



Security Culture Survey (SCS) Version 2 - Technical Document

This document describes the process utilized by KnowBe4 to develop, test, and validate the Security Culture Survey (SCS) Version 2, released from Beta in June of 2025. It also outlines the criteria by which it passed the beta testing phase and replaced the original version released in July of 2020. The process outlined is the culmination of many man-hours by a team at KnowBe4 in partnership with faculty at the University of Ljubljana over two years to provide the best possible instrument for measuring security culture. The team is comprised of the following members:

Principal Investigator:

- Dr. Gregor Petrič, Full Professor of Social Informatics and Chair of the Research Centre for Social Methodology and Informatics, University of Ljubljana

Research Colleagues:

- Dr. John N Just, Ed.D., Chief Learning Officer, KnowBe4
- Miha Matjašič, Lecturer - University of Ljubljana · Faculty of Education, University of Ljubljana
- Isida Drake, Senior Vice President of Security & Compliance eLearning, KnowBe4
- Elizabeth Smith, Senior Data Scientist, KnowBe4

Research Assistants:

- Tyler Balding, Senior Director of Courseware Quality Assurance, KnowBe4
- Priscila Mottola, Director of Language Quality, KnowBe4
- Gage VanGorden, Courseware Documentation and Assessment Specialist, KnowBe4
- Austin Pearlman, Senior Instructional Designer, KnowBe4

What is SCS, and how does it work?

SCS was introduced in 2020 as a method to measure an organization's security culture across seven dimensions that determine the human-related resilience of organizations against information security incidents, breaches, and attacks. The assessment of seven dimensions provides an organization with an overall security culture benchmark in addition to dimensional benchmarks, which reveal strengths and weaknesses of the human factor in organizational security. The results provide a baseline for managing human risks in the organization. Individuals at the organization are surveyed and asked to answer 31 items that address seven dimensions of security culture. The items were developed and refined following a strict scientific procedure of creating high-quality scales¹. Validity and reliability are the two most important criteria of high-quality scales; validity guarantees that the scale is accurate and measures what it is intended to measure, and reliability makes sure that the results of assessments are consistent across different conditions.

Technical development

The team reviewed the results, performance, and feedback received from three years of usage of the original version of SCS. A mixed methods approach² was used to analyze the experiences with the original SCS. With quantitative analysis, we analyzed the psychometric performance of individual items and analyzed the quality of data in terms of the presence of speedsters³ and straightliners⁴.

¹ DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage publications; Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, Vol 7(3), 309-319

² Blank, C. A. (2013). SAGE handbook of mixed methods in social & behavioral research. *Journal of music therapy*, 50(4), 321.

³ Respondents that complete the survey quicker than is theoretically necessary in order to read and answer all questions

⁴ Respondents that use the same response value for each item in a series of items 5 Orehek, Š., & Petrič, G. (2021). A systematic review of scales for measuring information security culture. *Information & Computer Security*, 29(1), 133-158.

Qualitative analysis (inductive thematic analysis) was used to analyze the feedback from the thousands of organizations that used the instrument over this timeframe. In addition, an extensive review of the literature on measuring security culture⁵ was performed. It was determined that no new domains were needed, but all new survey items would be developed to address the current cybersecurity landscape and consider the issues of clarity, difficulty, and appropriateness of the scale's responses.

The Security Culture Survey (SCS) was developed based on an information security culture and behavior instrument that was published in the *Journal of Cybersecurity*, a leading scientific publication in the field. This publication validates the SCS's scientific foundation and demonstrates its practical relevance. The following steps describe the technical procedure that was used to develop and test the revised and new items (See Figure 1 in the Appendix for a detailed visualization of the whole process).

Step one

On the basis of a review of performance and feedback on the original SCS and literature review, an initial pool of revised and new items was developed that contained 85 items measuring the domains of Attitudes, Behaviors, Cognition, Communication, Compliance, Norms, and Responsibilities. All items were first reviewed by multiple quality assurance (QA) team members and then critically evaluated for accuracy, clarity, relevance, and suitability by a panel of subject matter experts. The panel included assessment specialists, instructional designers, linguists, and multiple security experts. Content validity index (S-CVI/Ave⁶) was computed for all the items. On this basis, any items that did not reach high standard thresholds⁷ were either discarded or revised according to the panel's recommendations. All revisions were reviewed by the QA specialists again and reevaluated by the panel. The process resulted in a refined pool of 71 items.

