



Development of a Scale to Assess Knowledge about Suicide Postvention using Item Response Theory

INGO W. NADER, MSc, MEng, THOMAS NIEDERKROTENTHALER, MD, PhD, MMSc, ANNE H. E. SCHILD, MSc, INGRID KOLLER, PhD, ULRICH S. TRAN, DSc, MSc, NESTOR D. KAPUSTA, MD, PhD, GERNOT SONNECK, MD, AND MARTIN VORACEK, DSc, DMSc, PhD

Knowledge about suicide postvention (KSPV) is an important distal outcome in the evaluation of suicide prevention programs that focus on the bereaved. However, most scales are specifically tailored to the evaluation study in question and psychometric properties are often unsatisfactory. Therefore, we developed the KSPV scale. Scale properties were investigated with Rasch trees, a newly developed method in the framework of item response theory. Additionally, we provide cues for convergent validity. In summary, the scale shows satisfactory properties for assessing KSPV and could be used to evaluate suicide postvention programs more effectively.

Many countries have established suicide prevention plans (Beautrais, 2005; Taylor, Kingdom, & Jenkins, 1997). One suggested measure in these plans is education; that is, providing knowledge about suicidal behavior and about how to identify and support people in suicidal crises (e.g., Isaac et al.,

2009). Furthermore, acquisition of knowledge is an important goal in preparing caregivers for their work in suicide prevention (Ramsay, 2004).

Most evaluation studies of suicide prevention programs assess changes in knowledge about suicide, at least as a distal

INGO W. NADER, Department of Basic Psychological Research and Research Methods, School of Psychology, University of Vienna, Vienna, Austria; THOMAS NIEDERKROTENTHALER, Department of General Practice and Family Medicine, Center for Public Health, Medical University of Vienna, Vienna, Austria, and Department of Medical Psychology, Center for Public Health, Medical University of Vienna, Vienna, Austria; ANNE H. E. SCHILD, INGRID KOLLER, and ULRICH S. TRAN, Department of Basic Psychological Research and Research Methods, School of Psychology, University of Vienna, Vienna, Austria; NESTOR D. KAPUSTA, Department of Psychoanalysis and Psychotherapy, Medical University of Vienna, Vienna, Austria; GERNOT SONNECK, Department of Medical Psychology, Center for Public Health, Medical

University of Vienna, Vienna, Austria, and Crisis Intervention Center and Ludwig Boltzmann Institute for Social Psychiatry, Vienna, Austria; MARTIN VORACEK, Department of Basic Psychological Research and Research Methods, School of Psychology, University of Vienna, Vienna, Austria.

The authors would like to thank the Wiener Werkstätte for Suicide Research, especially Arno Herberth and Thomas Stompe, as well as Karin Böck, Renate Binder, and Sissy Hanke for their help with item construction.

Address correspondence to Ingo W. Nader, Department of Basic Psychological Research and Research Methods, School of Psychology, University of Vienna, Liebiggasse 5, A-1010 Vienna, Austria; E-mail: ingo.nader@univie.ac.at

outcome (Isaac et al., 2009). However, those evaluation studies often use subjective self-ratings of perceived (as opposed to actual) knowledge (e.g., Clark, Matthieu, Ross, & Knox, 2010) or ad hoc scales tailored to a specific program (e.g., Bean & Baber, 2011). Hardly any evaluation study used a scale with established psychometric properties, and many of the specifically tailored scales assess knowledge about epidemiological facts. These facts can be relevant for recognizing warning signs at population level, but are less suitable when it comes to offer support for people at risk at the individual level, like in tertiary prevention (i.e., suicide postvention). The latter is an important part of suicide prevention because of known contagion effects and the increased risk of psychological problems in survivors that have been exposed to suicide (Mackesy-Amiti, Fendrich, Libby, Goldenberg, & Grossman, 1996).

A recent review of the effectiveness of suicide postvention programs also reports knowledge as an important part of the program evaluation (Szumilas & Kutcher, 2011). For one of the scales used to quantify knowledge about suicide postvention (KSPV), an attempt was made to develop a stand-alone instrument, the Preparing for Crisis Knowledge Test (PFC-KT; Grossman et al., 1995). However, the test seems to be rather heterogeneous in its contents and has low internal consistency (Mackesy-Amiti et al., 1996). Moreover, the scale focuses on school-based programs (many items relate to structural prerequisites in schools that facilitate postvention after suicides); hence, the PFC-KT is only suitable for this specific population.

