# APPENDIX G NEXT-GENERATION SCORING SPECIFICATIONS



## MCAS Next-Gen & Legacy 2022–2023



## **Scoring Specifications**

Part A:

Client-Specific Scoring Specifications

Part B:

General Scoring Guidelines & Best Practices



## Part A: Client-Specific Scoring Specifications

Contract Name:	MCAS Gr 3-8 and HS		Year:	2022-2023		
Contract Code:	1637					
Contact Information:	Program Managers: Chris Clough, Dezarae Blossomgame					
	Scoring Services at Cognia:					
	ELA: Sandy Sinclair, Melissa Polansky, Sean Burke, Vince McGroary  Mathylania Magazine Magazine  Mathylania Magazine Magazine  Mathylania Magazine  Mat					
	Math: Janice Knox, Andrea Kuegel					
	Science: Rozanna Gaines					
	•	Goring Services at Pearson:  ELA: Stephen Hoffelt				
	•					
	Math: Trace  Project Management	ey Benvin				
	Project Management:	rah Melmed (in transition) & Ka	rin Evana (transitional cupr	oort)		
	•	aula Schwartz	IIII Evalis (transitional supp	oort)		
A dustriatuation or			UC Detect Diele 9 Dhu	since Manch 100 FLA 9 Moth Co. 40		
Administrations:	Retests; Spring '23 EL		nd Science Gr 5, 8, and H	sics; March '23 ELA & Math Gr 10 S Physics & Biology; Spring '23		
Scoring Plan:	This scoring plan summ		ing of all MCAS Next Gen	test administrations for all contents and		
	Technology/Engineering		by test duffills for 110 ocion	ice (Orientistry and		
		conducted applying a virtual/sy	nchronous scoring model.			
	-		-	ons to support their readiness and to		
		ology-related questions.	•			
	<ul> <li>Scorers will evalu</li> </ul>					
	Scorers will work in a non-public setting and will be required to be on camera during training and scoring. Scorers					
	may blur their backgrounds.					
	Training and all in					
	and/or via pre-red	orded training module or a reco	ording of live training.			
	<ul> <li>Breakout rooms n</li> </ul>	nay be used to facilitate scorer	training and individualized	coaching.		
	<ul> <li>DESE will have re</li> </ul>	emote access to the scoring sys	stem and Zoom/Teams links	s will be provided in advance to observe		
	training sessions	and scoring with technical assis	stance, if needed.			
	<ul> <li>A post-scoring su</li> </ul>	rvey will be sent out to all MCA	S scoring associates to elic	it feedback on their scoring experience.		
	The results will be shared with DESE.					
	• As a result of DESE's decision to remove the 0/1 rule for essays in grades 3-8, the following actions will be taken:					
	<ul> <li>The asterisk referring to the 0/1 rule will be removed from all rubrics for gr. 3-8.</li> </ul>					
				e, to show essays that score a 0 in Idea		
		arrant a range of Convention so				
		ntal responses will be provided		-		
	• • • • • • • • • • • • • • • • • • • •		ved and approved during th	ne conference calls scheduled to finalize		
	the 2023 scoring materials.  o Pearson will have the scoring director for gr. 7 attend the scheduled conference call on 3-2-23.					
		-				
T (1 D) (1		sible, the scoring directors for of	ner grades will also attend	the scheduled ELA conference calls.		
Testing Platform:	□ TestNav     □					
Scoring Platform:		10 ELA and Math, FT Gr 3-8 E	LA and Math, Gr 5, 8 and F	IS Science		
A.1	☑ ePEN: OP Gr 3-8 ELA and Math					
Admin Type:	☑ Operational	☑ Field Test:	Note: Embedded: Math, ELA, a	nd Science		
		☐ Standalone	Embouded. Math, ELA, a	ind Cololloc		
		⊠ Embedded				
Paguired Client Mestings	Danahmarkina	Notes: Panahmarking mosting	ngo will be held at most called	agracable dates and times in Chrise		
Required Client Meetings:	Benchmarking	Notes: Benchmarking meeting 2023 to determine the scoring		agreeable dates and times in Spring		
		2020 to dotomine the scomi	granos ioi an i i ileiris.			



			Table 1 -	Estimated Studen	t Count per Grade	, Content, and Test	Administration	
	Content	3	4	5	6	7	8	HS
Math	Nov. Retest							Total 12,287 PBT 201 CBT 12,086
	March Retest							Total 5,130 PBT 239 CBT 4,891
	Spring OP	Total 70,000 CBT only	Total 69,587 PBT 574 CBT 69,013					
ELA	Nov. Retest							Total 7,637 PBT 119 CBT 7,518
	March Retest							Total 3,278 PBT 160 CBT 3,118
	Spring OP	Total 70,000 CBT only	Total 69,241 PBT 458 CBT 68,783					
Science	Feb. Retest							Biology: 10,704 PBT 200 CBT 10,504 Physics: 6,000 PBT 86 CBT 5,914
	Spring OP			Total 70,000 CBT only			Total 70,000 CBT only	Biology: 63,693 PBT 686 CBT 63,007 Physics: 15,102 PBT 86 CBT 15,016
	Spring OP Legacy							Chem.: 1,058 Tech/Eng.: 500 PBT only
Estimated S	Spanish Test Takers:							
Math	March Retest							Total 779 PBT 46 CBT 733
	Spring OP							Total 1,303 PBT 68 CBT 1,235
Science	Feb. Retest							Biology: 364 PBT 36 CBT 328 Physics: 183 PBT 10 CBT 173
	Spring OP							Biology: 2,166 PBT 39 CBT 2,127 Physics: 460 PBT 19 CBT 441

