Technical Brief:

Effective Learning Environments Observation Tool, version 2.0

V1.0, 01 10, 2023

1. Introduction

The Effective Learning Environments Observation Tool (eleot™) is a classroom observation tool designed to observe students in grades K-12 in the context of seven learning environments. Each item contained in the eleot describes how, in a learner-centric setting, learners demonstrate active engagement, agency, and self-efficacy. Collectively, the items promote an environment in which the students are the center of the instruction and activities and are responsible for their learning and interactions with other learners. In this type of classroom setting, the teacher becomes a facilitator who assists students in their learning activities, as opposed to the teacher serving primarily as the provider of information and knowledge, placing learners in a passive or receptive role. In essence, the eleot items, once rated, measure the extent to which students' observable behaviors indicate their engagement in certain activities, demonstration of knowledge, attitudes, and dispositions in a classroom during a defined period.

2. Background

The eleot 1.0 was developed in 2012 by the Innovation Division at AdvancED (now R&D as Cognia) to provide accreditation teams with an instrument that produced “real-time” data and evidence for rating accreditation standards that described expectations for high-quality teaching and learning. The concept of developing an observation instrument that focused on learners and not the teacher was a departure from the intent and use of most observation instruments. To begin the conceptualization, the key authors immersed themselves in the mindset of learners in a classroom, focusing on these essential questions, “What do learners experience during class time? How do learners respond and react to the instruction? Are learners engaged in the learning, and if so, what are they doing? In what ways do they interact with their peers?”

After thousands of observations conducted using eleot 1.0 over a span of four years, data analysis, combined with information from institutions, staff, and volunteers, prompted the consideration of a revised eleot. After collecting and analyzing data, such as the highest-scoring and lowest-scoring items, anecdotal information about confusing wording, and items that needed refreshing and combining, the original authors and a committee comprised of practicing educators and staff members revised the eleot 1.0. As a result, the number of items reduced from 30 to 28, several items moved to a better-aligned learning environment, the item stems changed to focus on the learners, and items were added or combined for clarity and to reflect new research in learner-centric instructional approaches. Additionally, the eleot 2.0 was created for Cognia’s platform and used as an application.

3. Instrument Overview

The Effective Learning Environments Observation Tool (eleot®) comprises 28 items organized in seven environments based on Cognia's performance standards, research on effective teaching and learning, and a review on learner-centric tasks, attitudes, and dispositions. In essence, the eleot measures the extent to which certain learning behaviors, attitudes, and actions are present (or absent) in a classroom during a defined period of time. The environments examined are:

A. Equitable Learning
B. High Expectations
C. Supporting Learning
D. Active Learning
E. Progress Monitoring and Feedback
F. Well-Managed Learning
G. Digital Learning

4. Training and Administration

The eleot is used by observers trained and certified by Cognia. To become eleot certified, one must complete an asynchronous, online training course, conduct a virtual observation of a 20-minute classroom video segment using the eleot, and obtain a minimum of 90% agreement with expert ratings.

Cognia offers year-round access to the survey, and members can choose to administer eleot 2.0 at any point they consider appropriate. To date, nearly 700,000 observations across the world have been conducted using eleot. In the 2021-22 administration, the distribution of administration dates differed between U.S. schools and international schools. Many U.S. schools administered the eleot 2.0 in late fall and early spring, as shown in Figure 1. International schools, however, presented a different pattern: popular administration months were from October to May, as shown in Figure 2.

Figure 1: Distribution of Observation Responses Throughout the Year for U.S. Schools

