

GALLUP CONSULTING®

The Clifton StrengthsFinder® 2.0 Technical Report: Development and Validation

February 2007
Updated March 2009

Jim Asplund
Gallup

Shane J. Lopez
Clifton Strengths Institute

Tim Hodges
Gallup

Jim Harter
Gallup

This document contains proprietary research, copyrighted materials, and literary property of Gallup, Inc. It is for your guidance only and is not to be copied, quoted, published, or divulged to others outside of your organization. Gallup®, Gallup Consulting®, Gallup Panel®, StrengthsQuest™, StrengthsFinder®, Clifton StrengthsFinder®, and each of the 34 Clifton StrengthsFinder theme names are trademarks of Gallup, Inc. All other trademarks are property of their respective owners.

This document is of great value to both you and Gallup, Inc. Accordingly, international and domestic laws and penalties guaranteeing patent, copyright, trademark, and trade secret protection safeguard the ideas, concepts, and recommendations related within this document.

No changes may be made to this document without the express written permission of Gallup, Inc.

Abstract

Gallup's Clifton StrengthsFinder is an online measure of personal talent that identifies areas where an individual's greatest potential for building strengths exists.

The 177-item pairs were based on the theory and research foundation associated with semi-structured personal interviews that had been used by Selection Research Incorporated and Gallup (Harter, Hayes, & Schmidt, 2004; Schmidt & Rader, 1999) for over 30 years. The measure, developed through rational and empirical processes, has been repeatedly subjected to psychometric examination; a summary of reliability and validity evidence gathered to date is presented.

The primary application of the Clifton StrengthsFinder, as the evaluation that initiates a strengths-based development process in work and academic settings, is discussed.

The authors would like to thank Steve Sireci of the University of Massachusetts for his invaluable advice and comments about this research.

For more information, please contact Jim Asplund at jim_asplund@gallup.com or 952-806-0630.

Introduction

The Clifton StrengthsFinder (CSF) has been subjected to repeated psychometric scrutiny by its developers. The purpose of this manuscript is to describe the development and application of the CSF and to summarize its psychometric support to date, in accordance with The Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999).

The Purpose of the Clifton StrengthsFinder

The validity of an assessment must be evaluated with respect to its intended purpose. The CSF is an online measure of personal talent that identifies areas where an individual's greatest potential for building strengths exists. By identifying one's top themes of talent, the CSF provides a starting point in the identification of specific personal talents, and the related supporting materials help individuals discover how to build upon their talents to develop strengths within their roles. The primary application of the CSF is as an evaluation that initiates a strengths-based development process in work and academic settings. As an omnibus assessment based on positive psychology, its main application has been in the work domain, but it has been used for understanding individuals and groups in a variety of settings — employee, executive team, student, family, and personal development.

The CSF is not designed or validated for use in employee selection or mental health screening. Given that CSF feedback is provided to foster intrapersonal development, comparisons across profiles of individuals are discouraged.

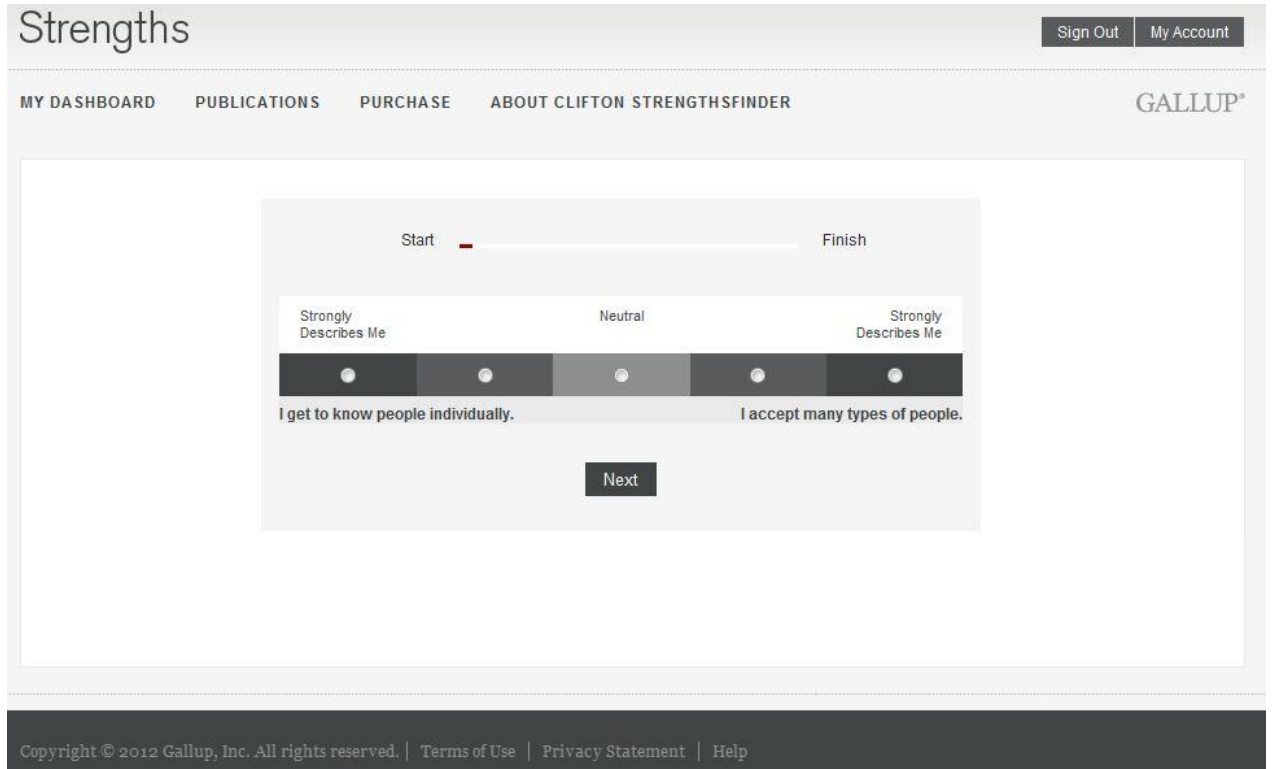
How the Clifton StrengthsFinder Is Scored

The precise scoring of the CSF is proprietary to Gallup, Inc. What follows is a general description of the scoring method so that readers can better understand the types of validity analyses that can and cannot be done.