#### Table 2 - Scope of Work by Number and Item Type per Grade and Administration

The 2022-2023 MCAS consists of both operational and matrix test items.

Cognia is responsible for all aspects of scoring with Pearson serving as sub-contractor for the operational scoring of Gr 3-8 ELA and Math. Pearson recruits scoring associates for their assigned scoring activities.

		Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade HS
ELA	Nov. Retest							2 OP ES 5-3
	March Retest							2 OP ES 5-3
	Spring OP	1 OP ES 4-3 1 OP CR3 2 EQ CR3 1 EQ ES 4-3 3 FT ES 4-3 6 FT CR3	1 OP ES 4-3 1 OP CR3 2 EQ CR3 1 EQ ES 4-3 3 FT ES 4-3 6 FT CR3	2 OP ES 4-3 2 EQ ES 4-3 6 FT ES 4-3	2 OP ES 5-3 2 EQ ES 5-3 6 FT ES 5-3	2 OP ES 5-3 2 EQ ES 5-3 8 FT ES 5-3	2 OP ES 5-3 2 EQ ES 5-3 6 FT ES 5-3	2 OP ES 5-3 1 EQ ES 5-3 20 FT ES 5-3
Math	Nov. Retest							4 OP OE4
	March Retest							4 OP OE4
	Spring OP	4 OP OE3 2 EQ OE3 7 FT OE3	4 OP OE4 2 EQ OE4 8 FT OE4	4 OP OE4 2 EQ OE4 8 FT OE4	4 OP OE4 2 EQ OE4 7 FT OE4	4 OP OE4 2 EQ OE4 7 FT OE4	4 OP OE4 2 EQ OE4 7 FT OE4	4 OP OE4 2 EQ OE4 24 FT OE4
Science	Feb. Retest							Biology: 5 OP CR4 Physics: 5 OP CR4
	Spring OP			2 OP CR2 4 OP CR3 1 EQ CR2 2 EQ CR3 4 FT CR2 10 FT CR3			2 OP CR2 4 OP CR3 1 EQ CR2 2 EQ CR3 4 FT CR2 10 FT CR3	Biology: 2 OP CR3 3 OP CR4 12 FT CR3 12 FT CR4 Physics: 2 OP CR3 3 OP CR4 12 FT CR3 12 FT CR4
	Spring OP Legacy							Chemistry: 5 OP CR4 Tech/Eng: 5 OP CR4
		OP = Operational FT = Field Test CR# = #-point Cons	structed Resp.	OE3 = 0-3 point of	open ended respon open ended respon		0-5 & 0-3 points	