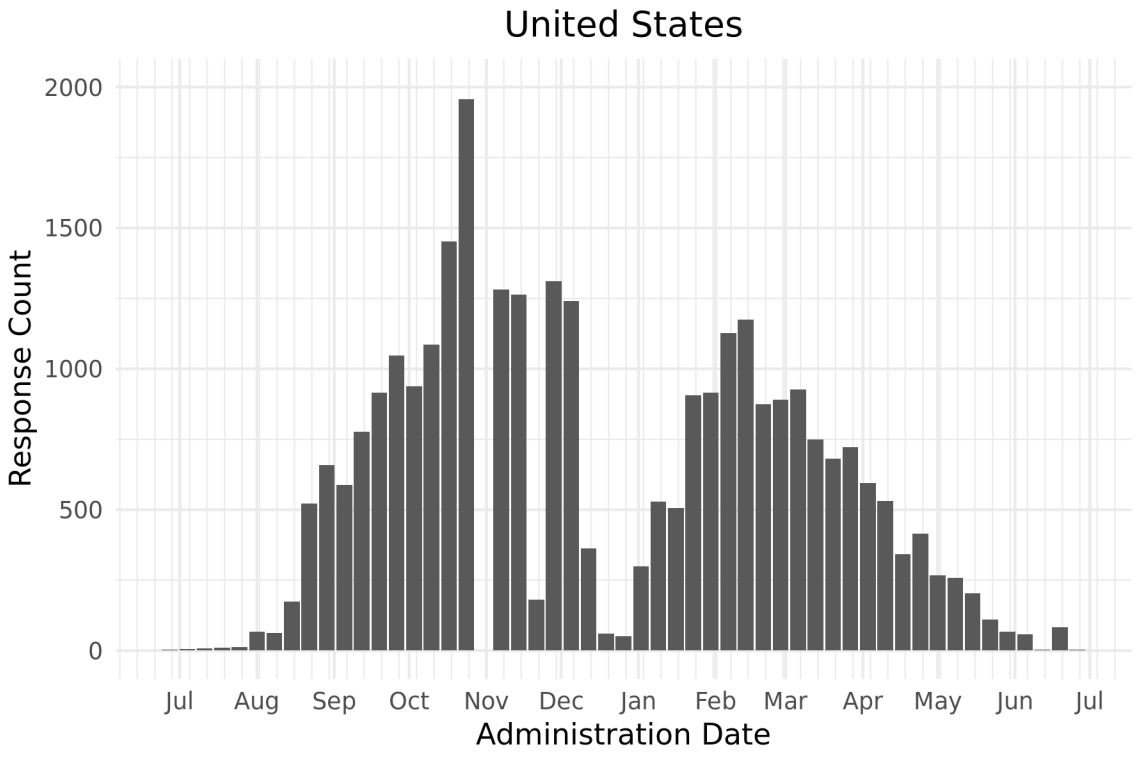
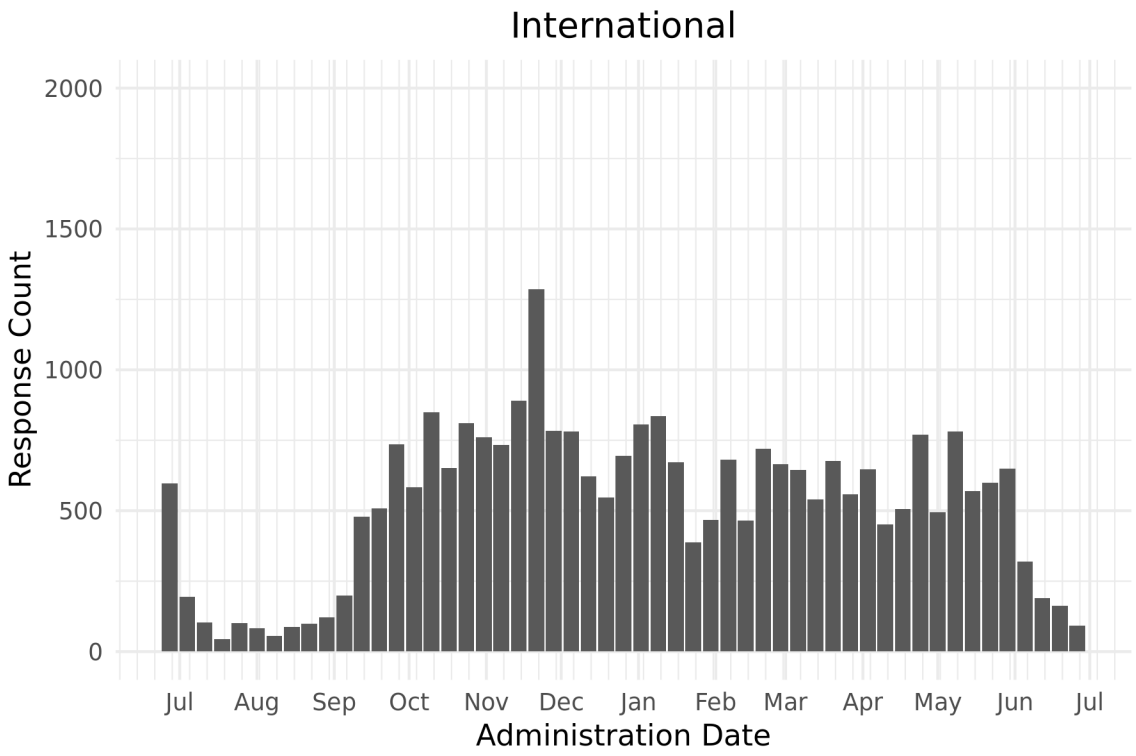


