



PANORAMA
EDUCATION

Reliability and Validity of The Panorama Well-Being Survey

Students' social and emotional well-being is an essential outcome for educators, families, and students themselves. This report gives in-depth information on the development of the [Panorama Well-Being Survey](#), including evidence of its reliability and validity.

Table of Contents

- Background**.....3
 - Well-being3
 - Emotions.....3
 - Social Support.....4
- Survey Development**5
 - Content Creation.....5
 - Pilot Study6
 - Scoring.....6
 - Scale Construction.....7
- Reliability**11
- Validity**11
 - Structural Validity12
 - Convergent and Discriminant Validity12
- Conclusion**.....15
- References**.....16

Background

The Panorama Well-Being Survey is composed of three scales, or groups of survey questions, each focused on a single construct, or topic: Positive Feelings, Challenging Feelings, and Supportive Relationships. The survey is based on the core social and emotional needs of students (see Gehlbach & Chuter, 2020) and designed for use in elementary and secondary school settings.

Although it is beyond the scope of this report to comprehensively review all the literature that informed our development of these survey scales, we hope that readers will benefit from brief, selective reviews of the key topics below.

Well-being

Based on the “disease model” of 20th century clinical psychology, well-being used to be thought of as the absence of risky behavior or mental illness. In the past two decades, however, the field of positive psychology has helped redefine well-being as the positive experiences, thoughts, and feelings that enable human prosperity and flourishing (Seligman & Csikszentmihalyi, 2000; VanderWeele, 2017). At Panorama, we also approach student well-being from this strength-based perspective, with the goal of helping educators build student happiness instead of diagnosing or forecasting student illness.

Although there are important objective components to student well-being—such as sleep, exercise, and stable housing (for review, see Pollard & Lee, 2003)—our approach to well-being centers on students’ subjective experiences. How are students interpreting their reality and experiencing their lives? Are their days mostly filled with moments of loneliness and anxiety, or with happiness and excitement? Do they feel safe and supported by a network of trusted friends, family, and adults at school? We emphasize students’ subjective experiences because they matter as ends unto themselves; are important mediators of objective health, social, and economic outcomes (for review, see Lyubomirsky et al., 2005); and, ultimately, because they are valued by educators, students, and families.

Emotions

Not surprisingly, emotions feature prominently in the literature on well-being (e.g., Ryan & Deci, 2001; Pollard & Lee, 2003; VanderWeele, 2017). In fact, some theorists and researchers define happiness as the relative proportion of positive versus negative affect that individuals experience over time (Kahneman et al., 1999; Keyes et al., 2002). In measuring emotional well-being, two issues persistently surface in the literature: 1) whether positive and negative emotion are separate constructs that merit separate measurement, and 2) whether one should measure the frequency or intensity of emotions.

For various reasons—including the fact that individuals who report more happiness tend to report less sadness, and that people smile when happy but frown when sad—it is reasonable to consider negative and positive emotions as two poles of a single, underlying dimension of human experience (see Larsen et al., 2001; Russell & Carroll 1999). But scholarship has shown that they are separate in important respects, including how they ebb and flow differently within individuals across time (Zevon & Tellegen, 1982), how they rely on different brain systems (Cacioppo et al., 1999), and how they relate differently to other variables (Bradburn & Caplovitz, 1965). For example, whereas anxiety disorders are characterized by negative emotions and (a lack of) positive emotions, clinical depression is characterized by only a lack of positive emotions (and not the presence of negative emotions; Watson et al., 1988). In the context of education, researchers have found that they can better predict student academic engagement from measures of both positive and negative affect than from only measures of negative affect (Lewis et al., 2009). These and similar findings led to the advent of positive psychology and its central tenet that the absence of human suffering is not the presence of human flourishing (Diener, 2000). Based on this evidence, we recommend that educators separately consider students’ experiences of positive and negative affect, and not simply view the presence of one as the absence of the other.

The subjective experience of emotions can be characterized not just in terms of negativity or positivity, but also frequency and intensity. Children, adolescents, and adults can experience infrequent moments of intense joy, as

well as frequent periods of calm contentment. When measuring well-being, we must decide whether to focus on emotional intensity, frequency, or both. It turns out that, somewhat counterintuitively, frequency matters more for well-being than intensity. How often individuals experience emotions is a better predictor of their happiness than how intensely they feel emotions (Diener, 2000; Diener et al., 1991). In the context of education, emotion frequency better predicts student outcomes, especially for positive emotions. For example, Hernández et al. (2015) found that the quality of kindergarteners' relationships with their peers was predicted by how frequently but not intensely they expressed positive emotion. Thus, although there may be benefits to measuring both emotional frequency and intensity (e.g., Kim et al., 2007), we decided to focus on frequency because of the costs (in time and interpretability) of measuring both and the research evidence that frequency matters more than intensity.