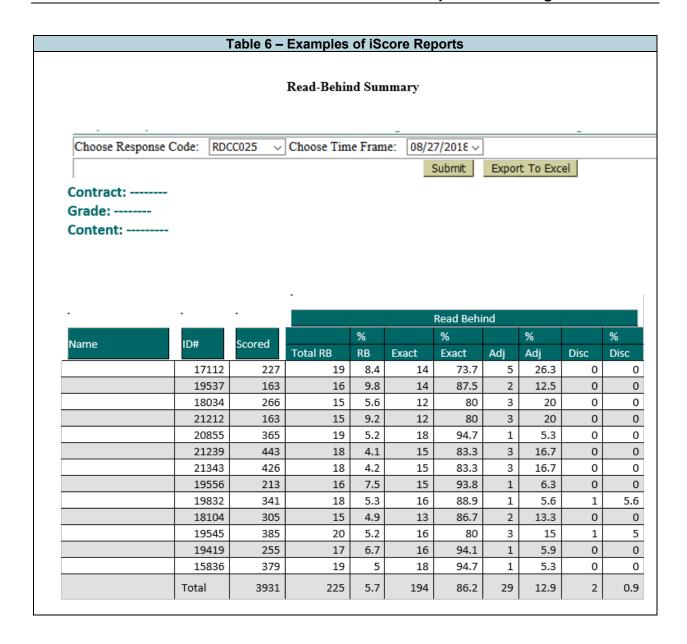
		Table 3	- Quality Contro	ol Tools		
Qualifying Sets	⊠ OP	QTY: 2 sets  Notes: Scorers are required to take Qualification Set 2 if the on Qualification Set 1.		s are required to take Qualification Set 2 if the threshold is not met n Set 1.		
	⊠ FT	QTY: 1 set	Notes:			
Qualification Threshold (%)	Leadership: Exact: 80% Exact + Adjacent: 90%,	, 1 Discrepant allowed		Scorers: Exact: 70% Exact + Adjacent: 90%, 1 Discrepant allowed		
	Clarification notes: For multi-trait ELA items,	the passing thresholds m	nust be met on e	each individual trait.		
Read-Behind Rate	Minimum daily requirer All Grades and Content A entire day.	Minimum daily requirement per Scorer: All Grades and Content Areas: 10 responses minimum for a full day. This number will be proportionate for shifts that do not last an entire day.				
Double-Blind Rate	Minimum (%): Operational scoring Grades 3-8 ELA & Math: 10% Operational scoring Sci 5 & 8: 10% Operational HS: 100%  Field Test 3-8 ELA: 20% Field Test 3-8 Math: 10% Field Test 5 & 8 Sci: 10% Field Test HS ELA, Math, and Science: 10%					
Recalibration Sets	⊠ Standalone  □ Embedded	Number of recalibration sets:	on 1 set dai	ly		
	□ N/A	Number of responses per set:	5 respon	ises		
	When Administered?	Beginning on the second day of operational scoring for each item and each day until scoring of each ite is complete.				
	Notes:	Please refer to comparison chart of scoring terminology and practices as applied by Cognia and by Pearson.				
Embedded Responses	When Administered?	Grades 5, 8, HS Science & Grade 10 ELA, Math: 10 responses deployed during the first 100 responses scored by a scorer				
Validity Responses	Required? ⊠ Yes □ N/A	Preset percentage: Operational Grades 3-8 ELA: 6% days 1 & 2, 4% day 3 Operational Grades 3-8 Math: 3% days 1 & 2, 2% day 3				
	Items requiring validity	Operational Grades 3-8 ELA & Math				
	Notes:	Please refer to comparison chart of scoring terminology and practices as applied by 0 Pearson.		oring terminology and practices as applied by Cognia and by		
Voiding	Threshold:	Grade HS ELA, Math, Sci and Grades 5 & 8 Sci: <70% based on daily Comp. Report Grades 3-8 ELA and Math: <70% based on cumulative validity performance				
	Frequency:	Daily				
	Threshold for scorer removal:	At the discretion of Sco	oring Leadership	)		
Equating Items						
	<b>Note:</b> To ensure scorer of polytomous.	consistency, seeded pape	ers will be inserte	ed for all equating items that are		

	Table 4 - Staffing Requirements:					
Staffing Level:	Minimum Education Requirements:	Specific Degree Requirements:				
Scorer	Grades 3-8:	Must be at least 18 years of age.     Cannot be under contract to Massachusetts schools, including as teachers, administrations, and paraprofessionals.				
Scoring Team Leader	Grades 3-8:	Must be at least 18 years of age.     Cannot be under contract to Massachusetts schools, including as teachers, administrations, and paraprofessionals.				
Scoring Supervisor	Grades 3-8:	Must be at least 18 years of age.     Cannot be under contract to Massachusetts schools, including as teachers, administrations, and paraprofessionals.				