Figure 2: Distribution of Observation Responses Throughout the Year for International Schools



5. Participation

Results described in this report are based on completed observations by Cognia-certified observers from 59,104 individual observations in grade K through grade 13 from July 2021 through June 2022.

Participation by Geographical Regions

These observations were conducted by 1,733 educational institutions globally. 73.5% of these institutions were from the United States and administered 53.2% of the total observations. The26.5% international institutions administered close to half (46.8%) of the total observations (Table 1).

Table 1. Participation by Geographical Regions

	Observations		Observers		Institutions	
Region	n	%	N	%	n	%
International	27746	46.8	1548	40.6	458	26.5
Mid-Atlantic	12857	21.8	698	18.4	342	19.7
Midwest	1767	3.0	166	4.4	113	6.5
Mountain	4536	7.7	283	7.4	171	9.9
Northeast	877	1.5	59	1.6	44	2.5
Pacific	762	1.3	67	1.8	36	2.1
Southeast	10558	17.9	979	25.8	569	32.8
Overall	59103	100	3800	100	1733	100

Participation by Grades

All grade levels of schools from pre-K through grade 13 were represented in the 2021-22 administration. An observation can cover more than one grade when the classroom has students from mixed grades. The number of observations were nearly evenly distributed across grades with a slightly higher number in grade 9, 3, 4, and 2. By grade span, 2,435 schools have early elementary grades (K-2), 2,529 schools have elementary grades (3-5), 2,272 schools have middle school grades (6-8), and 2,455 schools have high school grades (9 and up) (Table 2).

Table 2. Participation by Grades

	Observations	Observers	Institutions

Grade	n	%	N	%	n	%
K	4901	7.7	1109	6.7	741	7.6
1	5168	8.1	1357	8.2	839	8.6
2	5248	8.2	1367	8.3	855	8.7
3	5342	8.4	1390	8.4	856	8.7
4	5291	8.3	1417	8.6	851	8.7
5	4776	7.5	1373	8.3	822	8.4
6	4495	7.0	1305	7.9	769	7.8
7	4796	7.5	1290	7.8	753	7.7
8	4856	7.6	1241	7.5	750	7.6
9	5361	8.4	1217	7.4	656	6.7
10	4950	7.8	1167	7.1	624	6.4
11	4768	7.5	1113	6.8	601	6.1
12	3596	5.6	931	5.7	522	5.3
13	116	0.2	55	0.3	52	0.5
N/A	155	0.2	128	0.8	115	1.2

Participation by Subjects

Observations were performed in classes covering 15 different subject areas (Table 3). Over 50% of the observations were made in Language Art and Mathematics classes, followed by Science and Social Studies classes (11.5% and 7.7%, respectively).