The most widely used scale on knowledge about suicide is perhaps the Revised Facts on Suicide Quiz (RFOS; Hubbard & McIntosh, 1992). However, this scale focuses on knowledge about suicide in general and not on suicide postvention specifically. It contains a number of questions that relate to epidemiological facts about suicide (e.g., the most frequently used method). Furthermore, it also has relatively low values of internal consistency (e.g.,

Kölves, Tran, & Voracek, 2007; Voracek, Fisher, Loibl, Tan, & Sonneck, 2008; Voracek, Loibl, Swami, Vintila, Kölves, Sinniah, et al., 2008). It is hardly ever used in studies that evaluate suicide prevention or postvention programs.

To the best of our knowledge, there is no scale assessing KSPV that has satisfactory psychometric properties. Hence, the aim of this study was to develop a scale measuring KSPV, specifically focusing on support for the bereaved by suicide. As knowledge scales are known to be heterogeneous in content, we applied a new method of analysis in the framework of item response theory (IRT), the so-called Rasch trees (Strobl, Kopf, & Zeileis, 2011). This method allows identifying subgroups of participants who respond differently to specific items. This information can be used to gain additional insight into the construct of KSPV. We also investigated cues for validity of the newly constructed measure.

METHODS

Participants

For the construction of the KSPV scale, three independent samples were used. In the first sample, a preliminary version of the scale was tested, the second sample was used for item selection, and the third sample was used for confirmatory analyses and cross-validation.

Sample 1. The first sample consisted of 620 German-speaking volunteers (60.2% women, 2.3% missing). In terms of nationality, 74.5% were Austrian, 19.8% were German, and 3.2% had other nationality (2.4% missing). Mean age was 30.2 years ($SD = 13.0$ years). Regarding education, 7.3% reported primary education as their highest education, 19.4% had an apprenticeship diploma, 52.7% had completed secondary education, and 20.2% had a university degree (0.5% missing). About 39.5% of the sample were students, and 47.7% reported knowing someone who had engaged in suicidal behavior (see Demographics).

Sample 2. This sample consisted of 626 German-speaking volunteers (61.3% women, 0.2% missing). Participants were mainly Austrian (66.3%), 18.8% were German, and 12.8% reported other nationality (2.1% missing). Mean age was 29.5 years ($SD = 12.3$ years). Regarding their highest educational degree, 7.0% had primary education, 13.7% reported an apprenticeship diploma, 61.0% completed secondary education, and 17.3% reported having a university degree (1.0% missing). Of the sample, 54.3% were students, and 48.6% knew someone who had engaged in suicidal behavior.

For the Rasch tree analyses, participants who answered all items of the KSPV scale correctly or incorrectly were excluded (they conveyed no information on how difficult the items are in relation to each other). Also, participants with missing values had to be excluded because of software limitations. Hence, the final sample for these analyses consisted of 589 participants.

Sample 3. For the confirmatory analyses, we aimed at achieving a balanced sample regarding sex ($n = 681$; 49.9% women, 5.9% missing). Nationality was 81.2% Austrian, 7.8% German, and 4.4% other nationality (6.6% missing). Mean age was 31.7 years ($SD = 13.4$ years). For highest education, 4.4% reported primary education, 10.3% had an apprenticeship diploma, 53.2% completed secondary education, and 25.8% had a university degree (6.3% missing). About 38.8% of the participants were students, and 47.7% knew someone who had engaged in suicidal behavior. With the same exclusion criteria as for sample 2, the sample for the Rasch tree analyses comprised 593 participants.

Measures

Knowledge about Suicide Postvention. The first version of the KSPV scale consisted of 10 items. Item content was developed by the first author (IWN) by screening the scientific literature and then discussed among all authors and with further experts in the field of suicide research and crisis intervention

(who also offer counseling for bereaved by suicide). Relating to item content, the main aim was to provide items practically relevant to the counseling of individuals bereaved by suicide. For that reason, questions about epidemiological facts were not part of the questionnaire. After initial analyses in the first sample, items were rephrased where necessary and additional items were constructed, resulting in a 21-item version. Starting from this version, items were excluded based on their content and psychometric properties (see Data Analysis). The final version of the KSPV scale consisted of 15 items, with three response options each: *true*, *false*, and *don't know*. The answer "don't know" was treated as incorrect. For a complete list of items see Table 1. Psychometric properties are discussed in detail in the Results section.

Items were constructed and validated in German and later translated into English by three of the authors (IWN, AHES, and TN) using the parallel blind technique (Behling & Law, 2000). The resulting translations were discussed with native English speakers from the field of public health, and wording of items was adapted and revised accordingly.

Revised Facts on Suicide Quiz. For assessing general knowledge about suicide and to validate the KSPV scale, the RFOS was used (Hubbard & McIntosh, 1992; German form: Voracek, Tran, & Sonneck, 2008). This measure consists of 32 items assessing various facts about suicide and suicidal behavior. All items have three response categories (either *true/false/don't know*, or three possible answers; "don't know" was treated as incorrect). Cronbach α in the three samples was .47, .62, and .60, respectively.