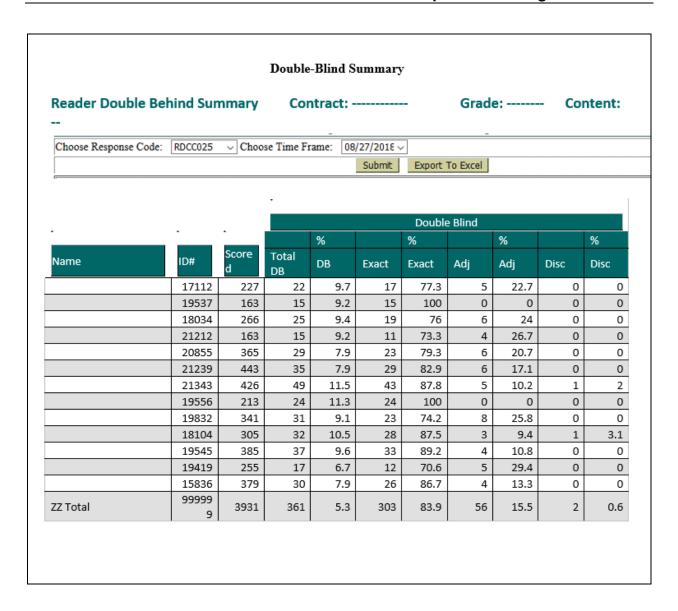


		Table 5 - Scoring	Platform Additional Set-up		
Al Scoring	☐ Yes (1st score)				
	⊠ Yes (2 <sup>nd</sup> score)	by DESE.			
Arbitration Rules	☐ Adjacent		ere is a discrepancy between two scores assigne		
	□ Discrepant		e than one point apart), the response is automationse and provides an arbitration score.	cally routed to scoring leadership who	
Practice Set within iScore / ePen	□ Yes 図 N/A		CAS practice sets are an integral part of scorer tr onse, revealing the actual score and explaining th		
Score-of-Record Rules		l			
	Scor	Source	Resolutio	on	
	Arbitration score and provided	read-behind score both	Latest read-behind score is the score-of-record	d	
	Arbitration score (no	read-behind performed)	Arbitration score is the score-of-record		
	Two read-behind sco adjacent (and therefo performed)		Grades 3-8: If Read-behind score is provided by 2 STLs, th score-of-record Grade HS: If Read-behind score is provided by 2 STLs, th score-of-record		
	Two read-behind sco (and resolution is pe	res that are discrepant formed)	If Read-behind score is provided by 2 STLs and they are discrepant (more than one score point apart), then resolution is required. The resolution score (Edit Score) is the score-of-record.		
	One read-behind sco	re	Read-Behind score is the score-of-record		
	Two Scores (provide double-blind process	•	Grades 3-8:  If the first score and second score are adjacen score is the score-of-record (Cognia setting) Grade HS:  If the first score and the second score are adjacen higher score is the score-of-record (MCAS set	cent (differ by one score point), the	
		Con	dition Codes		
☑ Crisis (41)					
⊠ Blank (B-21)		able (U-51)	☑ Wrong Location (W-52)	Non-English (F-53)	
☑ Off Topic (O-54)	☐ Illegib	e (I-55)	☐ Quarantine (Q-56)	☐ Insufficient Amount to Score (A-57)	
☐ Refusal to Score (R-58)	⊠ Repea	s the Prompt (P-59)	☐ Typed Sheet/NSR (T-60)	□Escalate (61)	
☐ No Score (N-62)					
NOTE: Science LEGACY us  Defining information of cond					
Demining innormation of cond	ilion codes can be lound in	rail D, Section 3.2			

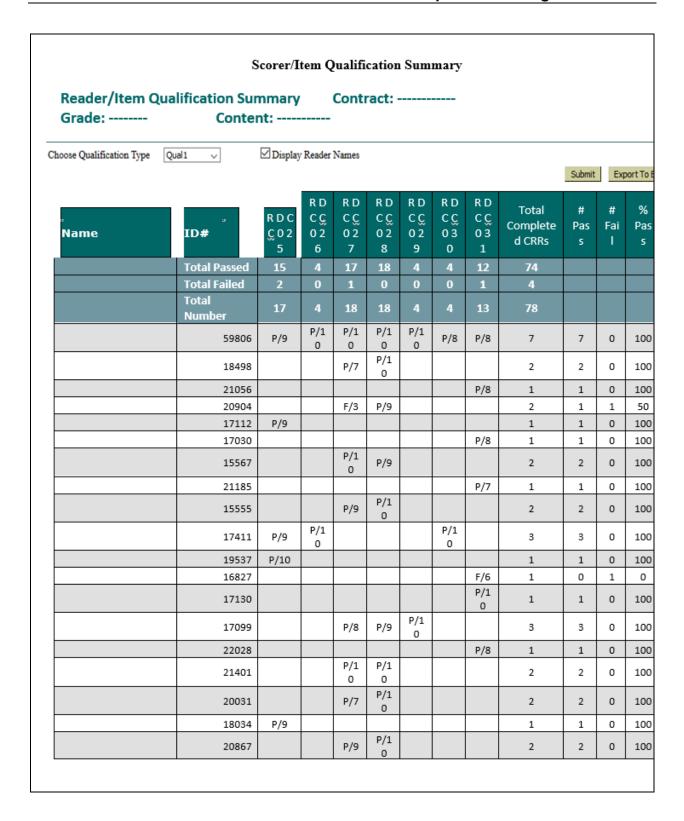




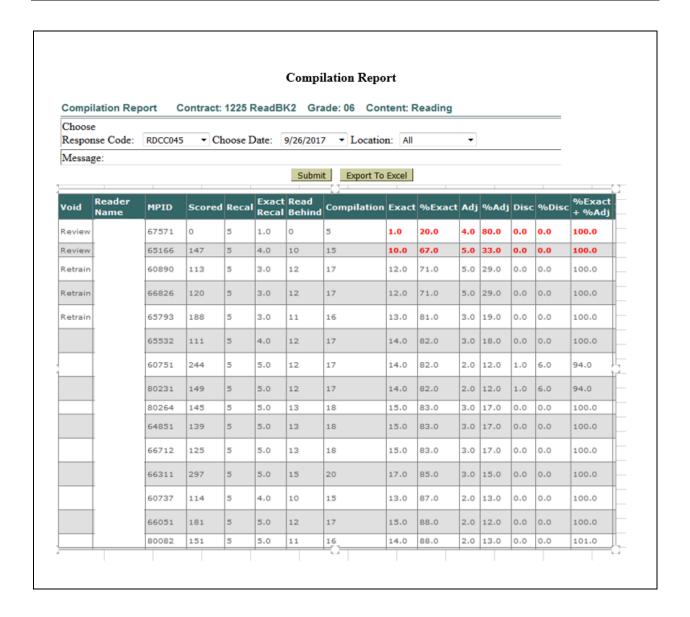












### **Pearson Al Scoring Process**

For training the Intelligent Essay Assessor (IEA) engine with a set of human-scored responses, typically 2/3 of the responses are used to train the engine and the remaining 1/3 are held out to evaluate performance. The MCAS models will be trained using 2023 field test prompts, responses, and human-scored data. Within that data, approximately 2,000 responses per prompt received a first human score and 20% received a blind 2nd score. IEA will be trained on ~1,300 responses per trait, randomly selected to represent the operational distribution. Once the engine is trained, the models will be evaluated based on the remaining ~650 responses per trait.