The CSF is an online assessment in which each respondent is presented with 177 stimuli and makes 177 responses. Each item lists a pair of potential self-descriptors, such as “I like to help people.” The descriptors are placed as if anchoring opposite poles of a continuum. From that pair, the respondent is asked to choose the statement that best describes him or her, and also the extent to which that chosen option is descriptive of him or her. The participant is given 20 seconds to respond to a given item before the system moves on to the next item (developmental research showed that the 20-second limit resulted in a negligible item non-completion rate.)

An example of the item format is presented in Figure 1. This illustrates how each item is presented on the screen:

Figure 1



Most of these descriptors are associated with a “theme.” A theme is a category of talents, which are defined as recurring and consistent patterns of thought, feeling, or behavior. The CSF measures the presence of talent in 34 distinct themes. (A complete set of theme descriptions is included in Appendix A). For example, one of these themes is “Positivity.” Several statements within the CSF measure “Positivity,” and there are 33 other themes configured in the same way; that is, multiple statements measuring each theme. The number of statements varies by theme, as shown in Table 1.

Table 1: Number of Statements Associated With Each Theme

CSF Theme	Total Number of Items	CSF Theme	Total Number of Items
Achiever	6	Futuristic	8
Activator	7	Harmony	5
Adaptability	8	Ideation	7
Analytical	11	Includer	7
Arranger	13	Individualization	6
Belief	11	Input	5
Command	9	Intellection	10
Communication	9	Learner	8
Competition	7	Maximizer	7
Connectedness	8	Positivity	12
Consistency	8	Relator	8
Context	4	Responsibility	11
Deliberation	8	Restorative	6
Developer	10	Self-Assurance	13
Discipline	14	Significance	12
Empathy	6	Strategic	4
Focus	12	Woo	9

Some statements are linked to more than one theme. Also, for some items, each of the two statements within that item is linked to a separate theme. Thus, one response on an item can contribute to two or more theme scores. A proprietary formula assigns a value to each response category. Values for items in the theme are aggregated to derive a theme score.

The calculation of scores is based on the mean of the intensity of self-description. Scores are recorded in Gallup’s database as theme means, standard scores, and percentiles (derived from Gallup’s database of more than 3.9 million respondents at the time of this writing).

Results are presented to the respondent as a ranked ordering of Signature Themes, where the five highest scoring themes are provided to the respondent. Absolute scores are used to rank the themes, with percentiles against the database norms and theme reliabilities used as subsidiary ranking factors. These theme-rank data are also recorded into the Gallup database. Given the intended use of the CSF for intrapersonal development, these theme-rank data are the focus of feedback that is given to the respondent.

Strengths Theory

When educational psychologist Donald O. Clifton first designed the interviews that subsequently became the basis for the CSF, he began by asking, “What would happen if we studied what is right with people?” Thus emerged a philosophy of using talents as the basis for consistent achievement of excellence (strength). Specifically, the strengths philosophy is the assertion that individuals are able to gain far more when they expend effort to build on their greatest talents than when they spend a comparable amount of effort to remediate their weaknesses (Clifton & Harter, 2003).

Clifton hypothesized that these talents were “naturally recurring patterns of thought, feeling, or behavior that can be productively applied” (Hodges & Clifton, 2004, p. 257). “Strengths” are viewed as the result of maximized talents. Specifically, a strength is mastery created when one’s most powerful talents are refined with practice and combined with acquired relevant skills and knowledge. The CSF is designed to measure the raw talents that can serve as the foundation of strengths. Thus the purpose of the instrument is to identify Signature Themes of talent that serve as a starting point in the discovery of talents that can be productively applied to achieve success.

Development of the Clifton StrengthsFinder

Gallup, widely known for its polls (Gallup, 2004; Newport, 2004) and employee selection research (Harter, Hayes, & Schmidt, 2004; Schmidt & Rader, 1999) developed numerous semi-structured interviews to identify talent that could be enhanced and used to pursue positive outcomes in work and school. In the 1990s, under the leadership of Donald O. Clifton, Gallup developed the CSF as an objective measure of personal talent that could be administered online in less than one hour.

Clifton, over his 50-year career at the University of Nebraska, Selection Research Incorporated, and Gallup, studied “frames of reference” (Clifton, Hollingsworth, & Hall, 1952), teacher-student rapport (Dodge & Clifton, 1956), management (Clifton, 1970; 1975; 1980), and success across a wide variety of domains in business and education (Buckingham & Clifton, 2000; Clifton & Anderson, 2002; Clifton & Nelson, 1992). He based his research and practice on straightforward notions that stood the test of time and empirical scrutiny.

First, he believed that talents could be operationalized, studied, and capitalized upon in work and academic settings. Talents are manifested in life experiences characterized by yearnings, rapid learning, satisfactions, and timelessness. These trait-like “raw materials” are believed to be the products of normal healthy development and successful experiences over childhood and adolescence. “Strengths” are viewed as extensions of talent. More precisely, the strength construct combines talents with associated knowledge and skills and is defined as the ability to consistently provide near-perfect performance in a specific task. (Though labeled the Clifton StrengthsFinder, the instrument actually measures the talents that serve as the foundations for strengths development.)

Second, Clifton considered success to be closely associated with personal talents and strengths in addition to the traditional constructs linked with analytical intelligence. In accordance with those beliefs, he worked to identify hundreds of “themes” (categories) of personal talents that predicted work and academic success, and he constructed empirically-based, semi-structured interviews for identifying these themes. When developing the interviews, Clifton and analysts examined the prescribed roles of a person

(e.g., student, salesperson, administrator), visited the job site or academic setting, identified outstanding performers in these roles and settings, and determined the long-standing thoughts, feelings, and behaviors associated with situational success. Many of the interviews developed provided useful predictions of positive outcomes (Schmidt & Rader, 1999). These interviews subsequently were administered by Gallup analysts to more than two million individuals for the purposes of personal development and employee selection. In the mid-1990s, when considering the creation of an objective measure of talent, Clifton and colleagues systematically reviewed these interviews and the data they generated to capitalize on the accumulated knowledge and experience of Gallup's strengths-based practice.

The prominence of dimensions and items relating to motivation and to values in much of the interview research informed the design of an instrument that can identify those enduring human qualities. An initial pool of more than 5,000 items was constructed on the basis of traditional validity evidence. Given the breadth of talent assessed, the pool of items was considered large and diverse. A smaller pool was derived subsequent to quantitative review of item functioning and a content review of the representativeness of themes and items within themes (with an eye toward the construct validity of the entire assessment). Specifically, evidence used to evaluate the item pairs was taken from a database of criterion-related validity studies, including over 100 predictive validity studies (Schmidt & Rader, 1999). Factor and reliability analyses were conducted in multiple samples to assess the contribution of items to measurement of themes and the consistency and stability of theme scores — thereby achieving the goal of a balance between maximized theme information and efficiency in instrument length. During development phases, a number of sets of items were pilot tested. The items with the strongest psychometric properties (including item correlation to theme) were retained.