Table 3. Participation by Subjects

	Observations		Observers		Institutions	
Subject	n	%	n	%	n	%
Art	1278	2.2	702	4.8	527	5.4
Business	322	0.5	184	1.3	142	1.5
Career and Technical Education	940	1.6	310	2.1	233	2.4
Computers and Technology	782	1.3	404	2.8	331	3.4

Foreign Language	3279	5.5	752	5.1	500	5.1
Health	224	0.4	149	1.0	140	1.4
Language Arts	18504	31.3	2679	18.3	1521	15.6
Mathematics	12836	21.7	2485	17	1442	14.8
Multi-Subject	893	1.5	356	2.4	317	3.3
Music	993	1.7	590	4.0	455	4.7
Physical Education	1544	2.6	746	5.1	571	5.9
Religious Studies	1234	2.1	374	2.6	276	2.8
Science	6821	11.5	1873	12.8	1171	12.0
Social Studies	4538	7.7	1557	10.6	1039	10.7
Other	4391	7.4	1172	8.0	784	8.1
N/A	525	0.9	324	2.2	272	2.8
Overall	59104	100	14657	100	9721	100

Participation by Class Segments

An observation can occur across multiple time points during classes (beginning, middle, and end). The majority of observations happened in the beginning and middle of classes (40.9% and 40.8%, respectively), while only 18.3% were made during the end of classes (Table 4).

Table 4. Participation by Class Segments

	Observations		Observers		Institutions	
Class Segment	n	%	n	%	n	%
Beginning	34909	40.8	3457	37.3	1646	35.5
Middle	34857	40.7	3428	37.0	1642	35.4
End	15686	18.3	2264	24.4	1240	26.8
N/A	135	0.2	116	1.3	107	2.3

6. Descriptive Analysis

The eleot 2.0 consists of items categorized by seven environments of observations: Equitable Learning (items A1 – A4), High Expectations (items B1 – B5), Supportive Learning (items C1 – C4), Active Learning (items D1 – D4), Progress Monitoring and Feedback (items E1 – E4), Well-Managed Learning (F1 – F4), and Digital Learning (G1 – G3). Point values from 1 to 4 correspond to four evidence levels: Not Observed, Somewhat Evident, Evident, and Very Evident.

6.1 Item Analysis

Table 5 provides item statistics including means (average response to an item), standard deviations ("sd"; dispersion of scores; how "spread out" the scores are for that item), item correlations ("corr w total"; relationship between item score and total scores), and answer option percentage (proportion of each answer option). Figure 3 displays the proportion of each answer options by item. Item text for each environment can be found in Appendix A.

Notable Findings

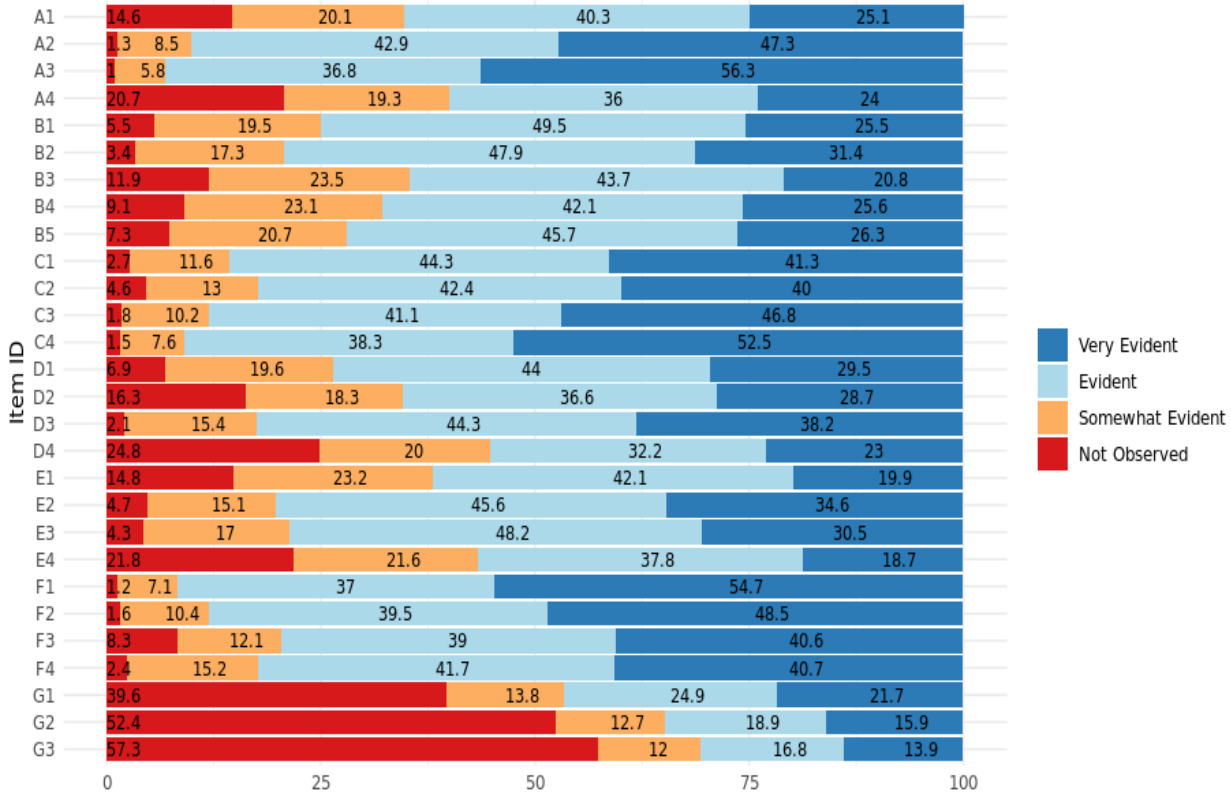
- Item analysis showed mean scores of items ranging from 1.87 (Item 3 in Digital Learning) to 3.48 (Item 3 in Equitable Learning).
- The median value of these means is 3.01. Standard deviation values were clustered relatively closely to the means (0.65 to 1.20).
- Most (21 out of 28) of the item scores were within the first standard deviation, indicating scores were scattered closely around the mean and less spread out.
- Item score to total score correlations ranged from 0.50 to 0.73. The lowest correlations were all in the Digital Learning environment.
- Item 3 in Equitable Learning (Item A3) and Item 1 in Well-Managed Learning (Item F1) had the highest proportion of "Very Evident," with 56.3% and 54.7% respectively.

