b>13</b>	<b>12</b>	4624.5	-2.319	0.02	-0.14	0.367
Participated in dementia training	Yes (n = 78)	No (n = 194)	Yes	No	U	z	p	r	K-S test
DKAT2-D	199.83	111.04	<b>17</b>	<b>13</b>	2626.5	-8.447	<0.001	-0.51	0.415
DKAS-D	197.97	111.78	<b>36</b>	<b>25.5</b>	2771.0	-8.177	<0.001	-0.50	0.68
DKAS20-D	198.86	111.43	<b>30</b>	<b>19</b>	2702.0	-8.296	<0.001	-0.50	0.71
KIDE-D	174.61	121.18	<b>13</b>	<b>12</b>	4593.5	-5.121	<0.001	-0.31	0.178

<sup>1</sup>U, U test statistic; <sup>2</sup>z, z statistic; <sup>3</sup>p, significance; <sup>4</sup>r, Pearson correlation coefficient for effect size; <sup>5</sup>K-S test, two-sample Kolmogorov-Smirnov test for variance homogeneity, significant K-S tests are bold highlighted, numbers highlighted in bold are the descriptive statistics to be interpreted according to the K-S test.

and typically, “I don’t know” was given as an option for 20.1% of the items. Comparing dementia training participants with non-participants showed that those who participated in a training answered on average 81.1% of questions correctly and chose the “I don’t know” option on average for 7.4% of the items. Those who did not participate answered on average 62% of the questions and chose the “I don’t know” option for 24.9% of the items.

The item-total correlations were mostly good, with a range of 0.111 to 0.532 and a mean item-total correlation of 0.331. Item 7 had a correlation with the total scale that was slightly low, at 0.111. The mean inter-item correlation was 0.14, ranging from 0.046 to 0.225, with no redundant items. Some items had a mean below 0.1, with item 7 having the lowest value at 0.046, indicating that it barely contributed any relevant information to the scale.

*Retest reliability*

The retest reliability was calculated with the subgroup (n = 88) after a four-week interval. Table 4 contains all intraclass correlation coefficients with their respective confidence intervals for all questionnaires. DKAT2-D achieved an ICC with the lower

Table 4  
Test-retest reliability with the subgroup (n = 88)

	Intraclass Correlation Coefficient	95% - Confidence interval	
		Lower bound	Upper bound
DKAT2-D	0.886	0.825	0.926
DKAS-D	0.928	0.891	0.953
DKAS20-D	0.931	0.895	0.955
KIDE-D	0.813	0.714	0.878

bound at 0.825, which is considered a good value [52].

*Floor and ceiling effects*

In the total sample, only one person managed to achieve the maximum score of 21, and there were no participants who achieved a score lower than 3. Therefore, no ceiling or floor effect was observed. In the dementia training group, no person achieved a score of 21, but 14.1% of participants reached a score of 20. The psychometric properties of item 7 clearly indicate that it should be removed from the scale, and since it is by far the most challenging item on the scale, this would result in 5.5% of the total sample achieving a maximum score (20 in this case)

and 12.8% of the dementia training group achieving a maximum score, resulting in a significant ceiling effect.

#### DKAS-D

DKAS-D had a mean score of 28.80 (SD = 9.46) on a scale from 0 to 50 points.

#### Internal consistency

DKAS-D achieved a Cronbach's alpha of  $\alpha = 0.873$ , which corresponds to a high common variance of the scale items [50].

#### Construct validity

DKAS-D was also able to distinguish 1) between individuals with and without prior experience with PwD:  $U = 3083.5$ ,  $z = -5.296$ ,  $p < 0.001$   $|r| = 0.32$  and 2) between participants and non-participants in dementia training:  $U = 2771.0$ ,  $z = -8.177$ ,  $p < 0.001$   $|r| = 0.50$ .

#### Item analysis

For the calculation of the difficulty level, a question was considered to be answered correctly if 1 or 2 points were scored. In the total sample, the questionnaire was answered correctly on average 70.71% of the time, with 14.30% of questions answered "I don't know".

The group that had not participated in any training was able to answer an average of 66.0% correctly, with "I don't know" being given for 17.9%. On the other hand, the group that attended a training course answered the questionnaire correctly with an average of 82.5%, with only 5.4% of the questions being answered with "I don't know". When examining individual items in the total sample, Item 6 proved too easy with 96% correct answers. However, there were no other outliers.

The mean item-total correlation was 0.433, with a range of 0.214 to 0.578, indicating a good reliability of the questionnaire and no item redundancy or any item with too little explanative power. The mean inter-item correlation was 0.215, ranging from 0.110 to 0.286, with items 4, 6, 18, 19, and 25 having the lowest means with 0.110–0.182.

#### Retest reliability

Retest reliability for DKAS-D was at 0.928, with its lower bound at 0.891 (see Table 4), which could be considered a very good reliability.

#### Floor and ceiling effects

The lowest score achieved was 8, and there were two individuals who achieved the maximum score of 50, so by our definition, there was neither a ceiling nor a floor effect present. Only one person in the dementia training group scored the maximum score, and there was no aggregation of scores near the maximum.