Social Support

Decades of scholarly research across a diversity of theoretical and applied fields have shown that humans of all ages benefit from meaningful connections with one another. When we are loved, valued, and supported by others, we feel happier (Ryan & Deci, 2001), manage stress better (Sapolsky, 2004), and live longer (House et al., 1988). Scholars commonly distinguish between three types of social support, all of which manifest in school settings: informational support (e.g., advice), instrumental support (e.g., services), and emotional support (e.g., reassurance; Taylor, 2011). The positive impacts of social support depend, in part, on matching the type of support given to the type of support needed (Thoits, 1995). Importantly, however, social support is a fundamentally subjective experience in which the perception of support matters just as much, and sometimes more, than any actual support received (e.g., Wethington & Kessler, 1986). In other words, merely knowing that one has support when needed is itself a form of support. Finally, a few close confidants offer as efficacious and, in some cases, more efficacious support than a larger social network (Taylor, 2011).

Students receive complementary forms of social support from three key sources: family members, peers, and teachers (Cauce & Srebnik, 1990; Dubow & Ullman, 1989; Wentzel, 1998). These sources of support benefit students in terms of their well-being (Chu et al., 2010), academic motivation (Wentzel, 1998), and academic achievement (Malecki & Demaray, 2006). Using a nationally representative sample of students, for example, Rosenfeld et al. (2000) found that students who felt supported by their families, peers, and teachers had fewer school absences, spent more time studying, were more engaged in school, and obtained better grades than other students. Evidence suggests that social support from teachers and other school personnel is particularly important, especially for underserved students (Becker & Luthar, 2002). These and related findings not only underscore the importance of examining supportive relationships when measuring well-being, but also suggest looking across different relationships and the support they provide.

Survey Development

To develop the survey, we interviewed district leaders and educators from across the country, reviewed relevant scholarship and instrumentation, and followed best practices in the science of survey design. After designing an initial survey, we pilot-tested it in 100 secondary schools, collecting initial data from over 25,000 students. Using these results and feedback from our pilot partners, we refined the survey to maximize its psychometric qualities and practical utility (see Fowler, 2013; Gehlbach & Artino, 2018; Gehlbach & Brinkworth, 2011). Our goal with these efforts was to create a reliable, valid, and practical instrument that educators can use at scale in schools to measure and support students' well-being, not to design a clinical instrument for diagnosing mental illness or predicting suicidality.

Content Creation

We created each survey item to adhere to best practices from the science of survey design (Gehlbach & Artino, 2018; Dillman et al., 2014; Fowler, 2013). For example, researchers have concluded that designing survey items as statements, particularly ones that require respondents to agree or disagree, is likely to inject additional measurement error and bias into responses. Numerous surveys used by educators unfortunately fail to adhere to these well-established survey design practices. For example, asking questions with response options that are linked to the underlying concept is the preferred practice (Dillman et al., 2014; Krosnick, 1999a; Saris et al., 2010)—e.g., rather than posing statements that respondents are supposed to agree or disagree with. Failing to label all response options, using numeric rather than verbal labels, and using too few response options are other commonly violated best practices (Artino et al., 2014; Dillman et al., 2014; Krosnick, 1999b; Weng, 2004). As a survey scale assessing a particular topic violates more of these best practices, the amount of measurement error and bias grows. The items that comprise Panorama’s survey instruments, including our Well-Being Survey, adhere to these best practices.

In creating content for this well-being instrument, several special considerations also impacted our item design to maximize the usefulness of these measures for schools. First, we wanted the survey to be especially short so that it could be administered frequently or alongside other survey content. Second, given the impacts of the coronavirus pandemic on schooling, we designed the survey questions so that they could accommodate remote, hybrid, and in-person learning environments.

Third, to ensure that the questions were maximally accessible to younger children or students with limited language proficiency, we relied on established corpora of child-friendly or high-frequency emotion words. For example, we used the word frequency lists developed by Hiebert (2005) to discriminate between competing emotion words (e.g., “worried” instead “anxious”) based on their developmental appropriateness. To optimize accessibility and comprehension, we also relied on understandability ratings collected from a convenience sample of parents, former educators, and survey experts at Panorama.

Finally, we aimed to make our survey items sensitive to change while allowing them to capture more than fleeting moods. Following the norm of Bradburn and Caplovitz (1965) and based on research into the Positive and Negative Affect Schedule scales (see Watson et al., 1988; Watson & Clark, 1999), we instructed students to reflect on their emotional experiences over the past week (as opposed to the past day or month).

We created alternative forms of each survey scale, one for younger students (Grades 3-5) and the other for older students (Grades 6-12). In drafting and revising survey content for age groups, we strived to create developmentally appropriate questions while, at the same time, minimizing content differences between the two survey forms.

Pilot Study

In the spring of 2020, we collected survey responses from 26,656 students enrolled in 100 public middle and high schools from 24 school districts in the Northeast, Midwest, South, and West. The majority of students (92%) completed all survey items, and 99% of students skipped four or fewer items. We excluded data from the 62 students (0.2%) who did not complete a majority of survey items.

For the vast majority of instances, we used administrative data to determine student demographics; when unavailable, we used available self-report data. Schools ranged in their demographics from 40-100% female, 0-95% White, 0-100% English-language learners (ELL), and 3-100% eligible for free or reduced-price lunch (FRPL). Overall, the sample was 60% female, 11% ELL, 24% FRPL-eligible, 42% White, 27% Latinx, 11% Asian, and 16% Black.