Table 5. Item Descriptive Statistics

						Answer Option Percentage			
Environment	Item ID	n	mean	sd	corr w total	p1	p2	p3	p4
Equitable Learning	A1	59104	2.76	0.99	0.62	14.6	20.1	40.3	25.1
	A2	59104	3.36	0.69	0.64	1.3	8.5	42.9	47.3
	A3	59104	3.48	0.65	0.62	1.1	5.8	36.8	56.3
	A4	59104	2.63	1.06	0.59	20.7	19.3	36.0	24.0
High Expectations	B1	59104	2.95	0.82	0.73	5.5	19.5	49.5	25.5
	B2	59104	3.07	0.79	0.73	3.4	17.3	47.9	31.4
	B3	59104	2.73	0.92	0.71	11.9	23.5	43.7	20.8

	B4	59104	2.84	0.91	0.70	9.1	23.1	42.1	25.7
	B5	59104	2.91	0.87	0.68	7.3	20.7	45.7	26.3
Supportive Learning	C1	59104	3.24	0.76	0.69	2.7	11.6	44.3	41.3
	C2	59104	3.18	0.83	0.67	4.6	13.0	42.4	40.0
	C3	59104	3.33	0.73	0.70	1.8	10.2	41.1	46.8
	C4	59104	3.42	0.70	0.66	1.5	7.6	38.4	52.5
Active Learning	D1	59104	2.96	0.88	0.67	6.9	19.6	44.0	29.5
	D2	59104	2.78	1.04	0.57	16.3	18.3	36.6	28.7
	D3	59104	3.19	0.77	0.72	2.1	15.4	44.3	38.2
	D4	59104	2.53	1.10	0.58	24.8	20.0	32.2	23.0
Progress Monitoring and Feedback	E1	59104	2.67	0.96	0.68	14.8	23.2	42.1	19.9
	E2	59104	3.10	0.82	0.71	4.7	15.1	45.6	34.6
	E3	59104	3.05	0.80	0.73	4.3	17.0	48.2	30.5
	E4	59104	2.53	1.03	0.65	21.8	21.6	37.8	18.7
Well-Managed Learning	F1	59104	3.45	0.68	0.65	1.2	7.1	37.0	54.7
	F2	59104	3.35	0.73	0.65	1.6	10.4	39.5	48.5
	F3	59104	3.12	0.92	0.62	8.3	12.1	39.0	40.6
	F4	59104	3.21	0.78	0.68	2.4	15.2	41.7	40.7
Digital Learning	G1	59104	2.29	1.20	0.50	39.6	13.8	24.9	21.7
	G2	59104	1.98	1.16	0.50	52.4	12.7	18.9	15.9
	G3	59104	1.87	1.13	0.50	57.3	12.0	16.8	13.9