#### Structural validity

The first CFA model included all 25 items and provided insufficient model fit. The  $\chi^2$  test was significant, which is expected with a large sample size,  $\chi^2(269, N = 272) = 463.23$ ,  $p < 0.001$ , and should not be considered for model fitting [72, 73]. The model achieved a root mean square error of approximation (RMSEA) of 0.052, 90% CI [0.044; 0.059] a comparative fit index (CFI) of 0.875, the Tucker-Lewis's coefficient (TLI) of 0.861, and a standard root mean square residual (SRMR) of 0.058. The CFI and TLI were below the recommended thresholds of  $>0.95$  for model fit criteria [54], the RMSEA suggests a good fit with the upper confidence interval  $<0.06$ , and the SRMR does too, with a value  $<0.08$ . After investigating standardized regression weights, squared multiple correlations, and standardized residual covariances, items 4, 6, 18, 19, and 25 were excluded from the model because they all had very low regression weights  $<0.4$  and very low squared multiple correlations  $<0.2$ . Moreover, items 25, 18, and 19 showed overall high residual covariances  $>2$  with many items. These are all indicators that these items are not suitable for the model and are a source of misfit [74]. These five items also showed the lowest mean inter-item correlations, all below  $<0.2$ , supporting their removal. Additionally, error variances from items 2 and 7 and items 8 and 9 were covaried because they showed high covariance.

#### Psychometric properties of the 20-item DKAS-D

After that, the 20-item version was retested, which is depicted in Fig. 1. Fit indexes improved slightly, although the  $\chi^2$  test was still significant with  $\chi^2(162, N = 272) = 263.5$ ,  $p < 0.001$ . The 20-item version achieved an RMSEA of 0.048, 90% CI [0.037; 0.058]. It also showed a higher CFI of 0.928, an improved TLI of 0.915, and an improved SRMR of 0.052.

The psychometric properties of the 20-item version were retested to compare the modified version with the original. The full results are given in Tables 2–4. Overall, the quality of the questionnaire did not change, and the shortened 20-item version has very



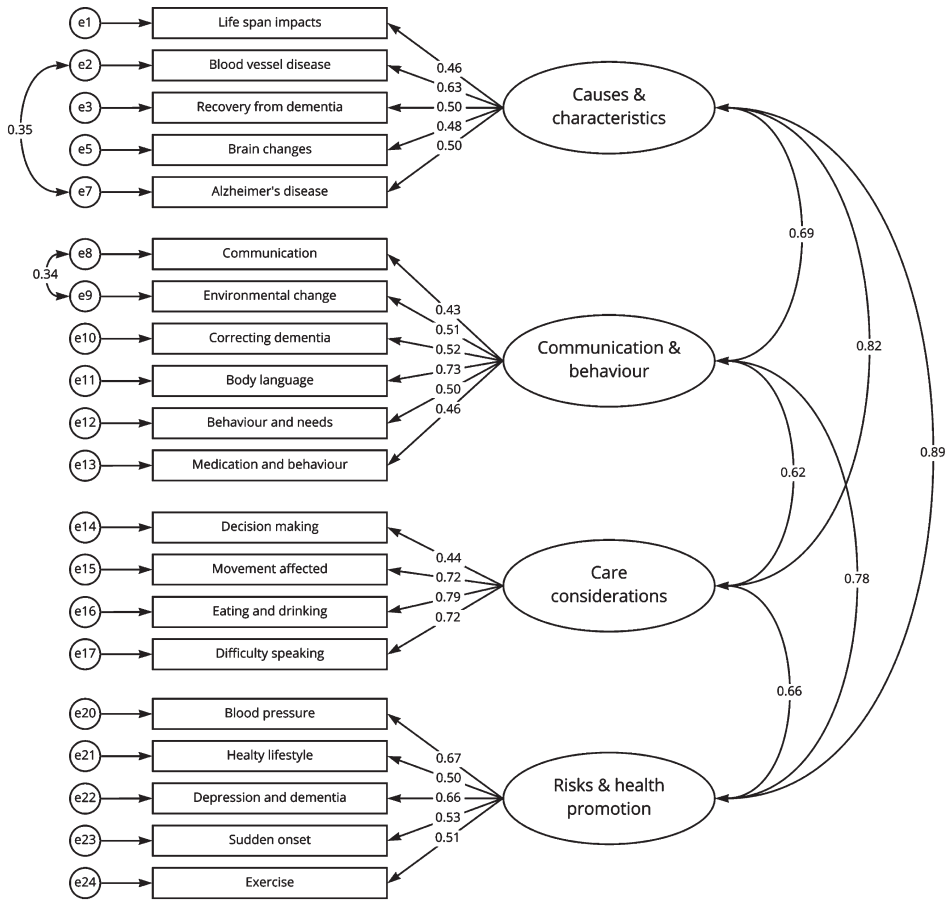


Fig. 1. Standardized estimates of the confirmatory factor analysis of the shortened DKAS-D without Items 4, 6, 18, 19, and 25.

similar properties. For construct validity, the WMW tests for dementia experience and dementia training were significant. The retest yielded nearly identical results to the 25-item version, and the item analyses yielded very similar results to the 25-item version, with no floor or ceiling effects.

**KIDE-D**

KIDE-D had a scale mean of 12.43 (SD = 2.05) on a scale from 0 to 16.

*Internal consistency*

KIDE-D achieved a Cronbach’s alpha of  $\alpha = 0.51$  for the total sample, indicating that the scale items have low shared variances and suggesting low reliability [50].

*Structural validity*

The Kaiser-Meyer-Olkin (KMO) value for the 16-item scale is 0.603, with a significant Bartlett test for

sphericity  $\chi^2(120) = 280.10, p < 0.001$ , making this dataset suitable for the PCA.