To minimize impression management and self-presentation biases, we included the following survey instructions:

These questions ask about how you've been feeling recently. Please respond honestly—there are no right or wrong answers because there are no right or wrong feelings! Your answers will help us better support you and other students, and will not affect your grades or show up on your report card. You can skip any question you don't feel comfortable answering.

The timing of this pilot study coincided with the onset of the coronavirus pandemic. Despite the likely impact of school closures, social distancing, and the pandemic on student well-being, we do not expect this timing to affect how items interrelate or group together psychometrically into measurement scales.

Scoring

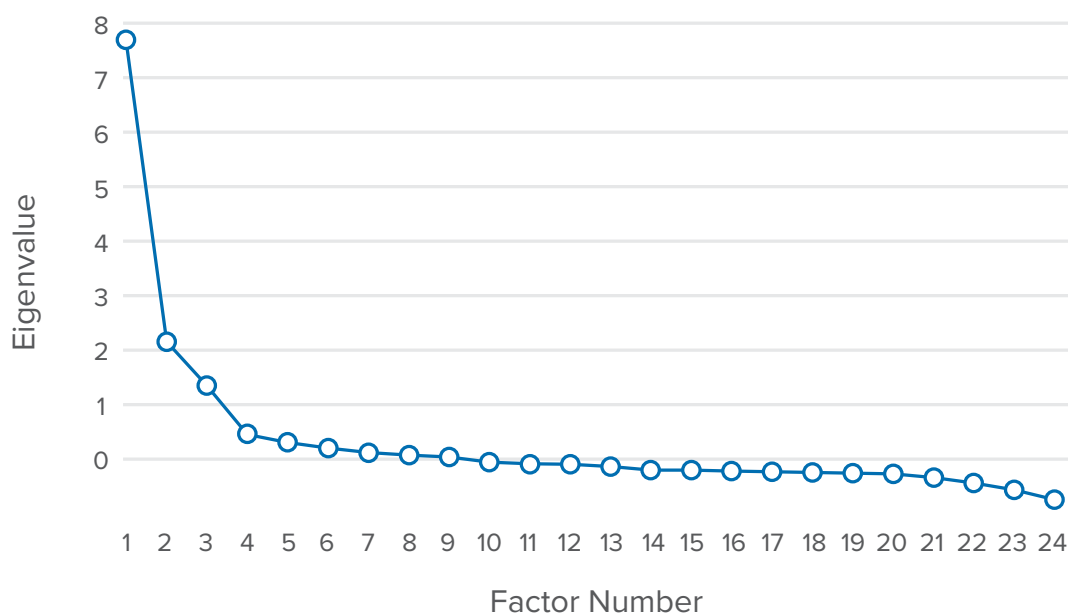
Most piloted survey items had five verbally-labeled response options, with some having only two options (“yes” or “no”). To score items for analytic purposes, we converted each response option to an integer value (starting at “1”) with higher integers always representing more positive responses. For example, students who reported “almost always” feeling happy or “almost never” feeling sad were assigned a numerical score of “5,” and those who reported “almost never” feeling happy or “almost always” feeling sad were assigned a score of “1.” To create topic scores, we simply averaged across item scores.

Scale Construction

We conducted exploratory factor analyses on one randomly-selected half of the data (stratified by school), and reserved the other half for confirmatory factor analyses. To assess appropriateness of the data for factor analysis, we conducted Bartlett's test of sphericity and calculated the Kaiser-Meyer-Olkin (*KMO*) measure of sampling adequacy. These analyses provided evidence of underlying latent factors and that the data are suited to factor analysis (Bartlett's $p < .0001$; *KMO* = 0.89).

Based on a priori theoretical considerations, the scree plot (see Figure 1), a parallel analysis that simulates eigenvalues for a randomly constructed dataset (Horn, 1965; see also Ledesma & Valero-Mora, 2007), and the interpretability of results, we decided to extract three factors.

Figure 1. Scree Plot



Note: Negative eigenvalues like those plotted here above can arise with factor analyses based on polychoric or tetrachoric correlations, especially when the data contain multiple binary questions (see Lorenzo-Seva & Ferrando, 2020).

Since most questions were on an ordinal (and not continuous) scale with some binary (yes/no) questions, we based the factor analysis on polychoric correlations (for those involving one or more ordinal questions) or tetrachoric correlations (for those involving two binary questions). This approach has been found to yield more accurate results in factor analysis than one based on Pearson correlations (Holgado-Tello et al., 2010).

As an additional stage of scale construction, we also collected ratings of each question's relevance to its targeted construct. Such feedback is recommended to help build validity into a survey from the outset (Gehlbach & Brinkworth, 2011). A convenience sample of twelve Panorama employees with relevant professional experience (e.g., working with state and district leaders to support child well-being, leading professional development SEL workshops with teachers) rated each question's relevance on a five-level scale, from "Not at all relevant" to "Extremely relevant." Those ratings were converted to a 1-5 integer scale and averaged together to create mean relevance scores for each item.

Loadings on the three factors and average relevance ratings are shown in Table 1.