In 1999, a 35-theme version of the CSF was launched. After several months of data were collected, researchers revisited the instrument and, based on analyses of theme uniqueness and redundancy, decided on 180 items and 34 themes. Since 1999, some theme names have changed, but the theme descriptions have not changed substantially. (See Appendix A for a listing and descriptions of the 34 themes.)

Today, the CSF is available in more than 20 languages and is modifiable for individuals with disabilities. It has been taken by more than 3.9 million individuals all over the world. It is appropriate for administration to adolescents and adults with a reading level of grade 10 or higher. In 2006, Gallup researchers undertook a comprehensive review of CSF psychometrics, which led to some revisions in the instrument. Confirmatory studies (presented in a subsequent section) validated the 34-theme structure in both adult and student populations. In the course of reviewing more than one million cases in multiple studies, some possible improvements in theme validities and reliabilities were identified. Some of these improvements involved rescoring of existing items, whereas others required the addition of new items. These new items were drawn from Gallup's library of talent-related items, and from researchers' experience in building structured interviews and providing talent feedback. Finally, there were items that had been included in the 180-item version of the CSF, but never used in theme scores. A thorough review of each of these items showed many to be unnecessary as either distracters or scored items. They were consequently removed. The result of all of these item changes was a slight reduction in the length of the instrument, from 180 items to 177.

Researchers both inside and outside Gallup contributed a number of the investigations into the CSF's continuing reliability, validity, and applicability to both the general population and college students in particular. Those most recent studies have included:

- Confirmatory studies:
 - Sireci (University of Massachusetts): n = 10,000
 - Lopez (University of Kansas), Hodges (Gallup), Harter (Gallup): n = 601,049
 - Asplund (Gallup): n = 110,438
 - Asplund: n = 250,000
 - Asplund: n = 472,850
- Reliability studies:
 - Schreiner (Azusa Pacific): n = 438
 - Lopez, Harter, Hodges: n = 706
 - Asplund: n = 110,438
 - Asplund: n = 250,000
 - Asplund: n = 472,850
 - Asplund: n = 2,219
 - Asplund: n = 46,905
- Other validity studies:
 - Lopez, Hodges, Harter: n = 297
 - Schreiner: n = 438
 - Stone (Harvard): n = 278
- Utility studies:
 - Asplund: n = 90,000 employees in over 900 business units
 - Various additional case studies

Separately, each of these studies affirms the ongoing viability of the CSF. More importantly, the collective evidence of all this work is convergent regarding the psychometric properties of the CSF, as well as regarding the details of its validity.

Notwithstanding the confirmatory evidence provided by this body of research, Gallup researchers identified some areas in which the CSF could be improved psychometrically. In particular, it was observed that some of the items could be improved, removed, or replaced. As a logical first step to improving the psychometrics, Gallup researchers thoroughly examined each unscored statement to see whether it could be used to improve the performance of the assessment. Unscored statements that showed no utility were removed, if possible. (Several of the unscored statements are paired with a scored statement, and therefore are not subject for removal at this time.)

Administration and Feedback

Feedback varies in accordance with the reason the person completes the CSF. Summary scores are not provided to respondents. In most cases the respondent receives a report listing his or her top five talent themes — those in which the person received his or her highest scores, in order of intensity — the aforementioned Signature Themes. In other situations the respondent may review his or her sequence of all 34 themes, along with “action items” for each theme, in a personal feedback session with a Gallup consultant or in a supervised team-building session with colleagues. In programs designed to promote strengths-based development, feedback is often accompanied by instruction, experiential learning, and mentoring activities designed to help people make the most of their talents (i.e., develop strengths associated with occupational or educational roles).

As part of this update to the CSF, a new, more detailed type of feedback is provided: talent descriptions that go beyond the Signature Themes by looking at item-level responses. These “strengths insights” provide a more customized version of the respondent’s Signature Themes report featuring a more in-depth dive into the nuances of what makes him or her unique, using more than 5,000 new personalized strengths insights that Gallup researchers have discovered in recent years. This feedback based on both theme and item-level data provides a more rich description of the particular combination of responses provided by the participant.

Application: Strengths-Based Development

The CSF is often used as a starting point for self-discovery in Gallup strengths-based development programs. After a respondent has completed the assessment and talent feedback is provided, a set of developmental suggestions is customized to the individual’s Signature Themes and to his or her role to help integrate his or her talents into a more informed view of self. As the identification and integration stages of strengths development unfold, behavioral change is encouraged. Specifically, the strengths-based development process encourages individuals to build strengths by acquiring skills (i.e., basic abilities) and knowledge (i.e., what you know, including facts and meaning-making from experiences) that can complement their greatest talents in application to specific tasks.

The CSF’s intended purpose is to facilitate personal development and growth. It is intended and used as a springboard for discussion with managers, friends, colleagues and advisers, and as a tool for self-awareness. CSF results are viewed as a preliminary hypothesis to be verified with the respondent. Accordingly, feedback about talents and strengths development often forms the basis of further interventions that help individuals capitalize on their greatest talents and apply them to new challenges. For this application, the psychometric properties of the instrument are more than adequate.

Reliability

The reliability of a score is an estimate of its stability, or the portion of the score not due to random variation. For instruments like the CSF, two types of reliability estimates are generally used:

- **Internal Consistency.** In general, this involves looking at how well the items designed to measure the same thing produce the same results. Cronbach's alpha is a commonly used measure of this type of reliability.
- **Test-Retest Reliability.** This is employed by administering the instrument to the same sample at two different time periods. It is generally more difficult to acquire these data, as one has to get the respondent to complete the instrument twice.

Practical concerns will limit the number of items that can be used, but mathematically speaking, the more items the better (within reason). The same is true for validity; having more items should usually imply more coverage of the construct domain. There will be a cumulative effect on validity because each item taps into a slightly different aspect of the construct in question, or the criterion being predicted.

Estimates of internal consistency reliabilities for the CSF are included in Table 2. Estimates are provided from two independent samples — a random sample of 46,902 respondents from 2008, and the 2,219 respondents from the test-retest study described in the following section. (Alphas shown are from the initial test.) Readers will note the strong similarity of the two sets of results.