Demographics. Participants were asked to provide demographic information on their age, sex, marital status, highest educational qualification, as well as about their field of study (for students; samples 2 and 3) and their occupation (sample 3 only). The latter two were later recoded into two categories (related to caring profession vs. other) to validate the KSPV scale by group comparisons.

TABLE 1

Items and Psychometric Properties of the KSPV Scale. Instructions: "The following statements refer to facts from research regarding support for the bereaved by suicide. Please tick the answer that you think is correct"

	Item Text	Sample 2		Sample 3	
		%c	r_{it}	%c	r_{it}
(1)	Bereaved should better not see the deceased person before the funeral (F)	.51	–	.48	–
2	Bereaved should not be contacted by crisis intervention teams shortly after the death (F)	.69	.30	.67	.29
3	Avoiding talking about the suicide of a family member/friend is a good coping strategy (F)	.84	.44	.85	.35
4	Pharmacological treatment (tranquilizers) in the first hours after the bereavement should be avoided by all means (F)	.42	.19	.45	.27
5	Touching the body of a deceased person is safe from a medical point of view (T)	.41	.24	.50	.29
6	All people react the same way to a sudden and unexpected death (F)	.78	.34	.78	.24
7	Anger with the deceased is a normal reaction (T)	.56	.35	.53	.32
8	Children should not, by any means, see the deceased person (F)	.31	.26	.31	.31
(9)	Children playing normally right after the death is a sign of traumatization (F)	.42	–	.43	–
10	Children playing "being dead" or "funeral" after a suicide is a sign of traumatization (F)	.42	.29	.44	.35
11	Suicide should be concealed from children (F)	.61	.43	.60	.43
12	It is best to cope with the suicide of a loved one or friend on your own (F)	.83	.37	.80	.38
(13)	The younger a child is at the time of a loved one's suicide, the higher the child's suicide risk (T)	.10	–	.09	–
(14)	The shock phase after a suicide lasts at least a week (F)	.18	–	.14	–
15	Media interviews of the bereaved immediately after the suicide are a first step in coping with the death (F)	.69	.38	.70	.27
16	Grief after suicide should be shared with children (T)	.71	.44	.72	.43
17	Farewell rituals (e.g., lighting a candle, talking to the deceased, ...) are unbearable after a suicide (F)	.79	.50	.85	.40
18	In the days after a loved one's suicide it is beneficial for children to maintain their usual daily routines (T)	.69	.16	.63	.17
(19)	In the shock phase after a suicide, the risk of a suicide is highest (F)	.25	–	.26	–
(20)	It is a good coping strategy to divert a bereaved person (e.g., by reminding him/her of good times) (F)	.15	–	.15	–
21	The bereaved should, generally, take tranquilizers in the first hours after the death (F)	.43	.23	.48	.20

T, True; F, False; %c, percentage of correct answers; r_{it} , item-restscore correlation (not given for excluded items; item numbers of excluded items are given in parenthesis).

Additionally, participants were asked whether they had any experience with family members or close friends engaging in sui-

cidal behavior (i.e., completed or attempted suicide), as knowledge has been found to be higher in that group (MacDonald, 2007).

Procedure

In all samples, research assistants recruited participants from their personal contacts and by approaching people at the university. All participants provided informed consent, took part voluntarily, and were not remunerated for participation. They were ensured that they remained absolutely anonymous, and data were treated confidentially. After completion of the questionnaire, participants were debriefed using a standardized debriefing page with information about the study goals as well as contact information of institutions that offer counseling.

For samples 1 and 3, the questionnaire was also presented online, and 37.8% and 31.5% of these samples completed the online version, respectively. Participants were recruited via social media sites and Internet message boards unrelated to suicide topics.

Data Analysis

Additionally to standard means; that is, investigating internal consistency (Cronbach α) and factor structure, we used the Rasch model (Fischer & Molenaar, 1995; Rasch, 1960). This model can ensure that the scale is unidimensional and that the person sum score is an appropriate measure. Furthermore, item difficulty parameters (i.e., estimated difficulty resulting from the Rasch analysis) are not confounded by the latent ability (in this case, knowledge) of the participants, which is a major advantage over classical methods.