Table 1. Factor Loadings and Relevance Ratings

Question ID	Question Text	Positive Feelings	Challenging Feelings	Supportive Relationships	Relevance Rating
excited	During the past week, how often did you feel excited?	0.61	0.18	-0.08	3.5
happy	During the past week, how often did you feel happy?	0.61	0.28	-0.03	4.5
loved	During the past week, how often did you feel loved?	0.66	0.09	0.02	4.2
safe	During the past week, how often did you feel safe?	0.49	0.17	0.01	4.6
hopeful	During the past week, how often did you feel hopeful?	0.81	-0.04	-0.11	3.5
helpful	During the past week, how often did you feel helpful?	0.76	-0.04	-0.06	2.8
kind	During the past week, how often did you feel kind?	0.78	-0.10	-0.06	2.6
optimistic	During the past week, how often did you feel optimistic?	0.73	0.08	-0.09	3.9
grateful	During the past week, how often did you feel grateful?	0.73	-0.17	-0.02	3.1
curious	During the past week, how often did you feel curious?	0.55	-0.26	-0.06	3.4
mad	During the past week, how often did you feel mad?	0.14	0.55	-0.05	3.9
lonely	During the past week, how often did you feel lonely?	0.21	0.61	-0.01	4.8
sad	During the past week, how often did you feel sad?	0.16	0.72	-0.04	4.2
worried	During the past week, how often did you feel worried?	-0.18	0.87	0.02	4.1
frustrated	During the past week, how often did you feel frustrated?	0.01	0.76	0.01	3.7
stressed out	During the past week, how often did you feel stressed out?	-0.02	0.80	-0.03	4.5
afraid	During the past week, how often did you feel afraid?	-0.17	0.74	0.08	4.6
bored	During the past week, how often did you feel bored?	0.15	0.38	-0.12	3.0
support teacher	Do you have a teacher or other adult from school who you can count on to help you, no matter what?	0.25	-0.02	0.35	4.5
support family	Do you have a family member or other adult outside of school who you can count on to help you, no matter what?	0.30	0.07	0.39	4.5
support peer	Do you have a friend from school who you can count on to help you, no matter what?	-0.12	-0.03	0.92	4.5
authentic teacher	Do you have a teacher or other adult from school who you can be completely yourself around?	0.17	-0.03	0.52	3.5
authentic family	Do you have a family member or other adult outside of school who you can be completely yourself around?	0.30	0.10	0.42	3.8
authentic peer	Do you have a friend from school who you can be completely yourself around?	-0.22	-0.01	0.98	4.3

As shown in Table 1, the questions about positive affect all loaded most highly on the first factor, the questions about negative affect all loaded most highly on the second factor, and the questions about social support all loaded most highly on the third factor. With this evidence for three conceptually-related topics, we decided on the names Positive Feelings, Challenging Feelings, and Supportive Relationships.

Although the research literature on well-being often refers to positive and negative affect, we favored the topic name “Challenging Feelings” over “Negative Feelings” to encourage accurate and equitable interpretations of student data. Students experiencing high levels of loneliness may be responding appropriately to challenging life circumstances, and students experiencing high levels of anger may be responding productively to systemic racial injustice. In neither case should we label those students as “lonely” or “angry,” their reactions as negative, or assume that they, as opposed to their circumstances, are responsible for their suffering. In other words, negative emotions are often positive reactions to challenging circumstances (e.g., Fischer & Manstead, 2008; Sell et al., 2009). By using “challenging” instead of “negative,” we hope to encourage an asset-based measurement approach that reduces self-presentation biases and stereotyped interpretations of student data (see Chemaly, 2019).

After identifying this tripartite structure for the survey instrument, we removed questions based on their relevance ratings, factor loadings, redundancy, and other issues. Table 2 lists the questions removed and the reasons for their removal.

Table 2. Eliminated Questions

Question ID	Reason(s)
helpful	Low relevance rating (less than 3.5 on 1-5 scale)
kind	Low relevance rating (less than 3.5 on 1-5 scale)
optimistic	Relatively high skip rate, particularly for younger grades
grateful	Low relevance rating (less than 3.5 on 1-5 scale)
curious	Low relevance rating (less than 3.5 on 1-5 scale)
bored	Low relevance rating (less than 3.5 on 1-5 scale)
stressed out	Redundancy with “worried” and lower factor loading [alternate form for Grades 3-5, “nervous,” also removed]
afraid	Redundancy with “worried” and lower factor loading

Table 3 documents the item intercorrelations for the retained questions, showing tetrachoric correlations for two binary questions and polychoric correlations for all other question pairs. With few exceptions, items show higher correlations within topics than between topics, as one would expect from structurally valid scales.