⁵ Orehek, Š., & Petrič, G. (2021). A systematic review of scales for measuring information security culture. *Information & Computer Security*, 29(1), 133-158.

⁶ Scale-level content validity index, based on average method

⁷ S-CVI/Ave at least 0.9

Step two

The 71 items were tested in two rounds on representative samples of anonymous employees that were sourced from a survey recruitment organization that we partnered with. The respondents were recruited from various regions and industries and across a diverse range of socio-demographic backgrounds including race, ethnicity, education level, age, and sex.

These were collected and analyzed for possible bias that might be included in the items and for investigation of measurement invariance of the scales across different groups. More specifically, the following validity tests and analytical procedures were conducted:

- a. Structural validity, which refers to the extent to which the items of a scale represent the underlying dimensionality of the construct, was assessed with the confirmatory factor analysis (CFA), exploratory factor analysis (EFA), and internal consistency statistics⁸;
- b. Construct validity, which refers to the extent to which the items of a scale accurately reflect the essential properties of a scale/dimension, was assessed by computing the convergent and discriminant validity;
- c. Measurement invariance was assessed by multigroup CFA and is used to test whether a given scale is measured in the same way across different groups.

After exclusion of several items the scale demonstrated good structural validity as the statistics of the measurement model demonstrated a good fit to the data - a value of root mean square error of approximation (RMSEA) was less than .08, a standardized root mean square residual (SRMR) was less than .08, and a comparative fit index (CFI) was greater than .90. Internal consistency of the dimensions is good to excellent as Cronbach's alphas and McDonald's omegas are in the range between .7 and .9⁹.

⁸ Package lavaan in R was used for statistical analyses

⁹ The value of .7 is usually taken as a threshold for Cronbach's alpha and McDonell's Omega to claim the reliability (in terms of internal consistency) of a studied scale

The scale also demonstrated both convergent and discriminant validity as average variance extracted (AVE) values are above 0.5 and the square root of AVE for the individual dimensions was higher than the correlations of particular dimensions with other dimensions. The evidence also suggests that the scale's properties are invariant across regions, industries, ethnicity, race, education, age, and sex, as differences in model fit for stricter models are not significant ($\Delta\text{CFI} < .1$ and $\text{SRMR} < .03$)¹⁰.

The above procedure was repeated twice with adjustments made to the items after analysis and review. The first participant number (N) = 3,252 and the second sample N = 1,227. The outcome of the pilot tests was a scale with confirmed seven dimensions and 47 items altogether.

Step three

All 47 items were selected for an expert panel review. The expert panel consisted of three customers and two KnowBe4 experts who were not on the development team. They were instructed to review the existing item set and provide feedback. A protocol session was held to kick this process off, the review was completed independently by all the experts, and a feedback session was held with the research team. Changes were made to the items based on this feedback. A parsimonious solution with 31 items (3-6 items per individual dimension) was proposed to address the balance between respondent burden and data quality.

Step four

Next, a closed customer Beta was conducted across three multinational organizations. Altogether N=922 employees completed the survey. Using EFA, CFA, and internal consistency procedures, the parsimonious solution with 31 items was empirically confirmed. These analyses and qualitative analyses of open feedback resulted in minor adjustments to the scale and the format of the questionnaire.

¹⁰ This means that we can compare results of SCS across these groupings

Step five

A final test was conducted on N=1,022 KnowBe4 employees with analysis - EFA, CFA, and internal consistency procedures - leading to a final set of 31 items released into an open Beta in the KnowBe4 Modstore in March of 2024.

In addition to the data quality that stems from the procedure described above, the data collection process incorporates further mechanisms that guarantee high quality of data, such as randomization of items, reversed wording, soothing statements, missing value inspection, exclusion of speedsters and straightliners. The quality of items is inspected annually on the basis of customer data from different industries and (national) cultures.

Beta Status

This version of the survey was in Beta for 14 months after the launch date. Minor modifications were made from the resulting analysis, and then localization (deep translation) began. The version remained in Beta until all the language versions were completed and thoroughly tested for measurement equivalence, validity, and reliability. This process concluded in early 2025. Analysis of localized items was conducted, with six items updated across five languages.

Figure: Visualization of creating SCS V2