The data will be evaluated based on the industry-standard criteria for automated scoring shown in the table below.

Measure	Threshold
Pearson R QWK	>= 0.70
Карра	>= 0.40
Exact Agreement	>= 65% (or greater than Human-Human)
By Score Point Agreement	>=50% (or greater than Human-Human)
SMD	Within  0.15

Training results will be presented to DESE for approval. Based on the engine training, IEA may recommend re-evaluating during the June human scoring window as more operational responses are available to supplement the IEA engine.

#### **Pearson Human Resource Recruitment Overview**

Pearson will recruit diverse professional individuals with experience and educational backgrounds that meet all contractual requirements. The Pearson School Assessments Human Resource business partners will ensure hiring of qualified and diverse individuals to fill scoring positions so that the workplace is equally represented with various experiences and skills.

All employees must undergo degree verification and criminal background checks. Pearson prioritizes previous hires to receive offers.

All employees will complete onboarding tasks including the latest Pearson Code of Conduct, Employee Handbook, and the technical requirements of their project. Candidates will be asked to sign and complete a confidentiality form. Employees must sign and agree to the terms as a requirement of employment.

Pearson will ensure completion of all onboarding tasks for each employee prior to their project start date. Notifications will be sent from Human Resources to remind individuals of any open tasks. Hiring records that display a candidate's status in the project will be provided to stakeholders on a regular basis.

Personal Information Guidelines are managed through a controlled document. Data is stored within the Human Resource system and requires secure access.

#### After converting this document to PDF, please insert here:

- Cognia-Pearson Scoring Terminology (page 13 & 14)
- MCAS Scoring Survey (page 15)
- Pearson Confidentiality and Acknowledgement (Page 16 & 17)

## Part B:

## Cognia General Scoring Guidelines & Best Practices

### Contents

Pearson Al Scoring Process	12
Preface15	
Scoring Services Staffing	15
Pre-Scoring Logistics	15
Employee Recruitment	15
The Benchmarking Process	16
Operational Benchmarking	16
Field Test Benchmarking	16
Benchmarking vs. Rangefinding Meetings	16
Scorer Training	17
Process and Materials	17
Training Sequence	17
The Qualification Process	18
Consensus Scoring Approach	18
Scoring System	18
Overview	18
Condition Codes	19
Quality Control	20
Read-Behind Scoring	20
Double-Blind Scoring	20
Validity Responses	21
Recalibration Sets	21
Voiding Scorer Work	22
Crisis and Alert Responses	22
Scorer Monitoring Reports	22
Distributed Scoring	23
Cognia Facilities	23
Addendum: Non-Mutual Non-Disclosure Agreement	24

#### **Preface**

This document represents Cognisa's comprehensive best practices and standard operating procedures for evaluating and scoring student work. Procedures will be implemented depending on the specific requirements of each client. All client-related details and applicable contractual requirements are specified in Part A of this document: Client-Specific Scoring Guidelines.

#### **Scoring Services Staffing**

The following table summarizes key positions held by members of Scoring Services and describes their general responsibilities.

Position	Description
Director, Scoring Content & Quality	Oversees the all content-related deliverables of the Scoring Content Specialists and their respective Scoring Content Group Manager.
Director, Scoring Operations & Logistics	Oversees and coordinates the operations and logistics of all scoring activities, creates budgets, and establishes scoring schedules.
Project Managers – Scoring	Manage scoring-related activities, deliverables, and scheduling of tasks.
iScore Operations Manager	Maintains Cognia's scoring platform (iScore), manages other scoring systems as needed, and coordinates data deliverables between Scoring Services and Reporting team.
Scoring Operations Managers	Oversee scoring logistics, recruitment of contingent workforce, facility requirements and security.
Scoring Content Group Managers	Manage Scoring Content Specialists within content areas of ELA/Social Studies and Science/Math, oversee workflow processes, and ensure quality and production of scoring.
Scoring Content Specialists	Supervise the scoring of their respective content areas within their assigned contracts.  Responsibilities include finalizing the selection of all scoring training materials and facilitating benchmarking and rangefinding meetings. They also train and supervise scoring leadership and monitor the training and scoring of items for their assigned projects. Scoring Content Specialists have the overall responsibility of ensuring accurate and consistent scoring according to the approved client guidelines for their content area and contracts.
Scoring Supervisors	Scoring Supervisors work under the guidance of a Scoring Content Specialist. They are responsible for training assessment items and ensuring consistency across assigned grades, content, and assessment administrations. They also respond to questions during scorer training and throughout scoring and monitor the quality and production of ongoing scoring.
Scoring Team Leader (STL)	Scoring Team Leaders work under the supervision of Scoring Supervisors and lead a small group of scorers. STLs are responsible for quality control by performing read-behinds and providing coaching as needed.
Scorers	Scorers review, evaluate, and assign scores to student work based on client-specific scoring standards.