The fit of the Rasch model is usually tested by comparing the item difficulty parameters of two or more groups. If they are similar, all items of the scale have comparable properties in the two samples and the test is fair. If item parameters differ, the affected items are said to show differential item functioning (DIF). DIF means that items work differently for participants of different groups, even when they have the same latent ability (knowledge). Additionally to the

statistical examination, it is important to examine whether DIF is also theoretically explainable. The statistical part is usually tested with global likelihood ratio (LR) tests (Andersen, 1973). The groups are defined by external or internal split criteria. External split criteria can be chosen freely, with participants' sex, education, or age as probably the most common. Regarding internal split criteria, the raw score is both the most common and rigorous one. This criterion tests whether the item discrimination is similar enough for the Rasch model to hold, which is important for the raw score to be a fair measure.

When the Rasch model does not fit for the whole data set, it can be investigated whether it fits for certain subgroups of the sample and whether characteristics of these subgroups can explain DIF. One means of identifying groups is to use mixture Rasch models (Rost, 1990) in which groups are identified automatically by means of latent class analysis. However, it is often difficult to interpret DIF, as it remains unknown what constitutes the latent classes in which the response behavior is different.

Recently, a new method has been proposed to elude this drawback. Rasch trees (Strobl et al., 2011) can be used to automatically identify subgroups of persons who differ in their response behavior based on external split criteria. The method identifies the most important criterion by estimating how much the item difficulties change across specific subgroups, splits the sample accordingly, and iteratively repeats the analysis for the remaining criteria within these subsamples. Hence, the method also considers combinations of criteria. This facilitates identification of DIF and also its explanation: due to the known subgroups, the reasons for DIF can be investigated more easily (Reise & Waller, 2009), which can be used to gain additional insight into the measured construct.

For item selection, Rasch tree analyses and model tests of the Rasch model in the identified subgroups were performed iteratively. First, homogenous subgroups

were identified by a Rasch tree analysis using the external criteria sex, age, education, marital status, and experience with suicidal behavior in relatives or close friends. To confirm the fit of the Rasch model in the identified subgroups, we performed global LR tests using the internal split criterion (raw score median split). In sample 3, we additionally used one external criterion (online vs. paper-pencil testing). Furthermore, we used Wald tests (Fischer & Molenaar, 1995) to identify items that did not satisfy the Rasch model and applied Martin-Löf tests (Martin-Löf, 1973) in the identified subgroups, to examine unidimensionality of the scale. For the Martin-Löf test, it is standard practice to compare a group of easy items to a group of difficult items. For the model tests, we used a significance level of $\alpha = .01$ (to correct for multiple testing). After exclusion of unfitting items, these steps were repeated until the LR and Martin-Löf tests indicated that the Rasch model was valid in all subgroups identified by the Rasch tree analysis.

Factor analyses were conducted in MPlus (Muthén & Muthén, 2008) using an estimator based on polychoric correlations, which is appropriate for analyzing dichotomous data (weighted least squares mean and variance adjusted [WLSMV] estimator). All other analyses were performed in R 2.15.0 (R Development Core Team, 2012), using packages eRm (Mair, Hatzinger, & Maier, 2011) and psychotree version 0.12-1 (Strobl et al., 2011).

RESULTS

Analyses in Sample 1

Cronbach α of the first version of the KSPV scale was .66, which was higher than that of the RFOS in the same sample (Cronbach $\alpha = .47$), in spite of the fact that this version of the KSPV scale consisted of 10 items only. Discriminatory power (correlations of single items and total score with the respective items excluded) ranged from

$r_{it} = .23$ to $.40$ (mean $r_{it} = .33$). The rates of correct answers to the items ranged from $.34$ to $.89$.

Regarding factor structure, an exploratory factor analysis was performed. The scree plot indicated a one-factor solution to be most suitable (36.8% of variance explained by first factor, 13.3% by second factor). The mean total score was 5.99, indicating no floor or ceiling effects. Total scores correlated moderately with RFOS scores ($r = .37$, $p < .001$).

Analyses in Sample 2 with All KSPV Items

Internal consistency of the 21-item version of the KSPV scale was $\alpha = .72$, which again was higher than that of the RFOS in the same sample ($\alpha = .62$). However, discriminatory power of some of the newly constructed items was low (around $r_{it} = .10$ for items 14 and 20) and zero for item 13 ($r_{it} = -.04$, $p = .33$; Table 1). An exploratory factor analysis revealed seven eigenvalues greater than 1, but the scree plot suggested a one-factor solution (25.5% of variance explained by first factor, 8.7% by second factor).

Item Selection by Means of Item Response Theory

According to the strategy of analysis described above, the following items were excluded (in that order): 13, 14, 20, 19, 9, and 1. Item 13 may be considered an epidemiological fact less relevant for practical counseling and therefore was excluded. Items 14 and 19 might be problematic because of the term “shock” (which can be understood in a psychological or in a physical sense). In item 9, the term “playing normally” might be problematic, because respondents could think differently about what is considered “normal” after a traumatic event. For items 1 and 20, the correct answer may depend on the situation. Hence, these items were excluded.