Table 3. Item Intercorrelations

	excited	happy	loved	safe	hopeful	mad	lonely	sad	worried	frustrated	support teacher	support family	support peer	authentic teacher	authentic family	authentic peer
excited		0.72	0.46	0.34	0.49	0.31	0.44	0.43	0.30	0.38	0.25	0.29	0.17	0.23	0.33	0.10
happy	0.72		0.61	0.46	0.52	0.41	0.52	0.55	0.38	0.46	0.25	0.36	0.20	0.27	0.39	0.16
loved	0.46	0.61		0.56	0.47	0.31	0.49	0.44	0.26	0.32	0.29	0.48	0.21	0.28	0.49	0.18
safe	0.34	0.46	0.56		0.41	0.24	0.37	0.38	0.33	0.27	0.23	0.36	0.13	0.24	0.40	0.15
hopeful	0.49	0.52	0.47	0.41		0.29	0.35	0.39	0.21	0.30	0.23	0.29	0.15	0.23	0.35	0.11
mad	0.31	0.41	0.31	0.24	0.29		0.43	0.53	0.38	0.61	0.15	0.20	0.08	0.13	0.23	0.05
lonely	0.44	0.52	0.49	0.37	0.35	0.43		0.71	0.49	0.49	0.19	0.31	0.12	0.16	0.35	0.11
sad	0.43	0.55	0.44	0.38	0.39	0.53	0.71		0.58	0.56	0.16	0.28	0.08	0.14	0.35	0.09
worried	0.30	0.38	0.26	0.33	0.21	0.38	0.49	0.58		0.57	0.10	0.19	0.06	0.11	0.20	0.06
frustrated	0.38	0.46	0.32	0.27	0.30	0.61	0.49	0.56	0.57		0.17	0.25	0.12	0.15	0.25	0.07
support teacher	0.25	0.29	0.29	0.23	0.23	0.15	0.19	0.16	0.10	0.17		0.49	0.39	0.64	0.32	0.27
support family	0.29	0.36	0.48	0.36	0.29	0.20	0.31	0.28	0.49	0.25	0.49		0.44	0.40	0.66	0.35
support peer	0.17	0.20	0.21	0.13	0.15	0.08	0.12	0.08	0.44	0.12	0.39	0.44		0.43	0.35	0.80
authentic teacher	0.23	0.27	0.28	0.24	0.23	0.13	0.16	0.14	0.43	0.15	0.64	0.40	0.43		0.50	0.49
authentic family	0.33	0.39	0.49	0.40	0.35	0.23	0.35	0.35	0.20	0.25	0.32	0.66	0.35	0.50		0.47
authentic peer	0.10	0.16	0.18	0.15	0.11	0.05	0.11	0.09	0.06	0.07	0.27	0.35	0.80	0.49	0.47	

As a final step in scale construction, we verified that the key findings presented above, which were based on survey content and data from secondary schoolers, replicated with the elementary school sample and survey form. To do so, we relied on a separate sample of 9,786 third-, fourth-, and fifth-grade students from 107 schools. Compared to the final survey for older children, the version for younger children was identical except for the exclusion of the “hopeful” and “frustrated” emotion items, and the three social support items that used the phrasing “Do you have a ... who you can be completely yourself around.” As the item-factor loading patterns and item intercorrelations remained essentially similar when using the data from the younger sample, we used the same scale structure in both survey forms (i.e., Grades 3-5 vs. Grades 6-12).

Reliability

Reliability, as assessed through Cronbach’s alpha, is essentially a measure of signal-to-noise (DeVellis, 2016) with higher values reflecting more “signal” and less “noise.” Put differently, Cronbach’s alpha measures how similarly students respond to the different items within the same scale, i.e., a scale’s internal consistency (Streiner, 2003). The Kuder and Richardson Formula 20 (*KR-20*) is used instead of Cronbach’s alpha for scales with binary items, as is the case for the Supportive Relationships topic; similar to Cronbach’s alpha, higher values reflect greater reliability. For Cronbach’s alpha, a value of .70 or higher indicates sufficient reliability among many experts whereas for *KR-20*, the threshold for adequate reliability is .50 (Frey, 2018). As shown in Table 4, all three topics exhibited “good” reliability by exceeding the typical sufficiency thresholds of their respective metrics in both the exploratory and confirmatory samples.

Table 4. Reliability

Topic (<i>statistic</i>)	Exploratory Sample	Confirmatory Sample
Positive Feelings (α)	.79	.80
Challenging Feelings (α)	.82	.82
Supportive Relationships (<i>KR-20</i>)	.65	.65

Validity

To investigate the validity of Panorama’s Well-Being Survey, we examined its structural, convergent, and discriminant validity. Unlike other validation studies for measures of child or adolescent well-being (e.g., Long et al., 2012), we examined data across a large, diverse sample of students, not a single school or district.

Structural Validity

Using the separate confirmatory dataset, we conducted confirmatory factor analyses (CFAs) to assess the fit of the proposed three-factor solution derived from the exploratory dataset and analyses (see Fabrigar et al., 1999). Specifically, we tested how well a three-factor solution fit the data and, as a reference point, tested the alternative of a single-factor solution (i.e., one topic instead of three).

Table 5 presents the two key statistics from each analysis: the comparative fit index (*CFI*) and the root mean square error of approximation (*RMSEA*). With a maximum possible value of 1, *CFI* measures how well the data from each scale fit a one-dimensional solution; historically, a *CFI* of .90 or greater has been considered sufficient, though a value closer .95 or above is preferred. *RMSEA* is a complementary measure of model fit, with lower values indicating better fit and a typical threshold of .08. (For a discussion of these metrics and their conventions, see Hu & Bentler, 1999; Laverdière et al., 2013). As shown in Table 5, the three-factor solution exceeds the established thresholds for both fit statistics, whereas the single-factor solution exhibits poorer, insufficient fit for both statistics.