When conducting the PCA, item 15 had to be excluded because it did not show any variance as it was answered correctly by all participants except one. The PCA produced a six-factor solution with an explained variance of 54.6%. Two factors were open to interpretation: one consisting of items 1 and 7, which related to aggressive behavior, and the other consisting of items 11 and 6, which addressed physical pain. In addition, no clear interpretations were possible, and item 4 had a very low communality ( $h^2 = 0.20$ ), so it did not load on any factor.

*Construct validity*

KIDE-D was able to distinguish between 1) individuals who had no experience with PwD and those who had, and 2) individuals who had attended dementia training and those who had not. For the groups with no dementia experience and the ones with experience,

the WMW test statistics were:  $U = 4624.5$ ,  $z = -2.319$ ,  $p = 0.02$   $|r| = 0.14$ . The WMW test for those who attended dementia training and those who did not was  $U = 4593.5$ ,  $z = -5.121$ ,  $p < 0.001$   $|r| = 0.31$ .

#### *Item analysis*

The average participant was able to answer 77.7% of the items correctly. Analysis of the difficulty of the items showed that many questions were answered correctly by most participants. The scale included five items that were answered correctly by over 90% of participants: item 15 (99.6% correct answers), item 14 (98.5%), item 13 (97.1%), item 16 (96.3%), and item 2 (93.4%).

When looking at the inter-item correlations, it was noticeable that there were many small negative correlations between the items, even though this was a single scale measuring a single variable. This was not the case with the other scales. There were also many correlations next to zero, suggesting that the items are generally unfit to map the construct for which they are intended. The mean correlation between items was 0.051, well below the suggested value of 0.2 to 0.4, and the mean item-total correlation was 0.159. However, we also detected some negative correlations that affected the reliability of the scale. Therefore, no absolute values were used, as this would distort the reliability of the scale, since the items should be positively correlated due to the measurement of a unidimensional construct.

#### *Retest reliability*

KIDE-D achieved an ICC of 0.813 with the lower bound at 0.714 for the retest reliability, which is still considered a good reliability.

#### *Floor and ceiling effects*

When examining the floor and ceiling effects, it became clear that the overall score showed a strong tendency toward higher scores, while low scores were virtually absent. The visualization of this can be seen in Fig. 2. In the total sample, only 4.4% of participants achieved the maximum score of 16, so there were no floor or ceiling effects according to our threshold. However, when assessing those who attended dementia training, 11.5% reached the maximum score, indicating a significant ceiling effect.

## **DISCUSSION**

The goal of this study was to validate the German versions of DKAS and KIDE for the general popula-

tion and compare them with the recently validated DKAT2-D to identify the best tool for measuring knowledge of the dementia disease.

All three questionnaires demonstrated a good construct validity by distinguishing inexperienced and experienced groups with respect to persons with dementia and by distinguishing those persons who participated in dementia training from those who did not. When looking at the other psychometric properties, it became clear that KIDE-D has significant weaknesses, and DKAT-D is mostly a good questionnaire, while DKAS-D shows the strongest characteristics in all areas.

DKAT2-D had acceptable Cronbach's alpha values for the general population, and DKAS-D achieved a very good value. The Cronbach's alpha values of DKAT2-D are identical to our previous publication [25] and comparable to the results of other groups' translations [32, 33, 75–77], but just like other researchers, we have also noticed that the internal consistency of DKAT2-D declines for samples with higher dementia knowledge [32, 33] as it was much lower for those who had attended a dementia training. DKAS-D's internal consistency was slightly higher compared to other translations [36–40] but comparable to what Annear et al. reported in the original [26, 35, 78], with no differences when training and non-training groups were inspected separately. KIDE-D had a lower consistency than reported by Elvish et al. (2014 & 2016) in the original study [27, 42] and, unfortunately, did not convince with an overall very low internal consistency.

For structural validity, KIDE-D and DKAT2-D did not reveal an underlying factorial structure. However, we were able to confirm the four-factorial structure of the DKAS-D, which was theorized by the authors [26].

Although we found that items 4, 6, 18, 19, and 25 could be removed without affecting the quality of the questionnaire, we also should note that the fit indexes for the CFA were ambiguous. The RMSEA and SRMR were good, but the TLI and CFI were a bit low ( $< 0.95$ ), suggesting that the latent variables are represented by choice of items, but there is still room for improvement. Nevertheless, the reported fit indexes are not poor, and some authors consider a CFI and TLI  $> 0.90$  as good, although these thresholds change over time [79].

When comparing our CFA to other translations, we found that Sung et al. (2021), who conducted the CFA for the traditional Chinese translation of DKAS [39], and Akyol et al. (2021), who performed the CFA

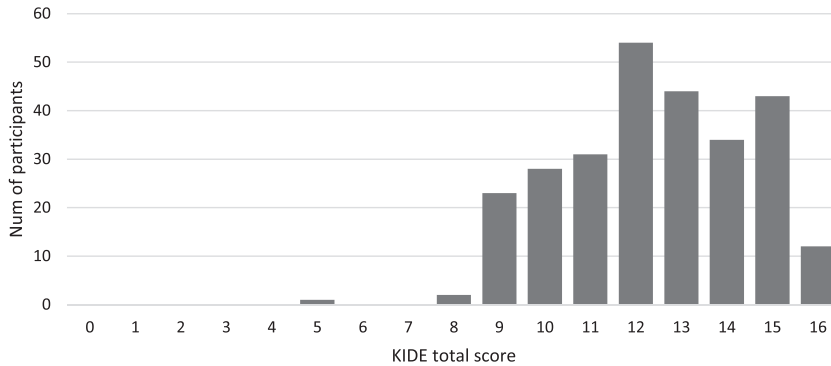


Fig. 2. Frequency of the KIDE-D total scores for the total sample.

for the Turkish version [40], both found a better fit for their models. However, Sung et al. (2021) argued that the CFA is dependent on the sample. Considering that the Chinese participants were already caring for someone with dementia, while the Turkish sample included a high proportion of nurses and nursing students, it is conceivable that their high model fit is due to a sample with higher general knowledge about dementia.