The final scale consisted of 15 items. The Rasch model did not fit for the whole data set. The subgroups identified in the Rasch tree analysis for this version of the

scale are depicted in Figure 1. Age was the most important criterion, followed by education and sex. Marital status did not affect

the response behavior to the items in the scale, and neither did experience with suicidal behavior in friends or family. LR tests

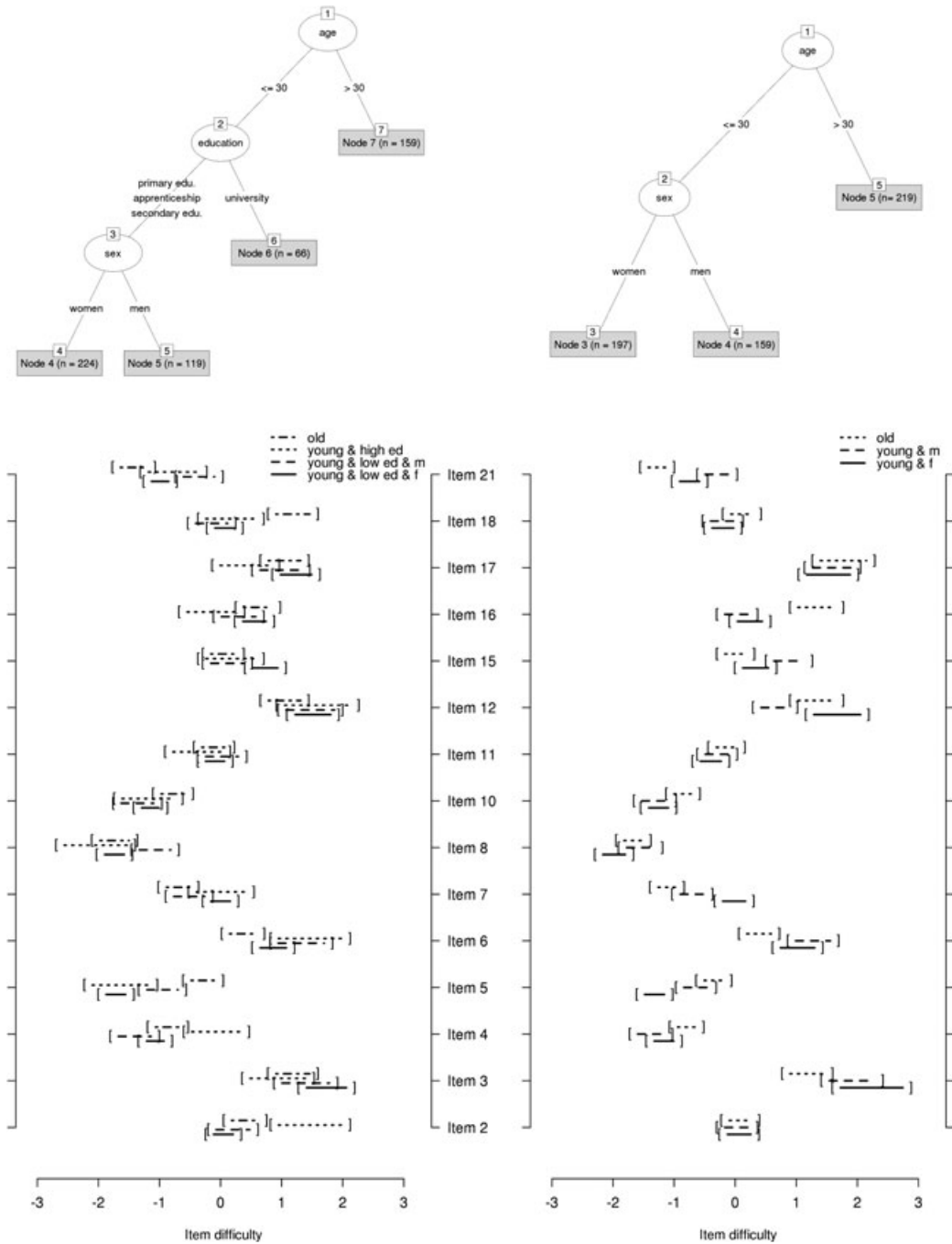


Figure 1. Groups of respondents identified by Rasch tree analysis and item difficulties in these groups with 95% confidence intervals (left: sample 2; right: sample 3). Only items of the final version of the KSPV scale are depicted.

in these subgroups were not significant (all $ps \geq .019$), and neither were the Martin-Löf tests (all $ps \geq .15$).