#### **Pre-Scoring Logistics**

#### **Employee Recruitment**

Cognia HR and its staffing partners are responsible for the recruitment of all scoring personnel. Cognia seeks to employ scoring staff with a wide range of educational backgrounds and professional experience. Cognia will recruit individuals who meet or exceed the contract-specific requirements to fill scorer and scoring leadership positions. All scoring associates are vetted for appropriate educational requirements through collection and review of their post-secondary transcripts. Candidates with backgrounds in education are also noted during this process. Depending on client preferences, Cognia will seek to customize the recruitment effort by including some or excluding all scoring associates from the client state. Potential associates must submit documentation, including transcripts and resumes, to



verify employment eligibility. Prior to hiring, all associates are advised of the scoring systems' minimum technical requirements.

If hired, all scoring associates will be required to sign and abide by a non-disclosure/confidentiality agreement which emphasizes the confidential and proprietary nature of all work and materials associated with all scoring activities. (see Attachment)

After hiring and before the onset of each scoring event, information on demographics and educational background will be collected again as additional employment verification measure. Further contractual specifics related to scoring associates' educational backgrounds are detailed in Part A of this document: Client-Specific Scoring Guidelines.

#### The Benchmarking Process

#### **Operational Benchmarking**

This activity occurs after operational administration of an assessment and prior to scoring it. It typically involves identifying additional suitable student responses (either from the pool of FT responses or from the pool of available OP responses to an item) in order to supplement existing scoring materials or to populate additional training or quality control materials.

#### Field Test Benchmarking

The activity of benchmarking occurs after administration of a Field Test and prior to scoring a Field Test. To prepare for benchmarking, scoring leadership review the assessment item and any associated stimuli, the scoring rubric, and scoring notes (when available). All students completed the assessment, their responses are loaded into the scoring system. Scoring leadership will log into the scoring system and start viewing student responses. After becoming familiar with both the assessment item and the student responses, scoring leadership will start assigning preliminary scores to appropriate responses and submit them to a separate folder in the scoring system. Within that folder, benchmarking staff can designate responses to specific sets of responses depending on the most appropriate use, e.g., anchor set, practice set, qualification set(s), or an extra set which stores responses for potential substitutions or for the assembly of supplemental training materials. Once the sets are created and reviewed, the benchmarking process for each field test item is completed and the item is ready for either benchmarking meetings or rangefinding meetings.

#### Benchmarking vs. Rangefinding Meetings

A difference between benchmarking and rangefinding meetings are the participating key stakeholders and the associated meeting facilitation. Key stakeholders in benchmarking meetings are representatives from Scoring Services, Content Development, and State Education Agency (SEA) content staff. In addition, rangefinding meetings also include participation by educators.

In a benchmarking meeting, it is the SEA content staff who define the scoring parameters for an item and they sign off on core training materials. The meeting itself is an open-forum discussion during which all meeting participants discuss how responses fare against the scoring rubric. While the goal is that all meeting participants agree on the scores after thorough discussions, it is the SEA content staff who have the final say and give final approval of the scores for all reviewed student responses.

In a rangefinding meeting, educators are the ones who provide the interpretive framework of the scoring standards. While the entire group (Scoring Services, Content Development, SEA, educators) reviews a



body of student work, it is the educators who are tasked with reaching consensus on the score(s) they assign to each reviewed response. In doing so, educators interpret the scoring rubric and thereby define the range of each score point level of the scoring rubric by consensus-scoring student work associated with an item.

The details as provided in Part A: Client-Specific Scoring Guidelines will outline the applicable meeting forum.

#### **Scorer Training**

#### **Process and Materials**

Scorer training will begin with an introduction to scoring and an overview of the assessment program. This could include the purpose and goal of the assessment program, any specific characteristics of the test and/or the testing population. There will also be a general discussion about the security, confidentiality, and proprietary nature of the assessment, all scoring materials, and Cognia's scoring procedures.