Table 5. Confirmatory Factor Analyses

Statistic	1-Factor Solution (separate analysis)			1-Factor Solution (combined analysis)	3-Factor Solution (combined analysis)
	Positive Feelings	Challenging Feelings	Supportive Relationships		
<i>RMSEA</i> (90% CI)	.12 (.11-.12)	.14 (.13-.14)	.11 (.10-.11)	.12 (.12-.13)	.07 (.06-.07)
<i>CFI</i>	.99	.99	.92	.92	.98

Convergent and Discriminant Validity

Convergent and discriminant validity assess how much a measurement instrument, in this case a survey scale, measures what it is designed to measure. A measurement scale demonstrates convergent validity when it correlates as expected with measures of theoretically similar constructs (or alternative measures of the same construct). Conversely, a scale demonstrates discriminant validity when it correlates minimally with measures of theoretically dissimilar constructs (for more, see Messick, 1995).

To examine convergent and discriminant validity, we analyzed how the well-being scales correlated 1) with each other, 2) with other Panorama survey scales, and 3) with demographic variables.

Based on existing research (e.g., see Larsen et al., 2001; Russell & Carroll 1999), we expected to find a positive correlation between the Positive Feelings and Challenging Feelings scales (given that the Challenging Feelings questions are reverse-scored). In other words, we expected students who reported more positive affect to also report less negative affect. As shown in Table 6, we found just that, with an effect size similar to what has been found in prior research, Spearman's $\rho = .56$, $p < .0001$. As expected based on social support literature (e.g., Asher et al., 1984; Reis et al., 2000), we found that students who reported more social support also reported more positive affect, $\rho = .33$, $p < .00001$, and less negative affect, $\rho = .21$, $p < .00001$.

We further examined the convergent and discriminant validity of our well-being topics by analyzing how they correlated with our pre-existing social-emotional learning and school climate topics. To do this, we relied on a separate, opportunistic sample of student responses from Panorama partner districts that used the well-being topics alongside other Panorama topics during Fall 2020 survey administrations. Because different districts selected different topics, the sample sizes for these correlations varied across topic pairs, from a minimum of 5,888 to a maximum of 37,380, with a median of 14,086 students. The results, also shown in Table 6, reveal correlations that vary in size proportionate to theoretical expectations. For example, one would expect very small correlations between students' growth mindset and their affective experience, or between their sense of school safety and their social support. On the other hand, one would expect strong relationships between students' sense of belonging and their affective experience or social support. The results match those expectations and, more generally, demonstrated stronger empirical relationships between more theoretically-related constructs.

Table 6. Topic Intercorrelations (Spearman's ρ)

	Positive Feelings	Challenging Feelings	Supportive Relationships
Positive Feelings		0.56	0.33
Challenging Feelings	0.56		0.21
Supportive Relationships	0.33	0.21	
Belonging	0.39	0.30	0.18
Engagement	0.35	0.24	
School Climate	0.30	0.24	
Self-Efficacy	0.30	0.26	
Emotion Regulation	0.30	0.34	
Social Awareness	0.29	0.17	
Self-Management	0.21	0.20	0.14
Grit	0.21	0.14	0.11
Growth Mindset	0.14	0.08	0.07
School Safety	0.10	0.19	0.07

Note: Green-shaded cells reflect topic correlations within the Well-Being Survey topics and pilot study, blue-shaded cells reflect correlations between well-being topics and other Panorama topics, and empty cells indicate correlations of topics with themselves or correlations lacking sufficient sample sizes (fewer than 5000 respondents).

Finally, we also examined how topic scores differed according to students' grade level and gender. Evidence from both cross-sectional and longitudinal studies reveal that as adolescents age, they experience positive emotions less frequently, negative emotions more frequently, and an increasing array of social-emotional challenges (for review, see Bailen et al., 2019; Eccles et al., 1993). And compared to boys, adolescent girls tend to experience more frequent negative emotions, particularly anxiety and sadness (Hankin et al., 1998; Nolen-Hoeksema & Girgus, 1994; Meininger et al., 2004). As shown in Figure 2, our results replicate these well-established effects. (Again, note that higher scores on the Challenges Feelings topic indicate less frequent negative emotions.)

Figure 2. Topic Means by Grade Level



By examining how students' responses to these well-being scales correlate with each other, other self-report measures, and student demographics, we find results entirely concordant with past research into well-being, emotion, and adolescent development. As such, we offer these findings as evidence of both convergent and divergent validity.

Conclusion

Students' social-emotional well-being matters. The positive and negative emotions students feel are essential characteristics of their psychology, indicators of their well-being, and mediators of their success in school and life. Supportive relationships with peers, school staff, and family members both indicate and promote student well-being. Panorama's Well-Being Survey measures these key aspects of students' social-emotional lives. It was developed based on an extensive review of scholarly literature, interviews with district leaders and educators from across the country, and best practices from the science of survey design. Data from hundreds of schools and tens of thousands of students were used to refine the instrument and document its psychometric properties. The final instrument, composed of three topics (Positive Feelings, Challenging Feelings, Supportive Relationships), meets or exceeds research standards for scale reliability and validity, and has already been used in hundreds of schools across the country.