Analysis of the Final Version of the KSPV Scale in Sample 2

For the final 15-item version of the KSPV scale, Cronbach α was .72. Discriminatory power of items was between $r_{it} = .16$ and .50 (mean $r_{it} = .33$). The percentage of correct answers to the items ranged from .31 to .84. In an exploratory factor analysis, the scree plot indicated a one-factor solution (33.1% variance explained by first factor, 9.6% by second factor). Additionally, a confirmatory factor analysis indicated an acceptable fit of a one-factor solution (root mean square error of approximation [RMSEA] = .046 [CI, .038; .054]; comparative fit index [CFI] = .923; standardized root mean square error [SRMR] = .078). The Martin-Löf test for unidimensionality was also not significant at $\alpha = .01$ ($p = .027$).

Confirmatory Analyses in Sample 3

A Rasch tree analysis with LR tests for model fit (as performed in sample 2)

yielded subgroups that were similar to those in sample 2, except that education was not replicated as an important factor for DIF (Figure 1). A scree plot suggested a one-factor solution (32.3% of variance explained by first factor, 9.6% by second factor). A confirmatory factor analysis yielded, again, an acceptable fit of a one-factor model (RMSEA = .043 [.035, .051], CFI = .910, SRMR = .079), and the Martin-Löf test attested unidimensionality ($p = .352$).

Using DIF to Investigate Response Behavior

The Rasch tree analyses identified subgroups of persons that responded differently to certain items (Figure 1, Table 2). For example, item 21 (bereaved should generally take tranquilizers) was more difficult for younger men compared with older participants. It could be speculated that this group is generally less willing to deal with intense negative emotions, and therefore believes that the use of tranquilizers is generally recommended. A similar argument might apply to item 5 (touching the deceased is medically safe), which is also more difficult

TABLE 2
Group Differences in Item Difficulties of Single Items

Item Number	Sample 2				Sample 3		
	Old (grp 1)	Young and high ed (grp 2)	Young and low ed and m (grp 3)	Young and low ed and f (grp 4)	Old (grp 1)	Young and m (grp 2)	Young and f (grp 3)
2		1, 3, 4					
4		3, 4					
5	2, 4		4		3	3	
6		1	1			1	
7				1			1, 2
12							2
15						1	
16					2, 3		
18	2, 3, 4						
21			1			1	

grp, group; ed, education; m, male; f, female.

Entries in cells indicate that the item in question (row) was more difficult for the group indicated in the column header compared with the groups stated in the cell. Example: In sample 2, item 5 was more difficult for group 1, compared with groups 2 and 4.

for younger men compared with younger women. Touch can be a very intense and intimate gesture, hence accepting this as a possibility might be difficult when less willing to deal with intense emotions. Items 6 and 7 (all people react the same way; anger is a normal reaction) were more difficult for younger respondents, maybe because they have less experience with grief in general. For most items, however, item difficulties were rather similar (overlapping confidence intervals, Figure 1).

Validity

As a first cue for convergent validity, scores of the final (15-item) version of the KSPV scale were correlated with the RFOS. This scale measures knowledge about suicide in general and is therefore less specific than the KSPV scale. The scores of the two measures correlated moderately in two independent samples (samples 2 and 3; $r = .53$ and $r = .46$, both $ps < .001$), indicating that the KSPV scale measures a related, but not identical construct.

Furthermore, sum scores of the KSPV scale were compared between different subgroups of participants (again, in two independent samples). First, groups of students were compared. The focal group consisted of students in disciplines related to caring professions (e.g., psychology, medicine, teaching, social work; $n = 173$ and $n = 104$, for samples 2 and 3, respectively). This group was compared with all other students ($n = 167$ and $n = 160$, respectively). Participants in the focal group scored significantly higher in the KSPV scale, both in sample 2 ($t_{(338)} = 6.71$, $p < .001$, $d = 0.73$) and in sample 3 ($t_{(262)} = 4.12$, $p < .001$, $d = 0.52$). Effects were of medium-to-large size and comparable to those of the RFOS ($d = 0.56$ and 0.70 for samples 2 and 3, respectively).

Second, groups in different occupational fields were compared. This comparison was only performed for sample 3, as occupation was not part of the survey for sample 2. Again, the focal group consisted of partici-

pants working in fields related to caring professions (e.g., psychologists, therapists, teachers, social workers, paramedics; $n = 82$). Performance of this group was compared with the remaining participants that were employed ($n = 301$). Again, participants in the focal group scored higher on the KSPV scale ($t_{(381)} = 4.69$, $p < .001$, $d = 0.59$). This medium-sized effect was slightly higher than when the RFOS was used to compare these subgroups ($d = 0.44$).