Training materials will be available to scorers during scoring and may include:

- Student prompt and associated stimuli
- Scoring rubric
- o Item sample response and training notes (when provided by content development team)
- Anchor Set
  - Clear examples that include mid-range student responses at each score point (when available)
  - Presented in score point order
- Practice Set
  - May include student work that demonstrates the cut-points between adjacent score points and/or atypical responses
  - May include examples of all score points (when available)
  - Presented in random order
  - Scorer accuracy can be captured and reported
  - Scoring Supervisor will review each practice set response (if required)

#### **Training Sequence**

A Scoring Content Specialist or Scoring Supervisor will lead the training for each item. Training may occur through a recorded, interactive training module, or through an online training system. Regardless of the method of training, the approach will follow this sequence:

- 1. Review of the student prompt, associated stimuli, the scoring rubric, associated sample responses, and training notes
- 2. Review of the anchor set
- 3. Analysis and discussion of each anchor response, its assigned score and associated, detailed scoring rationale
- 4. Scoring of responses in the practice set(s) to be scored independently to replicate the actual scoring process
- 5. Discussion of each practice response, revealing the actual score assigned to the student response and explaining the scoring rationale
- 6. Methodical review of all scoring criteria while paying particular attention to the fine lines that determine the cut-points between adjacent score points



- 7. Question and answer segment addressing any remaining scorer questions
- 8. Administration of a client-specific number of qualification sets, each consisting of 10 pre-scored responses, scored independently, and deployed randomly to each scorer
- 9. Review of qualification results after each set before scorers are admitted to subsequent qualification set(s)
- 10. Start scoring live student responses

#### The Qualification Process

Qualification sets are used to ensure that scorers have successfully internalized the scoring standards before they begin scoring each item. General qualification guidelines <u>for operational items</u> are:

- Each qualification set will contain 10 responses.
- The number of qualification sets administered are client-specific. Typically, operational items contain two qualifying sets to provide a second opportunity after retraining.
- Qualification sets are administered through Cognia's proprietary iScore system or another compatible scoring system. Responses are distributed to the scorers unscored and in random order.
- In order to qualify, scorers are required to meet the passing threshold as determined by the client and as specified in Part A: Client-Specific Scoring Guidelines
- Scorers who do not pass qualification will not be allowed to score the item. They will
  either be trained on a different item or dismissed from the scoring project.
- Responses included in the qualification set must be approved for use by the Scoring Content Specialist or Assistant Scoring Content Specialist. Depending on clientspecifications, responses may also have to be approved by the client and/or be part of materials approved in a range-finding or benchmarking meeting.

#### Note:

Scoring Team Leaders receive the same training and undergo the same qualification process as scorers. However, STLs may be trained on some or all items in advance during a separate leadership training. This provides an additional opportunity to absorb the training materials and it prepares them to fulfill their role during scorer qualification.

#### Consensus Scoring Approach

When the total number of student responses received is small, Cognia may recommend applying the consensus scoring approach. In this approach, a select group of highly experienced scorers will train and qualify on each item and then proceed by scoring the small number of student responses together in pairs, working side-by-side, and discussing each response to reach a consensus score. Using this approach, scorers are constantly calibrating with each other to provide accurate and consistent scoring for the small number of student responses. When the consensus scoring approach is used, quality control tools designed for high n-counts of student responses are not applicable.

#### **Scoring System**

#### Overview

The scoring of student responses will be conducted through Cognia's iScore or another compatible scoring system which displays images that are received through data transfer from the online computer-based testing platform or through scanned images of paper-based tests. In instances of rendering



issues with any paper-based test books, scoring will occur by referring to the actual test book and the scores will be manually entered into the scoring system.

The scoring system does not display any student or school identifiable information. Security is maintained during scoring through a highly secure server-to-server interface. It ensures that images are only accessible to those who will be scoring each item or to scoring management. All responses are tracked through a unique booklet code that is matched to the student records during data processing.

Each scoring day scorers are asked to review the anchor materials and the rubric of an ongoing item. There will also be a broader group refresher upon resumption of scoring following a recess (e.g., a weekend or disruption of delivery). Each scoring day typically concludes with a debrief meeting with the Scoring Content Specialist, the Scoring Supervisors, and, if desired, client staff members to recap the day and address any issues that may need resolution.

During the course of scoring, scorers may encounter student responses that indicate the possibility of cheating or some type of testing irregularity. Scorers will score this type of student response based on its own merits and then refer it to the Scoring Content Specialist and Project Manager for further processing and client notification. Any potential score change request by the client can be made prior to final reporting.

#### **Condition Codes**

Scoring Services makes every attempt to score each student response. However, when a response does not conform to the score point parameters as defined in the scoring rubric, condition codes can be employed. Responses that are flagged will receive a numeric score but will undergo supervisory review. Responses that are rejected will not receive a numeric score but will receive a second read.

#### Flags:

Crisis: Response indicates that a student may present a danger to themselves or others, the
student or another child is in danger, there are indications of sexual or physical abuse, or other
specific criteria as specified by the client. (Please refer to section 7 for the handling process)

#### Rejects:

- Blank: No deliberate marks in the answer space
- **Unreadable:** A rendering issue or obstructed student response
- Wrong Location: A clearly legitimate response to another item on the assessment
- Insufficient Amount to Score: The response contains an insufficient amount of student work to score
- **Illegible**: Tiny or poor handwriting (for PBT), spelling that cannot be deciphered, or other conditions that render the student work indecipherable
- Refusal: The response clearly indicates a refusal on the part of the student to address the prompt or participate in the assessment
- Repeats the Prompt: The response copies the prompt or portions of it and offers no attempt to respond to the task/prompt
- No Score: Any other circumstance (as defined by the client) that prevents the assignment of a numeric score
- Non-English: The response is written in a language other than English (or in a Spanish
  assessment in a language other than Spanish) or is a mix of English (Spanish) and another
  language but lacks sufficient English (Spanish) to provide a score.