Table 2: Estimates of Internal Consistency Reliabilities

Theme	Alpha: (n = 46,902)	Alpha: Retest Sample (n = 2,219)
Achiever	0.66	0.67
Activator	0.62	0.59
Adaptability	0.71	0.71
Analytical	0.72	0.75
Arranger	0.64	0.65
Belief	0.60	0.62
Command	0.69	0.68
Communication	0.73	0.72
Competition	0.73	0.71
Connectedness	0.65	0.66
Consistency	0.65	0.62
Context	0.61	0.62
Deliberation	0.73	0.74
Developer	0.65	0.70
Discipline	0.78	0.78
Empathy	0.61	0.63
Focus	0.71	0.68
Futuristic	0.73	0.70
Harmony	0.68	0.65
Ideation	0.71	0.69
Includer	0.61	0.63
Individualization	0.56	0.55
Input	0.52	0.57
Intellection	0.70	0.72
Learner	0.75	0.78
Maximizer	0.72	0.64
Positivity	0.78	0.76
Relator	0.54	0.60
Responsibility	0.66	0.68
Restorative	0.70	0.67
Self-Assurance	0.68	0.67
Significance	0.70	0.70
Strategic	0.69	0.66
Woo	0.79	0.76

Cronbach's alpha is heavily biased by the number of items in a theme. In fact, it is very difficult to obtain extremely high alphas for an instrument that measures 34 dimensions, such as the CSF. Because the goal of the CSF was to create an efficient assessment that optimized validity, efforts to increase the alphas could potentially be detrimental to the purpose of the CSF. That is, alphas could be optimized by making the instrument considerably longer. For example, items could be added to "Context" to give it higher alphas, but when you add items you are removing degrees of freedom in the scale and essentially building a theme that measures just one aspect of Context. The high alpha could therefore cost content validity and, in all likelihood, criterion validity. Criteria such as alpha are meaningful only to the extent to which they reflect improved validity. In cases like that of the CSF (measuring broad domains), they often do not.

Gallup recently conducted a test-retest study consisting of 2,219 members of the Gallup Panel, a nationally representative, probability-based panel of U.S. households that have agreed to participate in Gallup Panel surveys by phone, Web, or mail on any topic at any time.

Respondents were recruited to complete the CSF assessment in February of 2008. Those who completed the assessment received no feedback or output of any kind regarding their Signature Themes; nor were they informed that they were participating in a study of the CSF. This was done to enable as pure an evaluation of the CSF's test-retest reliabilities as possible. After completing the assessment, respondents were randomly assigned to one of three retest periods: (1) one month (n = 538), (2) three months (n = 390), and (3) six months after their first assessment (n = 376). The results of this study are shown in Table 3.

Table 3: Test-Retest Reliability Estimates

Theme	1-Month	3-Month	6-Month
Achiever	0.66	0.69	0.68
Activator	0.69	0.68	0.64
Adaptability	0.69	0.72	0.66
Analytical	0.77	0.76	0.75
Arranger	0.53	0.50	0.53
Belief	0.66	0.67	0.70
Command	0.61	0.58	0.64
Communication	0.69	0.69	0.70
Competition	0.73	0.71	0.67
Connectedness	0.70	0.71	0.71
Consistency	0.62	0.65	0.65
Context	0.70	0.71	0.69
Deliberation	0.79	0.79	0.80
Developer	0.67	0.63	0.54
Discipline	0.81	0.82	0.76
Empathy	0.73	0.71	0.65
Focus	0.72	0.73	0.60
Futuristic	0.66	0.67	0.61
Harmony	0.64	0.61	0.61
Ideation	0.75	0.73	0.71
Includer	0.69	0.67	0.67
Individualization	0.60	0.61	0.58
Input	0.79	0.78	0.75
Intellection	0.80	0.78	0.76
Learner	0.80	0.80	0.78
Maximizer	0.61	0.63	0.48
Positivity	0.76	0.77	0.69
Relator	0.63	0.63	0.67
Responsibility	0.69	0.72	0.65
Restorative	0.63	0.65	0.51
Self-Assurance	0.73	0.74	0.72
Significance	0.72	0.71	0.65
Strategic	0.68	0.67	0.71
Woo	0.80	0.82	0.76

As indicated in Table 3, test-retest correlations were generally consistent over the varying time intervals. Only a handful of themes showed notable changes over the longest retest period.

Recruiting from the Gallup Panel provided the opportunity to test the effects of a great number of covariates on CSF responses. Very few of these covariates were found to have any differential impact on the test-retest reliabilities:

- Significance means dropped slightly more among women in the 3-month and 6-month retests.
- Analytical means dropped slightly more among women in the 3-month retest, whereas men's scores increased slightly more in the 6-month retest.
- Individualization mean changes varied by education level in both the 3-month and 6-month retests, but because these changes were not directionally consistent, the effects of education on reliabilities appear artifactual.

Given that the "scores" presented to respondents are rank-ordered themes, the reliability of the score profile is also a critical issue. A Chi-Square test of independence was conducted on each theme, with the dichotomous variables labeled as "theme in top five during pretest" and "theme in top five during posttest." Of the 34 themes, 33 had significant Chi-Square results, indicating that their presence in the top five on the pretest was significantly related to their presence in the top five in the posttest. This finding provides evidence for the stability of the vast majority of the themes from the pretest to the posttest. However, one theme's posttest was independent of its pretest, meaning it was less stable over time in this sample. That theme was Self-Assurance, the rarest theme in the sample. It must be noted that, for most respondents, any new Signature Themes in the posttest had been in the respondent's top ten themes on the pretest, indicating that some of the apparent lack of temporal consistency is an artifact of how the results are reported.

Validity

From a validity standpoint, the CSF looks very strong. That is, it seems to measure what it is supposed to measure. Studies have produced evidence of congruence with the Big Five (Harter & Hodges, 2003), 16PF (Schreiner, 2006), and CPITM (Schreiner, 2006). Gallup researchers have also produced strong recent evidence of construct validity from large confirmatory studies looking at how the items "cluster." These will be examined in turn.

Content Validity

An assessment should be inclusive of all aspects of the domain it is measuring. It is difficult to provide content validity evidence for personality-type assessments. Don Clifton and other Gallup researchers spent more than 30 years studying the traits that led to optimal functioning in a broad array of areas — including schools, and numerous and varied work environments — and across a wide expanse of time. The assessments that were developed as part of this research have been used to select or develop well over 2 million individuals, giving Gallup researchers confidence in the content coverage of the CSF items and themes. Gallup continues to investigate this issue and welcomes any discussions about how to improve the content validity of any of the CSF themes, or the overall instrument.