In the item analyses, we found that DKAT2-D contained some items that were poorly correlated with the overall scale but still within the acceptable range. In contrast, the items of DKAS-D achieved very good values overall and had a coherent structure. However, we obtained poor inter-item correlation values for items 4, 6, 18, 19, and 25, which is consistent with our CFA's findings that these should be removed.

Comparing this with the results of other publications, it appears that DKAS contains excessive items and is often shortened after item analyses, with Annear et al. (2017) [36] removing a total of nine items in their Japanese translation due to poor performance in the item and PCA and Akyol et al. (2021) [40] removing eight items in their Turkish translation. However, not the same items had to be removed, which can be considered as an additional indication of the impact of the sample on item performance.

The items of KIDE-D hardly correlated with each other, and some of them could not contribute relevant information to the scale. It is currently not possible to compare the item analyses of DKAT2 and KIDE with other publications due to the lack of published materials on the subject.

By looking at the difficulty and mean scores of the questionnaires, it became clear that KIDE-D was overall too easy, and a high score was achieved too quickly. This is confirmed by Schneider et al. (2020)

and Jack-Waugh et al. (2018), finding a very high mean for their sample, while Lorio et al. (2016) reported a ceiling effect for their sample [28, 80, 81]. One problem with KIDE is the way the scale is constructed, as it lacks "I don't know" responses, which forces participants to guess when they do not know the answer and introduces a systematic error in the consistency of the scale that should be avoided by introducing a neutral response [82].

DKAT2-D had a similar problem: it was slightly too easy for people who already had knowledge about dementia, but it was in a good range overall. Compared to other studies, this does not seem to be an isolated case. In our previous study [25], we found that people with prior knowledge about dementia also answered a few items exclusively correctly. Both Parra-Anguita et al. (2018) [32] and Toye et al. (2013) [31], when validating the original scale, were able to identify some items that were answered more than 90% correctly by individuals with prior knowledge. However, considering the difficulty, DKAT2 performs well in more heterogeneous samples [25, 32, 33].

DKAS-D, on the other hand, convinced in terms of difficulty and showed no weaknesses in this regard. Moreover, DKAS-D's total score and the average percentage of correct answers in our sample is comparable to the findings of other translations [36–40], being slightly higher than what Sung et al. (2021) found in Taiwan but lower than what the sample of Annear et al. (2017) was able to achieve.

In terms of stability over time, all scales were in good to very good range, with DKAS-D and DKAT2-D achieving very similar and slightly higher scores on the test-retest than KIDE-D. Compared to the results of the first validation study in German, DKAT2-D achieved a slightly lower but very similar retest

reliability [25]. Our result regarding DKAS-D is consistent with other studies that have shown very good stability [37, 38, 40, 78], while no retest reliabilities have been reported for KIDE-D to date.

We did not observe any floor or ceiling effects in DKAT2-D for the total sample, but there was an aggregation of scores near the maximum for individuals who attended dementia training and were more educated about dementia. Out of all the respondents in this group, item 7 was only answered correctly by 14.1%. Given that it had an inter-item correlation close to zero and the lowest item-total correlation, it is an item that should be removed from the scale, which would result in DKAT2-D having a significant ceiling effect because so many participants were missing this one point for a maximum score. We previously acknowledged [25] that item 7 has conspicuous values, and there is a divergence of opinions about whether it actually contains a false statement [25]. However, we will refrain from further discussion on this matter at this point. For DKAS-D, no ceiling or floor effects could be observed, which is consistent with other studies. [36, 40, 83].

Similar to DKAT2-D, KIDE-D did not meet our definition of ceiling effects for the total sample, but examining the distribution diagram clearly underlines the low difficulty of the scale, as most participants were able to achieve a very high score, which hinders the scales' ability to detect small changes and makes it susceptible to ceiling effects. This was evident in the group that had attended dementia training, as they achieved the maximum score of 11.5% of the time, implying that a better-trained group experiences a significant ceiling effect. In a doctoral thesis on dementia knowledge by Gamble (2022) [84], several distribution diagrams of KIDE scores for different samples revealed that the scale experiences a strong ceiling effect in samples with higher pre-existing dementia knowledge. While MacRae et al. (2022) found significant ceiling effects for KIDE in health professionals [83], and given that DKAT2-D and DKAS-D showed that dementia training has a strong effect on dementia knowledge ( $|r| > 0.50$ ), KIDE-D would likely be sufficient to measure changes before and after dementia training in people who have not yet participated in any, respectively large effects. When looking at other publications, it becomes evident that this is possible to some extent, as Elvish et al. (2014 & 2016) had previously managed to distinguish pre- and post-intervention groups in their "Getting to Know Me" project [27, 42]. Jack-Waugh et al. (2018) successfully used KIDE to evaluate an intervention [81],

while Lorio et al. (2016) detected no pre-post difference in the KIDE scores even after an extensive 12-hour training, attributing this to a ceiling effect [80] because their participants were already well educated. Schneider et al. (2020) also expected a significant increase in the KIDE scores after an elaborate intervention on dementia knowledge but failed to demonstrate this [28]. Therefore, according to the studies mentioned above, future projects are recommended to evaluate KIDE-D in a sample that has low dementia knowledge to assess its ability to detect large effects.