Figure 3. Percentage of Each Answer Options by Item



6.2 Environment Analysis

Notable Findings

- Mean scores for all environments ranged from 2.05 to 3.31. Supportive Learning and Well-Managed Learning had the highest mean scores, while Digital Learning had the lowest mean score and the highest standard deviation.
- All environments but Digital Learning were highly correlated with total scores.
- Items in the Digital Learning environment also showed the highest proportion of “Not Observed” among all environments.
- Schools in the network performed the best in the Supportive Learning environment and the Well-Managed Learning environment. Items in these two environments had high mean scores, low standard deviations, and high proportions of “Evident” and “Very Evident.”

Table 6 shows descriptive statistics for each environment.

Table 6. Environment Descriptive Statistics

Environments	n	mean	sd	min	max	corr w total
A. Equitable Learning	59104	3.08	0.62	1	4	0.84
B. High Expectations	59104	2.90	0.69	1	4	0.88
C. Supportive Learning	59104	3.31	0.62	1	4	0.82
D. Active Learning	59104	2.89	0.71	1	4	0.83
E. Progress Monitoring	59104	2.86	0.72	1	4	0.85
F. Well-Managed Learning	59104	3.30	0.65	1	4	0.77
G. Digital Learning	59104	2.05	1.04	1	4	0.56

7. Reliability and Validity

7.1 Internal Consistency

Observation responses were used to investigate the evidence for the internal consistency reliability across the instruments. Reliability was calculated according to Cronbach's Alpha (α) and McDonald's omega (ω)[1]. eleot 2.0 demonstrates α values of 0.944, and ω values of 0.962. These values provide evidence that the eleot 2.0 demonstrates sufficient reliability to make decisions about effective learning enrironments[2]. Table 7 provides reliability for the whole instrument and each environment.

Table 7. Internal Consistency and Reliability of Environments

Reliability	A	B	C	D	E	F	G	Total
Alpha (α)	0.681	0.861	0.850	0.735	0.812	0.849	0.878	0.944
Omega (ω)	0.740	0.866	0.853	0.740	0.825	0.862	0.894	0.962

A: Equitable Learning; B: High Expectations; C: Supportive Learning; D: Active Learning; E: Progress Monitoring and Feedback; F: Well-Managed Learning; G: Digital Learning.

7.2 Inter-Rater Reliability

An important piece of evidence to evaluate for a classroom observation instrument such as eleot is inter-rater reliability. The results of an inter-rater reliability analysis indicate the degree of agreement

[1] Though Cronbach's Alpha is widely used and understood as a measure of reliability, McDonald's omega is a distinct reliability coefficient with the advantage of considering the strength of association between items, leading to a stronger measurement of the scale's internal consistency.

[2] Salvia, J., & Ysseldyke, J. E. (2004). Assessment in special and inclusive education (9th ed.). Boston: Houghton Mifflin.

among raters, for example the agreement between raters or observers on any single classroom observation evaluated using eleot. This is important to ensure that results can be attributed to the observed instructional practice and not as a function of the rater or observer.

One method for analyzing inter-rater reliability is the intra-class correlation coefficient (ICC) which is appropriate for categorical scales such as eleot ratings and is interpreted as the agreement, or correlation, between the raters if they were to conduct the same observation. Values of the ICC range from 0 to 1, where ICC = 1 indicates that any two raters would score observations the same whereas 0 indicates no agreement. Cicchetti (1994) provides commonly quoted guidelines for interpretation of ICC values: poor reliability (ICC < 0.40), fair (ICC < 0.60), good (ICC < 0.75), and excellent (ICC ≥ 0.75).

A total of n = 59,104 eleot observations were conducted by 3,787 certified observers at 1,733 institutions. The results of this analysis are presented in Table 8 for the overall score and for each Environment, including 95% confidence intervals.