Construct Validity

The paired-statement design of the CSF limits the methods that can be used to show construct validity. Some statements are linked to multiple themes, and when these statements are chosen, the respondent's score is counted multiple times, once for each theme. When statements within an item are treated as two different items, this builds a direct correlation between these "different" items that systematically biases inter-theme correlations.

Also, for those item pairs for which both statements are attached to themes, the selection of one statement assigns points to the themes aligned with both statements. This last type of statement pair, where endorsing one statement also means a negative score for the other statement, produces some of the properties of ipsativity in the data set.

What is "ipsativity," and what are its analytical ramifications? Kaplan and Saccuzzo (1982) provide a simple definition: Ipsative scores compare the individual against himself or herself and produce data that reflect the relative strength of each need for that person; each person thus provides his or her own frame of reference. One of the classical signs of ipsativity would be equal means and standard deviations in the themes. Because fewer than 30% of the items are ipsatively scored, we know that the instrument has limited ipsativity. Nevertheless, an examination of theme means and standard deviations was made to judge the amount of ipsativity present. This showed ipsativity not to be a problem in the interpretation of the overall instrument. For the purposes of this study, the primary ramifications of having some ipsatively scored items are that these built-in item dependencies limit the types of confirmatory analyses that can be performed. In particular, because knowing the scores of some items defines the scores on other items, the data matrix may be singular and incapable of being inverted.

Because some of the items are used in multiple themes, there is also the potential risk of multicollinearity in these data. Deeper investigations into this have found that multicollinearity is not a problem for the CSF instrument (see Plake, 1999), and the recent revisions to the instrument have reduced this further. But the multiple use of some items does mean that a traditional confirmatory factor analysis (CFA) is problematic. So to represent the internal structure of the CSF, and to show generalizability of the theme taxonomy, a different approach was taken.

The approach chosen was to look at themes in pairs, by performing a hierarchical cluster analysis using the items from two themes at a time, and repeating this process for all theme pairs in which the items are independent. This provided a good representation of how well the statements of a given theme cluster. This approach is similar to factor analysis, although it differs in the way variables are grouped. The between-groups linkage method measured with Pearson's correlation was employed because it uses information from all pairs of distances, not just the farthest or the nearest. The nearer to the origin the cluster combines, the stronger is the correlation between the statements. Sample dendrograms from these analyses are shown in Appendix B.

The results of the most recent series of cluster analyses are shown in Table 4. These results are from a sample of 472,850 respondents, all from 2006. Each cell represents the mean percentage of items in each theme that clustered together. For example: In the Achiever/Activator cell, 100% of the statements for each theme were clustered with the other statements that are linked to that theme. A score of 100% means that the cluster analysis perfectly replicated the statement combinations used in scoring their respective themes. For themes that share items, the shared items were removed prior to the analysis. Clearly, the shared items are already known to be associated with each theme, and the analysis was meant to show the results for all other items.

There is no standard criterion for determining what proportion of items measuring a theme or content area should be grouped together for the theme to be considered “validated.” Clearly, if all items in a theme are clustered and no items from other themes are in that same cluster, the results support the theory that the items are strongly associated enough to warrant a common designation (i.e., theme).

It is unrealistic to expect such perfect results across the entire instrument. In the content validity literature, where subject matter experts are used to group test items into content categories, a rule of thumb has been proposed (by Popham, 1992, and supported by Sireci, 1998): If 70% of the experts classify an item into its hypothesized category, the item should be considered matched to that category. O’Neil, Sireci, and Huff (2004), extended that criterion to content areas by considering an area congruent with its test specifications if at least 70% of its items were appropriately matched. For this analysis, themes were evaluated by determining the proportions of items that clustered, and comparing the results to this 70% criterion. Themes were considered validated if 70% of the items clustered in the two-cluster solution.

Applying this criterion to Table 4, the themes look to be quite distinct as a group. The vast majority of cells show a proportion much higher than the 70% criterion, but there are also a handful of theme combinations that fall below it. For example, Discipline and Consistency show less separation, with only 64% of the items clustering together. Given their conceptual similarity, this makes sense.

Table 4, taken as a whole, is convincing evidence of the validity of the CSF theme structure, with less than 2% of the theme pairs failing to meet the 70% criterion. It should be noted that this method has been replicated in multiple independent samples of CSF respondents, including one composed entirely of college students (Schreiner, 2006). The overall results are very positive, with the cluster analyses supporting the viability of the 34 themes.

In addition to the summary presented in Table 4, Appendix B presents sample dendrograms from the analysis. The vertical lines indicate the relative distance at which two clusters are combined. The two-cluster solution can be found by locating the highest horizontal line and seeing the two groups of items it combines. In some cases, all items within a theme clustered with one or two items from another theme. However, in general, few items from different themes clustered together, and no cross-theme clusters emerged in any of the 561 separate analyses of theme pairs.

This cluster approach circumvents the problem of the dependencies involved in items that measure more than one theme. It also, more appropriately, models the CSF theory, as there is no explicit structure of the CSF beyond the 34 distinct themes.

In addition to supporting the presence of all of the 34 CSF themes, this type of analysis can be used to evaluate all of the themes individually. For example, clusters of items within a theme could indicate subtleties of employees’ talents that have not yet been considered, or to identify subsets of items that need tweaking to become more congruent with the other items in the theme. This hierarchical approach was therefore one of the main methods used to reconfigure the CSF instrument into its current 177-item version.

Criterion-Related Validity

A construct validity study (Harter & Hodges, 2003) explored the relationship between the CSF and the five-factor model of personality in a sample of 297 undergraduate business students in a Midwestern university. The “Big Five” factors of personality are Neuroticism (which reflects emotional stability — reverse-scored), extroversion (seeking the company of others), Openness/Intellectence (interest in new experiences, ideas, and so forth), Agreeableness (likeability, harmoniousness), and Conscientiousness (rule abidance, discipline, integrity) (McCrae & Costa, 1987; McCrae, Costa, Lima, et al., 1999; McCrae, Costa, Ostendorf, et al., 2000). A priori hypotheses linking themes and personality variables included Conscientiousness correlating positively with Achiever, Deliberative, Discipline, Focus, and Responsibility; Extroversion correlating positively with Activator, Communication, and Woo; Agreeableness correlating positively with Harmony and Positivity; and Intellectence correlating positively with Ideation, Input, Intellection, and Strategic. Several of these expected associations between CSF themes and five-factor model constructs were found. For example, the Discipline theme correlates 0.81 with the measure of Conscientiousness. Theoretically, these constructs have similar definition in relation to orderliness and planning. Other examples include the 0.83 correlation between Woo and Extroversion, the 0.70 correlation between Ideation and Intellectence, and the 0.58 correlation between Positivity and Agreeableness.