We invite all educators interested in measuring and promoting student well-being to use this research-backed survey in their schools. In doing so, we believe it is essential for educators to contextualize these data with an appreciation for the personal, situational, and systemic factors that shape and support student well-being. Accurately interpreting and effectively acting on well-being data requires contextual knowledge about students and their environments. It's also important to appreciate that everyone's journey through school involves social-emotional ups and downs. Challenging moments of loneliness, frustration, or anxiety that serve as adaptive signals or growth opportunities are categorically different from chronic or extreme periods of suffering that interfere with daily functioning. In sum, student well-being deserves thoughtful measurement as well as thoughtful interpretation. Our survey content and platform enables the former, but educators' knowledge and wisdom ultimately determines the latter.

References

- Artino Jr., A. R., La Rochelle, J. S., Dezee, K. J., & Gehlbach, H. (2014). Developing questionnaires for educational research: AMEE Guide No. 87. *Medical Teacher*, 36(6), 463-474.
- Asher, S. R., Hymel, S., & Renshaw, P. D. (1984). Loneliness in children. *Child Development*, 55(4), 1456-1464.
- Bailen, N. H., Green, L. M., & Thompson, R. J. (2019). Understanding emotion in adolescents: A review of emotional frequency, intensity, instability, and clarity. *Emotion Review*, 11(1), 63-73.
- Becker, B. E., & Luthar, S. S. (2002). Social-emotional factors affecting achievement outcomes among disadvantaged students: Closing the achievement gap. *Educational Psychologist*, 37(4), 197-214.
- Bradburn, N. M., & Caplovitz, D. (1965). *Reports on happiness: A pilot study of behavior related to mental health* (No. 3). Chicago: Aldine Pub. Co.
- Chemaly, S. (2019, May 11). How women and minorities are claiming their right to rage. *The Guardian*.
- Cauce, A. M., & Srebnik, D. S. (1990). Returning to social support systems: A morphological analysis of social networks. *American Journal of Community Psychology*, 18, 609-616.
- Chu, P. S., Saucier, D. A., & Hafner, E. (2010). Meta-analysis of the relationships between social support and well-being in children and adolescents. *Journal of Social and Clinical Psychology*, 29(6), 624-645.
- DeVellis, R. F. (2016). *Scale development: Theory and applications* (Vol. 26). Thousand Oaks, CA: Sage.
- Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist*, 55(1), 34-43.
- Diener, E., Sandvik, E., & Pavot, W. (1991). Happiness is the frequency, not the intensity, of positive versus negative affect. In F. Strack, M. Argyle, & N. Schwarz (Eds.), *Subjective well-being: An interdisciplinary perspective* (pp. 119 -139). Oxford, England: Pergamon Press.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. Hoboken, NJ: John Wiley & Sons.
- Dubow, E. F., & Ullman, D. G. (1989). Assessing social support in elementary school children: The survey of children's social support. *Journal of Clinical Child Psychology*, 18(1), 52-64.
- Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. J. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, 48(2), 90-101.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299.
- Fischer, A. H., & Manstead, A. S. (2008). Social functions of emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (pp. 456-468). Guilford Press.
- Fowler, F. J., Jr. (2013). *Survey research methods*. Thousand Oaks, CA: Sage.
- Frey, B. (2018). *The SAGE encyclopedia of educational research, measurement, and evaluation* (Vols. 1-4). Thousand Oaks, CA: Sage.
- Gehlbach, H., & Artino, A. R. (2018). The survey checklist (manifesto). *Academic Medicine: Journal of the Association of American Medical Colleges*, 93(3), 360-366.