For purposes other than identifying large effects, either DKAT2-D or DKAS-D should be used. However, we clearly recommend DKAS-D as we consider it superior in all aspects. This is no surprise because the DKAS can be seen as a further development of the DKAT2 [78] and compares favorably with other dementia knowledge questionnaires, as reported by Annear et al. (2016) [35] and MacRae et al. (2022) [83]. Since the shortened form of DKAS-D with 20 items performs equally well in all parameters, is completed more quickly, and fits better with the proposed theoretical framework, we recommend it for further studies. As previously mentioned, it is crucial to regularly adapt questionnaires and evaluate their psychometric properties to maintain their quality and keep them in line with current research standards.

### *Limitations*

Although we have a diverse sample that includes people from all social groups and we encouraged people to let their social environment participate in the study, comparing our sociodemographic data to the German general population, it seems like our sample is likely to be more educated than the general population [85]. This is a known bias, probably due to the need for less educated people to engage in scientific projects. Also, the channels used to recruit participants are aimed at individuals who are more interested in dementia research.

The format of the online survey brings a set of problems with it, which should also be considered. Online surveys tend to attract certain types of respondents, such as those who are technologically adept or have a lot of free time [86, 87]. This can create a selection bias and skew the results. Furthermore, they lack the personal interaction that is possible with other methods of data collection, such as face-to-face interviews. This can limit the ability to probe for more detailed or nuanced responses [88]. And lastly, technical difficul-

ties such as slow loading times or problems with the survey software can create frustration for respondents and potentially impact response rates.

Accordingly, we would recommend larger samples for future projects and propose an expanded recruitment program. In general, convenience samples tend to be statistically unclear due to the data being WEIRD “Western, Educated, Industrialized, Rich, and Democratic“ [89]. Therefore, generalization and comparisons between cultures are limited.

One problem inherent to this type of study arises from the fact that completing three scales in succession can lead to fatigue. The order of the three scales was not randomized because the survey instrument we used does not support this method. The term “dementia training” was not further specified. Moreover, it should be noted that the quality, quantity, and frequency of these training programs can have a significant impact on a person’s knowledge about dementia.

## ACKNOWLEDGMENTS

This study is independent research. The views expressed in this publication are those of the authors. We want to thank Sabine Krause, Silvia Graf, Cornelia Lips, and Rita Müller from the Cantonal Hospital Winterthur for providing the translated DKAS. We also want to thank Michael Löhr of the Diaconia University of Applied Sciences Bielefeld, Germany, for permission to use the German version of the KIDE.

## FUNDING

The authors have no funding to report.

## CONFLICT OF INTEREST

B.T. is an Editorial Board Member of this journal but was not involved in the peer-review process nor had access to any information regarding its peer-review.

F.M. has no conflict of interest to report.

## DATA AVAILABILITY

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

## SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: <https://dx.doi.org/10.3233/JAD-230303>.

## REFERENCES

- [1] German Alzheimer’s Society (2022) Die Häufigkeit von Demenzerkrankungen, [https://www.deutsche-alzheimer.de/fileadmin/Alz/pdf/factsheets/infoblatt1\\_haeufigkeit\\_demenzerkrankungen\\_dalzg.pdf](https://www.deutsche-alzheimer.de/fileadmin/Alz/pdf/factsheets/infoblatt1_haeufigkeit_demenzerkrankungen_dalzg.pdf).
- [2] (2022) Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the Global Burden of Disease Study 2019. *Lancet Public Health* 7, e105-e125.
- [3] Statistisches Bundesamt (2021) Bis 2035 wird die Zahl der Menschen ab 67 Jahre um 22% steigen, [https://www.destatis.de/DE/Presse/Pressemitteilungen/2021/09/PD21\\_459\\_12411.html;jsessionid=BE2CDC1106D9799086B8E6130BB351C8.live?732](https://www.destatis.de/DE/Presse/Pressemitteilungen/2021/09/PD21_459_12411.html;jsessionid=BE2CDC1106D9799086B8E6130BB351C8.live?732).
- [4] Statistisches Bundesamt (2020) Pflegebedürftige nach Versorgungsart, Geschlecht und Pflegegrade, <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Pflege/Tabellen/pflegebeduerftige-pflegestufe.html>, Last updated December 15, 2020.
- [5] Schneider J, Hallam A, Murray J, Foley B, Atkin L, Banerjee S, Islam MK, Mann A (2002) Formal and informal care for people with dementia: Factors associated with service receipt. *Aging Ment Health* 6, 255-265.
- [6] Lamura G, Mnich E, Wojszel B, Nolan M, Krevers B, Mestheneos L, Döhner H (2006) Erfahrungen von pflegenden Angehörigen älterer Menschen in Europa bei der Inanspruchnahme von Unterstützungsleistungen: Ausgewählte Ergebnisse des Projektes EUROFAMCARE. *Z Gerontol Geriatr* 39, 429-442.
- [7] Schneekloth U (2006) Entwicklungstrends und Perspektiven in der häuslichen Pflege. Zentrale Ergebnisse der Studie Möglichkeiten und Grenzen selbständiger Lebensführung (MuG III). *Z Gerontol Geriatr* 39, 405-412.
- [8] Alves LCdS, Monteiro DQ, Bento SR, Hayashi VD, Pellegrini LNdc, Vale FAC (2019) Burnout syndrome in informal caregivers of older adults with dementia: A systematic review. *Dement Neuropsychol* 13, 415-421.
- [9] Cooper C, Balamurali TBS, Livingston G (2007) A systematic review of the prevalence and covariates of anxiety in caregivers of people with dementia. *Int Psychogeriatr* 19, 175-195.
- [10] Gérard P, Zech E (2021) Do informal caregivers experience more burnout? A meta-analytic study. *Psychol Health Med* 26, 145-161.
- [11] Joling KJ, van Marwijk HWJ, Veldhuijzen AE, van der Horst HE, Scheltens P, Smit F, van Hout HPJ (2015) The two-year incidence of depression and anxiety disorders in spousal caregivers of persons with dementia: Who is at the greatest risk? *Am J Geriatr Psychiatry* 23, 293-303.
- [12] Deutsche Alzheimer Gesellschaft e. V. (2022) Initiative Demenz Partner. Demenz Partnerinnen und Partner in Deutschland - Stand: 31. Dezember 2022, <https://www.demenz-partner.de/>.
- [13] Deutsche Alzheimer Gesellschaft e. V. Über unsere Kurse.