Third, we compared participants who had experience with suicidal behavior in friends or family ($n = 304$ and $n = 325$ for samples 2 and 3, respectively) with those who had no such experience ($n = 320$ and $n = 316$, respectively). For the KSPV scale, effects were small but significant in sample 2 ($t_{(622)} = 4.02$, $p < .001$, $d = 0.33$) as well as in sample 3 ($t_{(639)} = 3.20$, $p = .001$, $d = 0.25$). For the RFOS, effects were smaller in sample 2 ($d = .16$) and non-existent in sample 3 ($t_{(639)} = 1.69$, $p = .09$, $d = 0.13$).

DISCUSSION

In this study, we report on the construction and analysis of a new measure for assessing knowledge about suicide postvention. Both classical as well as IRT-based analyses suggest that measurement properties are satisfactory, and first cues of validity indicate that the measure is suitable for assessing knowledge about suicide postvention.

The newly constructed KSPV scale correlated with an already established measure of knowledge about suicide in general (RFOS), and results of groups comparisons were as expected (participants related to caring professions or with experience with suicidal behavior in friends or family had higher scores). Effects were not only comparable, but even slightly higher for the KSPV scale compared with the RFOS. Hence, the KSPV scale seems to be a valid measure of knowledge about suicide postvention.

The IRT analyses identified subgroups with different response behavior (DIF).

However, for most items, item parameters were not extremely different in the subgroups (overlapping confidence intervals). Items of a knowledge test are diverse in content (e.g., Voracek et al., 2008) and usually not exchangeable (each item assesses a related, but different aspect of knowledge). Therefore, it is only natural that some subgroups perform differently on certain items. Also, presence of item-level DIF does not necessarily lead to biased scale scores (Reise & Waller, 2009). Furthermore, both factor analyses and IRT analyses suggested that a one-factor solution was adequate. Hence, we argue that the KSPV sum score can be used as a measure for knowledge about suicide.

Despite using a composite score, the information resulting from the Rasch tree analysis might still be valuable. For those items where different response behavior was identified for specific subgroups (nonoverlapping confidence intervals), sex, and age seem to be important predictors for the difficulty of these items. For example, making clear that touching the deceased is not dangerous from a medical point of view (item 5) might be important particularly for younger men. Touch is an important part of farewell rituals and can help coping with the loss. For the loss of a child, there are even guidelines encouraging touching the deceased (Dyregrov, 1989), and we believe that this might generalize to the loss of adults. Knowing someone who was engaged in suicidal behavior was not identified as a significant influence on DIF, although it was found to influence knowledge in the current study and in prior research (MacDonald, 2007). Hence, the test is a fair measure to compare these groups.

REFERENCES

- ANDERSEN, E. B. (1973). A goodness of fit test for the Rasch model. *Psychometrika*, *38*, 123–140.
- BEAN, G., & BABER, K. M. (2011). Connect: An effective community-based youth suicide prevention program. *Suicide and Life-Threatening Behavior*, *41*, 87–97.
- BEAUTRAIS, A. L. (2005). National strategies for the reduction and prevention of suicide. *Crisis*, *26*, 1–3.
- BEHLING, O., & LAW, K. S. (2000). *Translating questionnaires and other research instruments: Problems and solutions*. Thousand Oaks, CA: Sage.
- The KSPV scale showed satisfactory psychometric properties and validity. Therefore, it could be used to assess knowledge as a distal outcome of suicide prevention programs. It may provide a more rigorous outcome evaluation than other instruments, as was demanded in prior research (Mann et al., 2005). The KSPV scale could also aid at comparing different suicide prevention programs in terms of knowledge gains.
- One limitation of the current study is that the Rasch trees method is rather sensitive to sample characteristics. Education was found to be a factor that influences response behavior only in sample 2, but this was not replicated in sample 3. The reason might be that sample 3 contained fewer participants with lower education. Hence, replication of results in different samples is needed, as well as further research regarding the theoretical foundations of subgroup differences. Another limitation of this study is that the Rasch model is rather rigorous in its assumptions. However, it was possible to identify subgroups for which these assumptions hold, even though knowledge is known to be a rather heterogeneous concept (e.g., Voracek et al., 2008). Thereby, the Rasch tree method allows for investigation of possible causes in differences of the response behavior. This could facilitate the identification of specific target groups for education campaigns, primary prevention, intervention, as well as for postvention.
- In summary, we demonstrated that the KSPV scale is a suitable measure of knowledge about suicide postvention. It might provide a means for a more rigorous outcome evaluation in evaluation studies.