- Gehlbach, H., & Brinkworth, M. E. (2011). Measure twice, cut down error: A process for enhancing the validity of survey scales. *Review of General Psychology, 15*(4), 380-387.
- Gehlbach, H., & Chuter, C. (2020). Conceptualizing the Core of “Social Emotional Learning”. ACCESS: *Contemporary Issues in Education, 40*(1), 24-33.
- Hernández, M. M., Eisenberg, N., Valiente, C., Spinrad, T. L., VanSchyndel, S. K., Diaz, A., Berger, R. H., Silva, K. M., Southworth, J., & Piña, A. A. (2015). Observed emotion frequency versus intensity as predictors of socioemotional maladjustment. *Emotion, 15*(6), 699–704.
- Hankin, B. L., Abramson, L. Y., Moffitt, T. E., Silva, P. A., McGee, R., & Angell, K. E. (1998). Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study. *Journal of Abnormal Psychology, 107*(1), 128-140.
- Hiebert, E. H. (2005). In pursuit of an effective, efficient vocabulary program. In E. H. Hiebert & M. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 243-263). Mahwah, NJ: Erlbaum.
- Holgado-Tello, F. P., Chacon-Moscoso, S., Barbero-García, I., & Vila-Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity, 44*(1), 153–166.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185.
- House, J. S., Landis, K. R., & Umberson, D. (1988). Social relationships and health. *Science, 241*(4865), 540-545.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55.
- Kahneman, D., Diener, E., & Schwarz, N. (Eds.). (1999). *Well-being: The foundations of hedonic psychology*. New York: Russell Sage Foundation.
- Keyes, C. L. M., Shmotkin, D., & Ryff, C. D. (2002). Optimizing well-being: The empirical encounter of two traditions. *Journal of Personality and Social Psychology, 82*, 1007-1022.
- Kim, G., Walden, T., Harris, V., Karrass, J., & Catron, T. (2007). Positive emotion, negative emotion, and emotion control in the externalizing problems of school-aged children. *Child Psychiatry and Human Development, 37*(3), 221-239.
- Krosnick, J. A. (1999a). Maximizing questionnaire quality. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of Political Attitudes* (pp. 37-57). San Diego: Academic Press.
- Krosnick, J. A. (1999b). Survey research. *Annual Review of Psychology, 50*(1), 537-567.
- Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology, 81*(4), 684–696.
- Laverdière, O., Morin, A. J., & St-Hilaire, F. (2013). Factor structure and measurement invariance of a short measure of the Big Five personality traits. *Personality and Individual Differences, 55*(7), 739-743.
- Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research & Evaluation, 12*(2), 1-11.
- Lewis, A. D., Huebner, E. S., Reschly, A. L., & Valois, R. F. (2009). The incremental validity of positive emotions in predicting school functioning. *Journal of Psychoeducational Assessment, 27*(5), 397-408.
- Long, R. F., Huebner, E. S., Wedell, D. H., & Hills, K. J. (2012). Measuring school-related subjective well-being in adolescents. *American Journal of Orthopsychiatry, 82*(1), 50-60.
- Lorenzo-Seva, U. & Ferrando, P. J. (2020). Not positive definite correlation matrices in exploratory item factor analysis: Causes, consequences and a proposed solution. *Structural Equation Modeling: A Multidisciplinary*

Journal, 1-10.

- Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin*, 131(6), 803-855.
- Malecki, C. K., & Demaray, M. K. (2006). Social support as a buffer in the relationship between socioeconomic status and academic performance. *School Psychology Quarterly*, 21(4), 375-395.
- Meininger, J. C., Liehr, P., Chan, W., Smith, G., & Mueller, W. H. (2004). Developmental, gender, and ethnic group differences in moods and ambulatory blood pressure in adolescents. *Annals of Behavioral Medicine*, 28(1), 10-19.
- Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, 14(4), 5-8.
- Nolen-Hoeksema, S., & Girgus, J. S. (1994). The emergence of gender differences in depression during adolescence. *Psychological Bulletin*, 115(3), 424-443.
- Pollard, E.L. & Lee, P.D. (2003). Child well-being: A systematic review of the literature. *Social Indicators Research* 61, 59-78.
- Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2000). Daily well-being: The role of autonomy, competence, and relatedness. *Personality and Social Psychology Bulletin*, 26(4), 419-435.
- Rosenfeld, L. B., Richman, J. M., & Bowen, G. L. (2000). Social support networks and school outcomes: The centrality of the teacher. *Child and Adolescent Social Work Journal*, 17(3), 205-226.
- Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin*, 125(1), 3-30.
- Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, 52(1), 141-166.
- Sapolsky, R. M. (2004). *Why zebras don't get ulcers: The acclaimed guide to stress, stress-related diseases, and coping*. New York: Henry Holt and Company.
- Saris, W. E., Revilla, M., Krosnick, J. A., & Shaeffer, E. M. (2010). Comparing questions with agree/disagree response options to questions with item-specific response options. *Survey Research Methods*, 4(1), 61-79.
- Seligman, M. E., & Csikszentmihalyi, M. (2000). Positive psychology. An introduction. *American Psychologist*, 55(1), 5-14.
- Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences*, 106(35), 15073-15078.
- Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, 80(1), 99-103.
- Taylor, S.E.. (2011). Social support: A review. In H. S. Friedman (Ed.), *The Oxford handbook of health psychology* (pp. 89-214). Oxford University Press.
- Thoits, P. A. (1995). Stress, coping, and social support processes: Where are we? What next?. *Journal of Health and Social Behavior*, 53-79.
- VanderWeele, T. J. (2017). On the promotion of human flourishing. *Proceedings of the National Academy of Sciences*, 114(31), 8148-8156.
- Watson, D., & Clark, L. A. (1999). *The PANAS-X: Manual for the positive and negative affect schedule-expanded form*.

- Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology, 97*(3), 346-353.
- Weng, L. J. (2004). Impact of the number of response categories and anchor labels on coefficient alpha and test-retest reliability. *Educational and Psychological Measurement, 64*(6), 956-972. doi: 10.1177/0013164404268674
- Wentzel, K. R. (1998). Social relationships and motivation in middle school: The role of parents, teachers, and peers. *Journal of Educational Psychology, 90*(2), 202-209.
- Wethington, E., & Kessler, R. C. (1986). Perceived support, received support, and adjustment to stressful life events. *Journal of Health and Social Behavior, 27*, 78-89.
- Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. *Journal of Personality and Social Psychology, 43*(1), 111-122.



PANORAMA

EDUCATION

Panorama Education 2021
Revised March 8, 2021