- [14] Alzheimer's Society (2017) What is a Dementia Friends Ambassador? <https://www.dementiafriends.org.uk/WEBArticle?page=what-is-a-champion#.Y-U6V3bMJaY>.
- [15] German Federal Ministry for Family Affairs, Federal Ministry of Health (2019) *The Alliance for People with Dementia - Results of the 2014-2018 Common Efforts - Short Report*, Berlin.
- [16] World Health Organization (2021) Global Dementia Observatory. Awareness and friendliness, <https://www.who.int/data/gho/data/themes/global-dementia-observatory-gdo/awareness-and-friendliness>.
- [17] World Health Organization (2018) Towards a dementia plan: A WHO guide, <https://www.who.int/publications/i/item/9789241514132>.
- [18] Rosi A, Govoni S, Del Signore F, Tassorelli C, Cappa S, Allegri N (2023) Italian Dementia-Friendly Hospital Trial (IDENTITÀ): Efficacy of a dementia care intervention for hospital staff. *Aging Ment Health* **27**, 921-929.
- [19] Annear MJ (2020) Knowledge of dementia among the Australian health workforce: A national online survey. *J Appl Gerontol* **39**, 62-73.
- [20] Perry M, Drašković I, Lucassen P, Vernooij-Dassen M, van Achterberg T, Rikkert MO (2011) Effects of educational interventions on primary dementia care: A systematic review. *Int J Geriatr Psychiatry* **26**, 1-11.
- [21] Zhao Y, Liu L, Chan HY-L (2021) Dementia care education interventions on healthcare providers' outcomes in the nursing home setting: A systematic review. *Res Nurs Health* **44**, 891-905.
- [22] Moore KJ, Lee CY, Sampson EL, Candy B (2020) Do interventions that include education on dementia progression improve knowledge, mental health and burden of family carers? A systematic review. *Dementia (London)* **19**, 2555-2581.
- [23] Sullivan KA, Mullan MA (2017) Comparison of the psychometric properties of four dementia knowledge measures: Which test should be used with dementia care staff? *Australas J Ageing* **36**, 38-45.
- [24] Spector A, Orrell M, Schepers A, Shanahan N (2012) A systematic review of 'knowledge of dementia' outcome measures. *Ageing Res Rev* **11**, 67-77.
- [25] Teichmann B, Melchior F, Kruse A (2022) Validation of the adapted German versions of the dementia knowledge assessment tool 2, the dementia attitude scale, and the confidence in dementia scale for the general population. *J Alzheimers Dis* **90**, 97-108.
- [26] Annear MJ, Toye C, Elliott K-EJ, McInerney F, Eccleston C, Robinson A (2017) Dementia knowledge assessment scale (DKAS): Confirmatory factor analysis and comparative subscale scores among an international cohort. *BMC Geriatr* **17**, 168.
- [27] Elvish R, Burrow S, Cawley R, Harney K, Graham P, Pilling M, Gregory J, Roach P, Fossey J, Keady J (2014) 'Getting to Know Me': The development and evaluation of a training programme for enhancing skills in the care of people with dementia in general hospital settings. *Aging Ment Health* **18**, 481-488.
- [28] Schneider J, Schönstein A, Teschauer W, Kruse A, Teichmann B (2020) Hospital staff's attitudes toward and knowledge about dementia before and after a two-day dementia training program. *J Alzheimers Dis* **77**, 355-365.
- [29] Schneider J, Teichmann B, Kruse A (2019) The impact of dementia training on hospital staff's knowledge and attitudes. *Innov Aging* **3**, S727-S727.
- [30] Polit DF, Beck CT (2011) *Nursing research: Generating and assessing evidence for nursing practice*, 9th ed., Wolters Kluwer, Philadelphia.
- [31] Toye C, Lester L, Popescu A, McInerney F, Andrews S, Robinson AL (2014) Dementia Knowledge Assessment Tool Version Two: Development of a tool to inform preparation for care planning and delivery in families and care staff. *Dementia (London)* **13**, 248-256.
- [32] Parra-Anguita L, Moreno-Cámara S, López-Franco MD, Pancorbo-Hidalgo PL (2018) Validation of the Spanish version of the dementia knowledge assessment tool 2. *J Alzheimers Dis* **65**, 1175-1183.
- [33] Gkioka M, Tsolaki M, Papagianopoulos S, Teichmann B, Moraitou D (2020) Psychometric properties of dementia attitudes scale, dementia knowledge assessment tool 2 and confidence in dementia scale in a Greek sample. *Nurs Open* **7**, 1623-1633.
- [34] Piovezan M, Miot HA, Garuzi M, Jacinto AF (2018) Cross-cultural adaptation to Brazilian Portuguese of the Dementia Knowledge Assessment Tool Version Two: DKAT2. *Arq Neuropsiquiatr* **76**, 512-516.
- [35] Annear MJ, Eccleston CE, McInerney FJ, Elliott K-EJ, Toye CM, Tranter BK, Robinson AL (2016) A new standard in dementia knowledge measurement: Comparative validation of the Dementia Knowledge Assessment Scale and the Alzheimer's Disease Knowledge Scale. *J Am Geriatr Soc* **64**, 1329-1334.
- [36] Annear MJ, Otani J, Li J (2017) Japanese-language Dementia Knowledge Assessment Scale: Psychometric performance, and health student and professional understanding. *Geriatr Gerontol Int* **17**, 1746-1751.
- [37] Carnes A, Barallat-Gimeno E, Galvan A, Lara B, Lladó A, Contador-Muñana J, Vega-Rodríguez A, Escobar MA, Piñol-Ripoll G (2021) Spanish-dementia knowledge assessment scale (DKAS-S): Psychometric properties and validation. *BMC Geriatr* **21**, 302.
- [38] Zhao Y, Eccleston CE, Ding Y, Shan Y, Liu L, Chan HYL (2022) Validation of a Chinese version of the dementia knowledge assessment scale in healthcare providers in China. *J Clin Nurs* **31**, 1776-1785.
- [39] Sung H-C, Su H-F, Wang H-M, Koo M, Lo RY (2021) Psychometric properties of the dementia knowledge assessment scale-traditional Chinese among home care workers in Taiwan. *BMC Psychiatry* **21**, 515.
- [40] Akyol MA, Gönen Şentürk S, Akpınar Söylemez B, Küçükgüçlü Ö (2021) Assessment of Dementia Knowledge Scale for the nursing profession and the general population: Cross-cultural adaptation and psychometric validation. *Dement Geriatr Cogn Disord* **50**, 170-177.
- [41] Hambleton RK (2001) The next generation of the ITC test translation and adaptation guidelines. *Eur J Psychol Assess* **17**, 164-172.
- [42] Elvish R, Burrow S, Cawley R, Harney K, Pilling M, Gregory J, Keady J (2016) 'Getting to Know Me': The second phase roll-out of a staff training programme for supporting people with dementia in general hospitals. *Dementia (London)* **17**, 96-109.
- [43] Nienaber A, Wabnitz P, Volmar B, Löhr M (2015) Erste empirische Ergebnisse der wissenschaftlichen Begleitung des Projektes 'Interdisziplinärer Demenzkoordinator'. In: "Sprachen": Eine Herausforderung für die psychiatrische Pflege in Praxis - Management - Ausbildung - Forschung; Vorträge, Workshops und Posterpräsentationen, Schoppmann S, ed., Verlag Berner Fachhochschule, Fachbereich Gesundheit: Bern, pp. 148-151.