Schreiner conducted an independent study of construct validity among 438 college students (Schreiner, 2006). In this study as well, the vast majority of a priori hypotheses were confirmed when correlating CSF themes with their expected counterparts on other well-validated personality instruments.

Utility

Successful strengths-based development results in desired behavioral change (Clifton & Harter, 2003; Hodges & Clifton, 2004). Indeed, Gallup (Black, 2001; Connelly, 2002; Krueger, 2004) reports that client-sponsored studies have provided evidence that strengths-based development relates to various positive outcomes, including increases in employee engagement and productivity. Furthermore, managers who create environments in which employees are able to make the most of their talents have more productive work units with less employee turnover (Clifton & Harter, 2003). Studies also show that strengths-based development increases self-confidence, direction, hope, and altruism (Hodges & Clifton, 2004) in college students.

Ongoing research continues to explore the benefits of strengths-based development on desired outcomes in both work and academic settings. In a recent study of the gains made by individuals and work units within Gallup clients, significant gains were observed by those individuals or teams that invested in their own strengths development. Specifically:

- Evidence was accumulated across client data to estimate the average performance increase experienced by them as a result of applying strengths-based management practices.
- Eleven companies were included, representing an estimated 90,000 employees across 900 business units, in 5 different industries. None of the performance measures were available across the entire population, but adequate data existed in multiple sub-populations to indicate significant gains in employee engagement, productivity, profit, and employee retention.

- Most of the individuals in the study were sited in North America, but at least one of the studied companies has a sizable international workforce scattered across Europe, Asia, and South America.

Employee Engagement

In 896 business units, pre-post measures of employee engagement were available in the form of survey data from Gallup's Q¹². (For information on the Q¹² see Wagner & Harter, 2006). The core Q¹² survey consists of 12 Likert items rated on a scale of 1-5. In interpreting the amount of growth on the Q¹² GrandMean (calculated as the mean of the responses to the 12 statements) to consider substantial growth, Gallup researchers have considered a number of different criteria, including various sources of possible error (sampling, measurement, transient), and the relationship of changes in engagement to changes in business outcomes. Considering all of this information, Gallup researchers have adopted, as a general guideline, using 0.20 as criteria for business unit growth, or a 0.10 improvement for larger groups with over 1,000 employees.

Among the 896 business units with Q¹² data, those whose managers received a strengths intervention (generally involving some personalized feedback, but not universally) showed 0.16 more improvement on their Q¹² GrandMean relative to those units where the manager received nothing. This was a simple wait-list control rather than a placebo-controlled study, but given the size of these workgroups (less than $n = 1,000$ but generally larger than $n = 100$), this indicates some evidence of significant increase in engagement from the strengths intervention. This is particularly notable because only the managers of these groups received strengths feedback during the study period — the other 100+ employees in both the study and control groups received nothing.

Data on individual engagement responses were also available for 12,157 employees. Among those employees receiving a strengths intervention, engagement improved by 0.33 relative to employees without the intervention. This was also largely a simple wait-list control, where most of the "control" employees in this study subsequently received strengths feedback and coaching as well. Nevertheless, the substantial gains in employee engagement among the employees receiving strengths feedback are a very positive indication of the utility of the intervention.

Employee Turnover

Turnover data were available for 65,672 employees. Among employees receiving some strengths feedback, turnover rates were 14.9% lower than for those employees receiving nothing (controlling for job type and tenure). Presumably, some of this gain in utility flows through the improvement in engagement discussed previously, given the large body of evidence linking employee engagement to employee turnover. Gallup researchers intend to explore the structure of this multivariate relationship among strengths, engagement, and performance as data become available.

Productivity

There were 530 business units with productivity data. Those whose managers received strengths feedback showed 12.5% greater productivity post-intervention relative to those units where the manager received nothing. Similar to the engagement data discussed above, this is particularly notable because only the managers of these groups received strengths feedback during the study period, with the remainder of the employees in both

the study and control groups receiving nothing in most cases. Also similar to the engagement studies, the “control” managers here were wait-list controls.

Data on the productivity of 1,874 individual employees were examined for the effects of strengths feedback as well. Most of these employees were engaged in sales functions, where the productivity data represent sales. Among those employees receiving a strengths intervention, productivity improved by 7.8% relative to employees without the intervention. This was also largely a simple wait-list control, where many of the “control” employees in this study subsequently received strengths feedback and coaching as well. Nevertheless, the substantial gains in productivity among the employees receiving strengths feedback are a very positive indication of the utility of the intervention. There is also thought to be a significant amount of range restriction in the measurable talents of many of these individuals, as a large percentage of them were selected for their current position via a strengths-based selection instrument. That is, participants were required to possess at a minimum the required levels of the talents measured by these selection instruments to be eligible for the strengths intervention in the first place.

Profitability

Profit data were available for 469 business units, ranging from retail stores to large manufacturing facilities. Those units whose managers received strengths feedback showed 8.9% greater profitability post-intervention relative to units where the manager received nothing. Again, this is extremely positive evidence of the utility of investing in talent; only the managers of these groups received strengths feedback during the study period, with the remainder of the employees in both the study and control groups receiving nothing in most cases. Also similar to the engagement studies, the “control” managers here were wait-list controls for the most part.

Closing Comments

Since 1998, the CSF has been used as Gallup’s talent identification tool in development programs with various academic institutions, faith-based organizations, major businesses, and other organizations. As mentioned previously, Gallup researchers continue to examine the psychometric properties of the instrument and modify it based on research findings.

The CSF has been used to facilitate the development of individuals across dozens of roles including: executive, student, teacher, manager, customer service representative, salesperson, administrative assistant, nurse, lawyer, pastor, leader, and school administrator. Strengths-based development programs, grounded in traditional Gallup practices, are now being refined based on the principles of Positive Psychology, the scientific study of and evidence-based promotion of optimal human functioning (as summarized in Cameron, Dutton, & Quinn, 2003; Keyes & Haidt, 2003; Linley & Joseph, 2004; Lopez & Snyder, 2003; Snyder & Lopez, 2002).