According to the criteria above, evidence for fair reliability of overall scores is demonstrated (ICC1 = 0.569, 95%CI = 0.564-0.575) [3]; per environment agreement ranges ICC1 = 0.464 to 0.561, which is categorized as fair consistency. It must be noted that the environment scores are the result of subsets of items on the eleot instrument and, therefore, measures of reliability are expected to demonstrate some deflation. Additionally, these results provide comparable or improved agreement over similar classroom observation instruments (ICC = 0.26-0.51; Pianta, Hamre, & Mintz, 2012). These results provide evidence supporting reliability of eleot for use in evaluating classroom practice.

Table 8. Intra-Class Correlation Results and Confidence Intervals for eleot Overall Score and by Environment

eleot Environment	ICC1 (95% CI)
Overall	0.569 (0.564-0.575)
A	0.517 (0.511-0.522)
B	0.471 (0.464-0.478)
C	0.485 (0.479-0.491)
D	0.423 (0.417-0.429)
E	0.522 (0.514-0.529)

[3] There are a variety of methods for calculating ICC, depending on the data collection design. For this analysis, each eleot observation was conducted by a single rater; therefore, ICC1 is calculated according to a one-way random effects model evaluating inter-rater reliability (Shrout & Fleiss, 1979).

eleot Environment	ICC1 (95% CI)
F	0.423 (0.418-0.429)
G	0.463 (0.453-0.472)

7.3 Validity Evidence

Validity evidence for the measurement of eleot 2.0 is presented in Table 9 as correlations between environments for the whole instrument. Specifically, correlations of responses across environments are presented in the off-diagonal cells and represent discriminant validity evidence (indicating the degree to which environments are separate and distinct from one another), while convergent validity evidence is presented in the cells along the diagonal as reliability (i.e., ω; indicating the consistency of measurement within the environments).

Within-environment reliabilities range from 0.740 to 0.894, demonstrating sufficient and strong convergent validity evidence. Across-environment correlations are moderate to strong and range from 0.570 to 0.774, with the exception of Digital Learning environment; however, the across-environment reliabilities (i.e., discriminant validity evidence) are consistently lower than the associated within-environment reliabilities. These results indicate that the eleot 2.0 consistently measures effective learning environment within each of the seven environments, which are related to each other, but demonstrate sufficient independence as to be considered separate traits or constructs.

Table 9. Convergent and Discriminant Validity of Environments

Environments	Equitable Learning	High Expectations	Supportive Learning	Active Learning	Progress Monitoring	Well-Managed Learning	Digital Learning
Equitable Learning	**0.740**						
High Expectations	0.703	**0.866**					
Supportive Learning	0.693	0.675	**0.853**				
Active Learning	0.675	0.709	0.645	**0.740**			

Progress Monitoring	0.662	0.774	0.638	0.669	**0.825**		
Well-Managed Learning	0.630	0.622	0.731	0.570	0.589	**0.862**	
Digital Learning	0.370	0.379	0.261	0.361	0.397	0.253	**0.894**

7.4 Construct Validity

Additional validity evidence for Cognia's eleot 2.0 is demonstrated using statistical techniques to evaluate the underlying theoretical structure, or latent factor model. The relationship among items, environments (A: Equitable Learning; B: High Expectations; C: Supportive Learning; D: Active Learning; E: Progress Monitoring and Feedback; F: Well-Managed Learning; G: Digital Learning), and overall observation on effective learning environment can be represented according to two latent factor models. The Unidimensional Model conceives effectiveness observation score on items as simply correlated to overall observation score; the First-Order Model suggests that effectiveness observation on each item is correlated with a specific learning environment and those environments are correlated with each other.

Confirmatory Factor Analysis (CFA) is the statistical technique used to evaluate whether the Unidimensional Model or the First-Order Model fits the data best and to confirm that the target model fits the data well. Model comparison is conducted according to a Likelihood-Ratio Test, with significant differences indicating that the more complex (i.e., First-Order) model fits best; model evaluation is conducted by comparing standard model fit statistics (RMSEA and χ^2/df) against critical values in an empirical sampling distribution that was generated according to a bootstrap resampling procedure (replications = 100; α = 0.05).