- [44] Löhr M, Baumeister M, Meißner B, Noelle R (2015) Lern von mir! Ein Trainingsprogramm zur Schulung von Klinikpersonal in Allgemeinkrankenhäusern im Umgang mit Menschen mit Demenz. In: "Sprachen": Eine Herausforderung für die psychiatrische Pflege in Praxis - Management - Ausbildung - Forschung; Vorträge, Workshops und Posterpräsentationen, Schoppmann S, ed., Verlag Berner Fachhochschule, Fachbereich Gesundheit: Bern, pp. 124-128.
- [45] Fachhochschule der Diakonie Knowledge in Dementia Scale - deutsche Version, [https://www.fh-diakonie.de/obj/Bilder\\_und\\_Dokumente/Lern\\_von\\_mir/Evaluationsblatt.pdf](https://www.fh-diakonie.de/obj/Bilder_und_Dokumente/Lern_von_mir/Evaluationsblatt.pdf).
- [46] IBM Corp (2020) *IBM SPSS Statistics for Windows*, IBM Corp, Armonk, NY.
- [47] Arbuckle JL (2014) *Amos*, IBM SPSS, Chicago, IL.
- [48] Tavakol M, Dennick R (2011) Making sense of Cronbach's alpha. *Int J Med Educ* **2**, 53-55.
- [49] Cronbach LJ (1951) Coefficient alpha and the internal structure of tests. *Psychometrika* **16**, 297-334.
- [50] Taber KS (2018) The use of Cronbach's alpha when developing and reporting research instruments in science education. *Res Sci Educ* **48**, 1273-1296.
- [51] McGraw KO, Wong SP (1996) Forming inferences about some intraclass correlation coefficients. *Psychol Methods* **1**, 30-46.
- [52] Koo TK, Li MY (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* **15**, 155-163.
- [53] Jackson DL, Gillaspay JA, Purc-Stephenson R (2009) Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychol Methods* **14**, 6-23.
- [54] Schreiber JB, Nora A, Stage FK, Barlow EA, King J (2006) Reporting structural equation modeling and confirmatory factor analysis results: A review. *J Educ Res* **99**, 323-338.
- [55] Möhring W, Schlütz D (2013) *Handbuch standardisierte Erhebungsverfahren in der Kommunikationswissenschaft*, Springer Fachmedien Wiesbaden; Imprint: Springer VS, Wiesbaden.
- [56] Kaiser HF, Rice J (1974) Little jiffy, mark IV. *Educ Psychol Meas* **34**, 111-117.
- [57] Tabachnick BG, Fidell LS (2014) *Using multivariate statistics*, International edition, Pearson, Harlow, Essex.
- [58] Mann HB, Whitney DR (1947) On a test of whether one of two random variables is stochastically larger than the other. *Ann Math Statist* **18**, 50-60.
- [59] Divine GW, Norton HJ, Barón AE, Juarez-Colunga E (2018) The Wilcoxon-Mann-Whitney procedure fails as a test of medians. *Am Stat* **72**, 278-286.
- [60] Hart A (2001) Mann-Whitney test is not just a test of medians: Differences in spread can be important. *BMJ* **323**, 391-393.
- [61] (2008) Kolmogorov-Smirnov Test. In *The Concise Encyclopedia of Statistics*, Springer New York, New York, NY, pp. 283-287.
- [62] Fritz CO, Morris PE, Richler JJ (2012) Effect size estimates: Current use, calculations, and interpretation. *J Exp Psychol Gen* **141**, 2-18.
- [63] Faul F, Erdfelder E, Lang A-G, Buchner A (2007) G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* **39**, 175-191.
- [64] Faul F, Erdfelder E, Buchner A, Lang A-G (2009) Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behav Res Methods* **41**, 1149-1160.
- [65] Kang H (2021) Sample size determination and power analysis using the G\*Power software. *J Educ Eval Health Prof* **18**, 17.
- [66] Rattray J, Jones MC (2007) Essential elements of questionnaire design and development. *J Clin Nurs* **16**, 234-243.
- [67] Ferretich S (1991) Focus on psychometrics. Aspects of item analysis. *Res Nurs Health* **14**, 165-168.