The preponderance of the validity evidence to-date shows strong evidence of the utility of these strengths-based development programs, with large identified gains in performance among those studied. Gallup continues to evaluate these relationships as data become available from clients or research partners.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (AERA/APA/NCME). (1999). *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Association.
- Asplund, James W. (2009). Strengths. In S. Lopez (Ed.) *Encyclopedia of Positive Psychology*, Volume 2. Wiley-Blackwell (Oxford, UK and Malden, MA), 2009.
- Black, B. (2001). The road to recovery. *Gallup Management Journal*, 1(4), 10-12.
- Buckingham, M., & Clifton, D. O. (2000). Now, discover your strengths. New York: Free Press.
- Cameron, K. S., Dutton, J. E., & Quinn, R. E. (Eds.). (2003). *Positive organizational scholarship*. San Francisco: Berrett-Koehler.
- Clifton, D. O. (1970, March). *The magnificence of management*. A reprint of an address presented to the 8th Annual Life Agency Management Program. Boston, Mass.
- Clifton, D. O. (1975). Interaction is: Where the action is. A reprint of a report prepared by Donald O. Clifton and presented at the 1972 Chartered Life Underwriters (CLU) Forum.
- Clifton, D. O. (1980). *Varsity Management: A way to increase productivity*. A reprint of an address presented to the 29th Annual Consumer Credit Insurance Association (CCIA) Program on June 24, 1980. Napa, California.
- Clifton, D. O., & Anderson, E. (2002). *StrengthsQuest: Discover and develop your strengths in academics, career, and beyond*. New York: Gallup Press.
- Clifton, D. O., & Harter, J. K. (2003). Strengths investment. In K. S. Cameron, J. E. Dutton, & R. E. Quinn (Eds.), *Positive organizational scholarship*. (pp. 111-121). San Francisco: Berrett-Koehler.
- Clifton, D. O., Hollingsworth, F. L., & Hall, W. E. (1952). A projective technique for measuring positive and negative attitudes towards people in a real-life situation. *Journal of Educational Psychology*, 43.
- Clifton, D. O., & Nelson, P. (1992). *Soar with your strengths*. New York: Delacorte Press.
- Connelly, J. (2002). All together now. *Gallup Management Journal*, 2(1), 13-18.
- Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- Dodge, G. W., & Clifton, D. O. (1956). Teacher-pupil rapport and student teacher characteristics, *Journal of Educational Psychology*, 47, 6.
- Gallup, G. (2004). *The Gallup Poll: Public opinion 2003*. Lanham, MD. Roman and Littlefield.

- Harter, J. K., Hayes, T. L., & Schmidt, F. L. (2004). *Meta-analytic predictive validity of Gallup Selection Research Instruments* [technical report]. Omaha, NE: Gallup.
- Harter, J. K., & Hodges, T. D. (2003) *Construct validity study: StrengthsFinder and the Five Factor Model* [technical report]. Omaha, NE: Gallup.
- Hodges, T. D., & Asplund, Jim (2009). Strengths Development in the Workplace. In A. Linley, P.A., Harrington, S., & Page, N. (Eds.) *Oxford handbook of positive psychology and work*. New York: Oxford University Press.
- Hodges, T. D., & Clifton, D. O. (2004). Strengths-based development in practice. In A. Linley & S. Joseph (Eds.), *Handbook of positive psychology in practice*. Hoboken, New Jersey: John Wiley and Sons, Inc.
- Kane, M.T. (1992a). An argument-based approach to validity. *Psychological Bulletin*, 112,527-535.
- Kaplan, R. M., & Saccuzzo, D. P. (1982). *Psychological testing*. Monterey, CA: Brooks/Cole.
- Keyes, C. L. M., & Haidt, J. (Eds.). (2003). *Flourishing: Positive psychology and the life well-lived*. Washington, DC: APA.
- Krueger, J. (2004, November). How Marriott Vacation Club International engages talent. *Gallup Management Journal*, 4.
- Linley, A., & Joseph, S. (Eds.). (2004). *Positive psychology in practice*. Hoboken, NJ: John Wiley & Sons, Inc.
- Lopez, S. (2009). *The Encyclopedia of Positive Psychology*. Malden, MA: Wiley-Blackwell.
- Lopez, S., Hodges, T., & Harter, J. (2005, January). *The Clifton StrengthsFinder technical report: Development and validation*. Unpublished report.
- Lopez, S. J., & Snyder, C. R. (Eds.). 2003. *Positive psychological assessment: A handbook of models and measures*. Washington, DC: American Psychological Association.
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81–90.
- McCrae, R. R., Costa, P. T., de Lima, M. P., et al. (1999). Age differences in personality across the adult life span: Parallels in five cultures. *Developmental Psychology*, 35, 466–77.
- McCrae, R. R., Costa, P. T., & Ostendorf, F., et al. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology*, 78, 173–86.
- Newport, F. (2004). Polling matters. New York: Warner Books.

- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum.
- Plake, B. (1999). *An investigation of ipsativity and multicollinearity properties of the StrengthsFinder Instrument* [technical report]. Lincoln, NE: Gallup.
- Schmidt, F. L., & Rader, M. (1999). Exploring the boundary conditions for interview validity: Meta-analytic validity findings for a new interview type. *Personnel Psychology, 52*, 445-464.
- Schreiner, Laurie A. (2006). *A Technical Report on the Clifton StrengthsFinder with College Students*. <https://www.strengthsquest.com/Content/?CI=25195>
- Snyder, C. R., & Lopez, S. J. (Eds.). (2002). *The handbook of positive psychology*. New York: Oxford University Press.
- Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55*(1), 5-14.
- Sireci, S. G. (2001). Standard setting using cluster analysis. In C. J. Cizek (Ed.), *Standard setting: Concepts, methods, and perspectives* (pp. 339-354). Mahwah, NJ: Lawrence Erlbaum.
- Sireci, S. G. (1998a). Gathering and analyzing content validity data. *Educational Assessment, 5*, 299-321.
- Sireci, S. G. (1998b). *The construct of content validity*. *Social Indicators Research, 45*, 83-117.

Appendix A: Brief Descriptions of the 34 Themes of Talent Measured by the Clifton StrengthsFinder

Achiever

People especially talented in the Achiever theme have a great deal of stamina and work hard. They take great satisfaction from being busy and productive.

Activator

People especially talented in the Activator theme can make things happen by turning thoughts into action. They are often impatient.

Adaptability

People especially talented in the Adaptability theme prefer to "go with the flow." They tend to be "now" people who take things as they come and discover the future one day at a time.

Analytical

People especially talented in the Analytical theme search for reasons and causes. They have the ability to think about all the factors that might affect a situation.