The model evaluation found that RMSEA and χ^2/df values did not exceed the empirically determined significance values; therefore, the models all fit the data sufficiently well to enable further comparison and interpretation[4]. Model comparison according to the results of Likelihood-Ratio Tests between the Unidimensional Model and First-Order Model found that First-Order Model fits significantly better than the Unidimensional Model. Together, these results provide evidence supporting measurement and reporting of Cognia's eleot 2.0 observation results according to the seven learning environments. Table 10 summarizes model evaluation and comparison results.

[4] Kline, R. B. (2013). *Beyond significance testing: Statistics reform in the behavioral sciences, 2nd Edition*. Washington, DC: American Psychological Association.

Table 10. Confirmatory Factor Analysis Model Summary

Model	N	df	RMSEA	χ2/df	Notes
Unidimensional	59104	350	0.143	1211.02	
First-Order	59104	329	0.063	238.66	*

** Preferred model based on LRT results*

For Additional Information

For any additional questions about this study contact the authors:

Xiaonan Zhang, Measurement Scientist: xiaonan.zhang@cognia.org

Matthew Gushta, Ph.D., Vice President, Research & Development: matthew.gushta@cognia.org

Research & Development Department: researchdevelopment@cognia.org

Appendix A

Table A1: Cognia's eleot 2.0 Instrument Item Text

Environment	Item	Item Text
Equitable Learning	A1	Learners engage in differentiated learning opportunities and/or activities that meet their needs
	A2	Learners have equal access to classroom discussions, activities, resources, technology, and support
	A3	Learners are treated in a fair, clear, and consistent manner
	A4	Learners demonstrate and/or have opportunities to develop empathy/respect/appreciation for differences in abilities, aptitudes, backgrounds, cultures, and/or other human characteristics, conditions and dispositions
High Expectations	B1	Learners strive to meet or are able to articulate the high expectations established by themselves and/or the teacher
	B2	Learners engage in activities and learning that are challenging but attainable
	B3	Learners demonstrate and/or are able to describe high quality work
	B4	Learners engage in rigorous coursework, discussions, and/or tasks that require the use of higher order thinking (e.g., analyzing, applying, evaluating, synthesizing)
	B5	Learners take responsibility for and are self-directed in their learning
Supportive Learning	C1	Learners demonstrate a sense of community that is positive, cohesive, engaged, and purposeful
	C2	Learners take risks in learning (without fear of negative feedback)
	C3	Learners are supported by the teacher, their peers and/or other resources to understand content and accomplish tasks
	C4	Learners demonstrate a congenial and supportive relationship with their teacher
Active Learning	D1	Learners' discussions/dialogues/exchanges with each other and the teacher predominate
	D2	Learners make connections from content to real-life experiences
	D3	Learners are actively engaged in the learning activities
	D4	Learners collaborate with their peers to accomplish/complete projects, activities, tasks and/or assignments

Progress Monitoring and Feedback	E1	Learners monitor their own learning progress or have mechanisms whereby their learning progress is monitored
	E2	Learners receive/respond to feedback (from teachers/peers/other resources) to improve understanding and/or revise work
	E3	Learners demonstrate and/or verbalize understanding of the lesson/content
	E4	Learners understand and/or are able to explain how their work is assessed
Well-Managed Learning	F1	Learners speak and interact respectfully with teacher(s) and each other
	F2	Learners demonstrate knowledge of and/or follow classroom rules and behavioral expectations and work well with others
	F3	Learners transition smoothly and efficiently from one activity to another
	F4	Learners use class time purposefully with minimal wasted time or disruptions
Digital Learning	G1	Learners use digital tools/technology to gather, evaluate, and/or use information for learning
	G2	Learners use digital tools/technology to conduct research, solve problems, and/or create original works for learning
	G3	Learners use digital tools/technology to communicate and/or work collaboratively for learning

References

Cicchetti, D.V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment, 6(4), 284–290.

Pianta, R.C., Hamre, B.K., and Mintz, S. (2012). Upper Elementary/Secondary CLASS® Manual Technical Appendix: Web Version. Teachstone.

Shrout P.E., Fleiss J.L. (1979). Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin, 86, 420–428.