- [68] Piedmont RL (2014) Inter-item correlations. In *Encyclopedia of Quality of Life and Well-Being Research*, Michalos AC, ed., Springer: Dordrecht, pp. 3303-3304.
- [69] Nunnally JC, Bernstein IH (1995) *Psychometric theory*, 3. ed., 3. print, MacGraw-Hill series in psychology, McGraw-Hill, New York.
- [70] Šimković M, Tráuble B (2019) Robustness of statistical methods when measure is affected by ceiling and/or floor effect. *PLoS One* **14**, e0220889.
- [71] Mondschein CF, Monda C (2019) *Fundamentals of Clinical Data Science: The EU's General Data Protection Regulation (GDPR) in a Research Context*, Cham (CH).
- [72] Bagozzi RP, Yi Y (2012) Specification, evaluation, and interpretation of structural equation models. *J Acad Mark Sci* **40**, 8-34.
- [73] Bagozzi RP, Yi Y (1988) On the evaluation of structural equation models. *J Acad Mark Sci* **16**, 74-94.
- [74] Maydeu-Olivares A, Shi D (2017) Effect sizes of model misfit in structural equation models. *Methodology* **13**, 23-30.
- [75] Čoso B, Mavrinac S (2016) Validation of Croatian version of Dementia Attitudes Scale (DAS). *Suvr Psihol* **19**, 5-22.
- [76] Çetinkaya A, Elbî H, Altan S, Rahman S, Aydemir Ö (2020) Adaptation of the dementia attitudes scale into Turkish. *Noro Psikiyatrisi Ars* **57**, 325-332.
- [77] Peng A, Moor C, Schelling HR (2011) Einstellungen zu Demenz: Übersetzung und Validierung eines Instruments zur Messung von Einstellungen gegenüber Demenz und demenzkranken Menschen (Teilprojekt 1), 1-33.
- [78] Annear MJ, Toye CM, Eccleston CE, McInerney FJ, Elliott K-EJ, Tranter BK, Hartley T, Robinson AL (2015) Dementia Knowledge Assessment Scale: Development and preliminary psychometric properties. *J Am Geriatr Soc* **63**, 2375-2381.
- [79] Xia Y, Yang Y (2019) RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behav Res Methods* **51**, 409-428.
- [80] Lorio AK, Gore JB, Warthen L, Housley SN, Burgess EO (2016) Teaching dementia care to physical therapy doctoral students: A multimodal experiential learning approach. *Gerontol Geriatr Educ* **38**, 313-324.
- [81] Jack-Waugh A, Ritchie L, MacRae R (2018) Assessing the educational impact of the dementia champions programme in Scotland: Implications for evaluating professional dementia education. *Nurse Educ Today* **71**, 205-210.
- [82] Choi B, Pak A (2005) A catalog of biases in questionnaires. *Prev Chronic Dis* **2**, A13.
- [83] MacRae R, Gamble C, Ritchie L, Jack-Waugh A (2022) Testing the sensitivity of two dementia knowledge instruments in dementia workforce education. *Nurse Educ Today* **108**, 105210.
- [84] Gamble CM (2022) *Dementia Knowledge - Psychometric Evaluation in Healthcare Staff and Students*, Doctoral Thesis.
- [85] Statistisches Bundesamt (2019) Bevölkerung im Alter von 15 Jahren und mehr nach allgemeinen und beruflichen Bildungsabschlüssen nach Jahren,

- <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Bildungsstand/Tabellen/bildungsabschluss.html>.
- [86] Wright KB (2005) Researching internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *J Comput Mediat Commun* **10**, JCMC1034.
- [87] Hunter L (2012) Challenging the reported disadvantages of e-questionnaires and addressing methodological issues of online data collection. *Nurse Res* **20**, 11-20.
- [88] Ball HL (2019) Conducting online surveys. *J Hum Lact* **35**, 413-417.
- [89] Henrich J, Heine SJ, Norenzayan A (2010) The weirdest people in the world? *Behav Brain Sci* **33**, 61-83; discussion 83-135.