Arranger

People especially talented in the Arranger theme can organize, but they also have a flexibility that complements this ability. They like to figure out how all of the pieces and resources can be arranged for maximum productivity.

Belief

People especially talented in the Belief theme have certain core values that are unchanging. Out of these values emerges a defined purpose for their life.

Command

People especially talented in the Command theme have presence. They can take control of a situation and make decisions.

Communication

People especially talented in the Communication theme generally find it easy to put their thoughts into words. They are good conversationalists and presenters.

Competition

People especially talented in the Competition theme measure their progress against the performance of others. They strive to win first place and revel in contests.

Connectedness

People especially talented in the Connectedness theme have faith in the links between all things. They believe there are few coincidences and that almost every event has a reason.

Consistency

People especially talented in the Consistency theme are keenly aware of the need to treat people the same. They try to treat everyone in the world with consistency by setting up clear rules and adhering to them.

Context

People especially talented in the Context theme enjoy thinking about the past. They understand the present by researching its history.

Deliberative

People especially talented in the Deliberative theme are best described by the serious care they take in making decisions or choices. They anticipate the obstacles.

Developer

People especially talented in the Developer theme recognize and cultivate the potential in others. They spot the signs of each small improvement and derive satisfaction from these improvements.

Discipline

People especially talented in the Discipline theme enjoy routine and structure. Their world is best described by the order they create.

Empathy

People especially talented in the Empathy theme can sense the feelings of other people by imagining themselves in others' lives or others' situations.

Focus

People especially talented in the Focus theme can take a direction, follow through, and make the corrections necessary to stay on track. They prioritize, then act.

Futuristic

People especially talented in the Futuristic theme are inspired by the future and what could be. They inspire others with their visions of the future.

Harmony

People especially talented in the Harmony theme look for consensus. They don't enjoy conflict; rather, they seek areas of agreement.

Ideation

People especially talented in the Ideation theme are fascinated by ideas. They are able to find connections between seemingly disparate phenomena.

Includer

People especially talented in the Includer theme are accepting of others. They show awareness of those who feel left out, and make an effort to include them.

Individualization

People especially talented in the Individualization theme are intrigued with the unique qualities of each person. They have a gift for figuring out how people who are different can work together productively.

Input

People especially talented in the Input theme have a craving to know more. Often they like to collect and archive all kinds of information.

Intellection

People especially talented in the Intellection theme are characterized by their intellectual activity. They are introspective and appreciate intellectual discussions.

Learner

People especially talented in the Learner theme have a great desire to learn and want to continuously improve. In particular, the process of learning, rather than the outcome, excites them.

Maximizer

People especially talented in the Maximizer theme focus on strengths as a way to stimulate personal and group excellence. They seek to transform something strong into something superb.

Positivity

People especially talented in the Positivity theme have an enthusiasm that is contagious. They are upbeat and can get others excited about what they are going to do.

Relator

People especially talented in the Relator theme enjoy close relationships with others. They find deep satisfaction in working hard with friends to achieve a goal.

Responsibility

People especially talented in the Responsibility theme take psychological ownership of what they say they will do. They are committed to stable values such as honesty and loyalty.

Restorative

People especially talented in the Restorative theme are adept at dealing with problems. They are good at figuring out what is wrong and resolving it.

Self-Assurance

People especially talented in the Self-Assurance theme feel confident in their ability to manage their own lives. They possess an inner compass that gives them confidence that their decisions are right.

Significance

People especially talented in the Significance theme want to be very important in the eyes of others. They are independent and want to be recognized.

Strategic

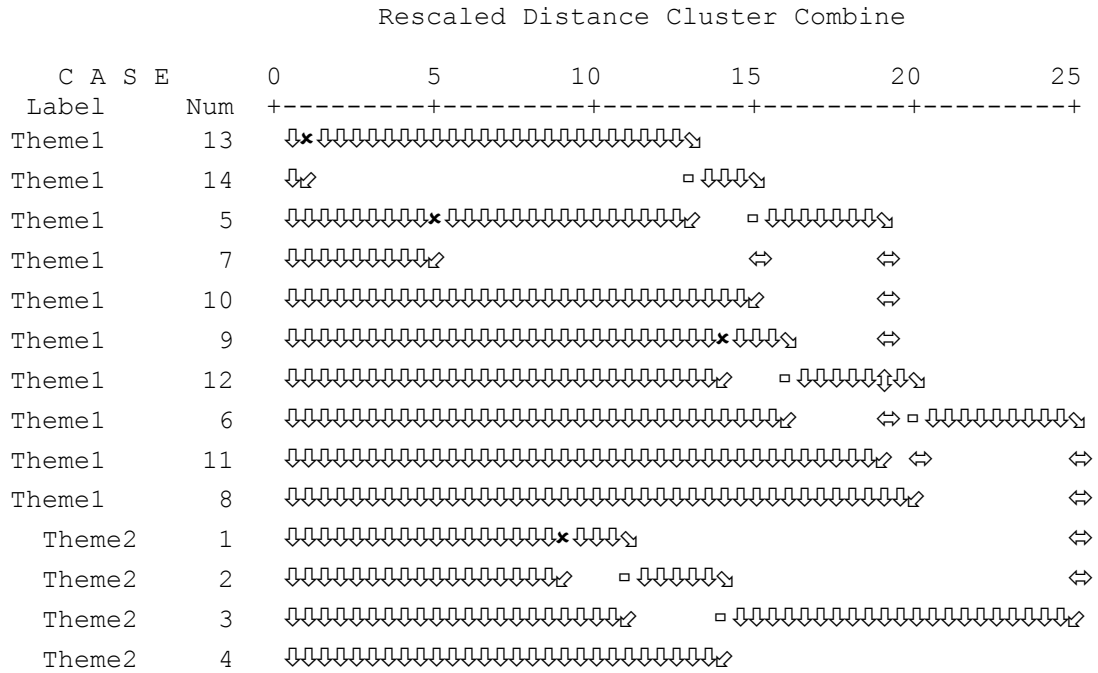
People especially talented in the Strategic theme create alternative ways to proceed. Faced with any given scenario, they can quickly spot the relevant patterns and issues.

Woo

People especially talented in the Woo theme love the challenge of meeting new people and winning them over. They derive satisfaction from breaking the ice and making a connection with another person.

Appendix B: Example Dendrograms From Hierarchical Cluster Analysis

This dendrogram shows 100% clustering of items in the correct themes.



This dendrogram shows an imperfect clustering of items by theme.