- CLARK, T. R., MATTHIEU, M. M., ROSS, A., & KNOX, K. L. (2010). Training outcomes from the samaritans of New York suicide awareness and prevention programme among community- and school-based staff. *British Journal of Social Work, 40*, 2223–2238.
- DYREGROV, A. (1989). Guidelines for help to families after the death of a child. *Tidsskrift for den Norske Laegeforening, 109*, 3408–3457.
- FISCHER, G. H., & MOLENAAR, I. W. (1995). *Rasch models: Foundations, recent developments, and applications*. New York: Springer.
- GROSSMAN, J., HIRSCH, J., GOLDENBERG, D., LIBBY, S., FENDRICH, M., MACKESY-AMITI, M. E., et al. (1995). Strategies for school-based response to loss: Proactive training and postvention consultation. *Crisis, 16*, 18–26.
- HUBBARD, R. W., & MCINTOSH, J. L. (1992). Integrating suicidology into abnormal-psychology classes: The revised facts on suicide quiz. *Teaching of Psychology, 19*, 163–166.
- ISAAC, M., ELIAS, B., KATZ, L. Y., BELIK, S.-L., DEANE, F. P., ENNS, M. W., et al. (2009). Gatekeeper training as a preventative intervention for suicide: A systematic review. *Canadian Journal of Psychiatry, 54*, 260–268.
- KÖLVES, K., TRAN, U. S., & VORACEK, M. (2007). Knowledge about suicide and local suicide prevalence: Comparison of Estonia and Austria. *Perceptual and Motor Skills, 105*, 3–7.
- MACDONALD, M. G. (2007). Undergraduate education majors knowledge about suicide. *Perceptual and Motor Skills, 105*, 373–378.
- MACKESY-AMITI, M. E., FENDRICH, M., LIBBY, S., GOLDENBERG, D., & GROSSMAN, J. (1996). Assessment of knowledge gains in proactive training for postvention. *Suicide and Life-Threatening Behavior, 26*, 161–174.
- MAIR, P., HATZINGER, R., & MAIER, M. (2011). *eRm: extended Rasch modeling*. Retrieved from <http://CRAN.R-project.org/package=eRm>
- MANN, J. J., APTER, A., BERTELOTE, J., BEAUTRAIS, A. L., CURRIER, D., HAAS, A., et al. (2005). Suicide prevention strategies: A systematic review. *Journal of the American Medical Association, 294*, 2064–2074.
- MARTIN-LÖF, P. (1973). *Statistical models: Notes from seminars 1969–1970, prepared by Rolf Sundberg*. Stockholm: Institutet för Försäkringsmatematik och Matematisk Statistik vid Stockholms Universitet.
- MUTHÉN, L. K., & MUTHÉN, B. O. (2008). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.
- R Development Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- RAMSAY, R. (2004). New developments in suicide intervention training. *Suicidologi, 9*, 10–12.
- RASCH, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Pedagogiske Institut.
- REISE, S. P., & WALLER, N. G. (2009). Item response theory and clinical measurement. *Annual Review of Clinical Psychology, 5*, 27–48.
- ROST, J. (1990). Rasch models in latent classes: An integration of 2 approaches to item analysis. *Applied Psychological Measurement, 14*, 271–282.
- STROBL, C., KOPF, J., & ZEILEIS, A. (2011). *A new method for detecting differential item functioning in the Rasch model*. Working Papers in Economics and Statistics, Research Platform Empirical and Experimental Economics, No. 2011-01. University Innsbruck. Retrieved from <http://EconPapers.RePEc.org/RePEc:inn:wpaper:2011-01>
- SZUMILAS, M., & KUTCHER, S. (2011). Post-suicide intervention programs: A systematic review. *Canadian Journal of Public Health, 102*, 18–29.
- TAYLOR, S., KINGDOM, D., & JENKINS, R. (1997). How are nations trying to prevent suicide? An analysis of national suicide prevention strategies. *Acta Psychiatrica Scandinavica, 95*, 457–463.
- VORACEK, M., FISHER, M. L., LOIBL, L. M., TAN, H., & SONNECK, G. (2008). Beliefs about the genetics of suicide in Canadian students: Cross-language validation of the Beliefs in the Inheritance of Risk Factors for Suicide Scale (BIRFSS). *Psychiatry and Clinical Neurosciences, 62*, 271–278.
- VORACEK, M., LOIBL, L. M., SWAMI, V., VINTILA, M., KÖLVES, K., SINNAH, D., et al. (2008). The Beliefs in the Inheritance of Risk Factors for Suicide Scale (BIRFSS): Cross-cultural validation in Estonia, Malaysia, Romania, the United Kingdom, and the United States. *Suicide and Life-Threatening Behavior, 38*, 688–698.
- VORACEK, M., TRAN, U. S., & SONNECK, G. (2008). Psychometric properties of the Revised Facts on Suicide Quiz in Austrian medical and psychology undergraduates. *Death Studies, 32*, 937–950.

Manuscript Received: July 2, 2012

Revision Accepted: October 17, 2012