KnowBe4
Human error. Conquered.

KnowBe4 Security Culture Survey

This document describes how the Security Culture Survey (SCS) was created and the methodology that is continuously applied to ensure its validity and reliability.

Developing scales (metrics) which prove to measure social science phenomena reliably requires a strict and rigorous scientific procedure[1]. Validity and reliability are the two most important criteria of quality scales; validity guarantees that the scale is accurate and measures what is intended to measure, and reliability takes care that the scale gives consistent results.

This procedure is based on a combination of theoretical deduction, synthesis, and strong empirical validation, on the basis of contemporary rigorous statistical methods. It is globally acknowledged by scientific academics as the most sophisticated and substantiated methodology for developing valid and reliable scales, and it encompasses the following steps:

1. The development of a large initial pool of assessment items;
2. Expert improvement and assessment of content validity;
3. Pilot testing;
4. Multiple cross-validations and removal of bias;
5. Validity and reliability testing;
6. Identification of the final set of items with sets of equivalent (banking) items.

To ensure the integrity of the assessment, KnowBe4 does not provide copies of the survey items[2].

A large pool of 139 items was initially developed on the basis of synthesis of scientific papers on security culture and theoretical application of relevant organizational culture models, which resulted in our own seven-dimensional model of security culture. This pool of items was evaluated for clarity, readability, and social desirability bias by experts trained in relevant subject matters: social psychology, survey design, social science methodology, social informatics, and information security.

On the basis of this evaluation, the items were optimized; some reworded and some excluded. A set of 101 optimized items was again evaluated by experts for relevance, and the estimation resulted in a Content Validity Index (S-CVI/Ave) of 0.96, indicating excellent content validity (threshold is 0.9).

In the next step, a pilot study was conducted on two companies (110 employees) in order to test the empirical behavior of the developed set of 101 items. A number of different statistical methods were used on the obtained data to test the proposed 7-dimensional model of security culture and to identify items prone to social desirability bias and satisficing. Items were excluded that correlated highly, had not enough variability, long completion times, or generally showed low levels of data quality.

1. De Vellis, R. F. (2003). Scale Development Theory and Applications (2nd ed., Vol. 26). Thousand Oaks, CA Sage Publications.
2. 'Items' are the questions or agreement statements that the survey respondents are asked and/or presented with.

Exploratory Factor Analysis (EFA) with a principal axis factoring method was conducted to obtain a solution with high enough commonalities (>0.2), a high enough measure of Kaiser-Meyer-Olkin (KMO), and significant Bartlett's test of sphericity (BTS)[3]. After further exclusion of items with too low communalities and/or factorial weights, a set of 45 items was further tested with Confirmatory Factor Analytical (CFA) approach, which is the strongest empirical test, giving direct insight into the empirical validity and reliability of proposed items.

The results of CFA suggest a very good fit of the model to the data with CFI=0.96, RMSEA=0.08 and SRMR=0.05. Cronbach's alphas for all individual dimensions was between 0.7 and 0.9, suggesting reliable (internally consistent) scale. Moreover, the value of Average Variance Extracted (AVE) was above 0.5, indicating good convergent validity of the scale and Maximum Shared Variance (MSV) was lower than AVE, indicating good discriminant validity.

Next, another (much larger) study was conducted on a sample of 38 companies and 5671 employees. The methodological intention of this study was to obtain a parsimonious set of items and identify banking[4] items. CFA approach was again used to reaffirm the 7-dimensional security culture model and the quality of the items.

On the basis of inter-item correlations and factorial weights, a set of replaceable items was selected. This procedure resulted in a parsimonious set of 28 items, and an additional set of 17 banking items. Structural equivalence was investigated between original and banking items by conducting multigroup CFA and the obtained solution demonstrated good fit to the data (CFI=0.93, RMSEA=0.09 and SRMR=0.06), suggesting that original and banking items are equivalent. Further sets of banking items are also being developed.

In addition to the data quality that stems from the procedure described above, the process of collecting data 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.

Strict adherence to the above procedures guarantees that the results of assessments are scientifically valid and bias-resistant. We consider this to be a valid and valuable assessment for the purposes of measuring security culture. The survey can be used multiple times to re-assess security culture and track changes over time.

3. After further exclusion of items with too low commonalities and/or factorial weights, we obtained a 7-factor solution with KMO=0.78 (well above the threshold of 0.5), confirming sampling adequacy of the analysis. BTS was statistically significant (P<0.001), indicating relevant factor solution. All communalities were above 0.2 and all factor weights were above 0.4 on individual dimensions.
4. Banking items are those that can be used to substitute or replace another assessment item in the survey without compromising the quality of the assessment.