Responses that are identified as Unreadable or Wrong Location undergo a separate resolution process. They will be routed to the Scoring Content Specialist or Scoring Supervisor. Responses will be reviewed, and the appropriate score assigned. Furthermore:

- Unreadable responses (PBT only) will be reviewed by consulting the student's original test booklet or by requesting a re-scan of the student work. If the response can be read through either method, the appropriate score will be assigned. Completely unreadable responses will not receive a numeric score.
- Wrong Location responses (PBT only) will be reviewed by a Scoring Supervisor or Scoring
  Content Specialist. Their broader access to the scoring system allows them to review all
  student work and assign the appropriate score for each response. Wrong locations can only be
  scored when the student was evidently attempting to respond to another item on the
  assessment.

#### **Quality Control**

Note: not all quality control measures listed in this section are applicable to every client contract.

While all scorers must first train and qualify to gain access to scoring student work, they must also maintain acceptable levels of accuracy to continue scoring. The scoring system provides the opportunity to employ multiple quality control tools in order to monitor accuracy and consistency throughout scoring.

Depending on client specifications, STLs may also score responses each day. In doing so, they are also subject to all quality control tools and statistics. While in a scoring capacity, the Scoring Supervisor or Scoring Content Specialist will conduct read-behinds on STLs. STLs may also encounter validity papers during their course of scoring.

#### **Read-Behind Scoring**

Read-behind scoring allows the STLs and Scoring Supervisors to monitor the performance of each scorer. It provides an immediate real-time snapshot of a scorer's accuracy and the opportunity to provide individualized coaching or re-training as needed.

Read-behinds are generated in the scoring system at the request of the STL. Scorers are not aware which responses are designated for read-behinds. Cognia's scoring platform allows for blind scoring of read-behinds. The STL conducts each read-behind without prior knowledge of the assigned score. After the STLs submit their score, they can reveal the score assigned by the scorer and provide counseling as needed.

The number of read-behinds conducted per scorer will vary and STLs will focus their attention on scorers as needed. Conducting read-behinds is an ongoing process throughout the day. STLs will conduct more read-behinds on scorers who are at the lower threshold of accuracy and require counseling. Cognia will adhere to contract requirements as outlined in Part A.

To further ensure the accuracy of the STLs, scoring leadership has the ability to review their readbehind work. The Scoring Supervisor has access to all responses that were reviewed and may compare scores to verify the accuracy and consistency of scoring.

#### **Double-Blind Scoring**



While read-behinds measure scorer accuracy in relationship to leadership, double-blind scoring provides statistics on scorer-to-scorer agreement, or inter-rater reliability. Double-blind scoring is the practice that refers to a method whereby the same response is routed to two scorers. The response is independently and anonymously reviewed by each scorer. In double-blind scoring, scorers do not know which response will be (or already has been) scored by another randomly selected scorer.

#### **Validity Responses**

The deployment of validity responses can provide an additional opportunity to compare and monitor the quality of scoring. The process is set up to meet the following criteria:

- Validity responses are identified from a pool of responses and pre-scored according to the scoring standards as expressed in the anchor set and the scoring rubric
- Pre-scored validity responses are loaded into the live scoring queue
- Validity responses look identical to live student responses such that scorers can't tell the difference between the two
- Validity responses can be launched at any time during the scoring project
- The insertion rate of validity responses is fully customizable in the scoring platform. Please refer to the Client-Specific Scoring Guidelines in Part A of this document. Scoring leadership may select validity responses either from recently scored responses, unscored responses, rangefinding meeting materials, or they may use previously administered validity responses for the item. In order to qualify as a validity response, it must be approved for use by the Scoring Content Specialist or other designated leadership staff. Depending on contract specifics, validity papers may also either be part of the approved rangefinding set or be approved by the client.

#### Recalibration Sets

Another option in Cognia's suite of quality control measures is the administration of recalibration sets. Beginning on the second day of scoring an item, scorers will take a recalibration set prior to starting scoring to ensure they remain calibrated to the scoring standards. Recalibration sets consist of prescored responses. Recalibration sets will include a variety of score points, but they will not always include an example of each score point.

Recalibration sets reinforce the scoring decisions of the training materials and prevents scorer drift throughout the project. Scorers who demonstrate continued understanding of the scoring standard will be allowed to start scoring for the day. Scorers who struggle with the recalibration responses will review them with scoring leadership, comparing the responses to the Anchor Set responses and the scoring rubric. Once the review is complete, scoring leadership will determine whether the scorer may begin scoring the item for that day.

Scoring leadership may select recalibration responses from recently scored responses, unscored responses, rangefinding meeting materials, or they may use previously administered recalibration responses for the item. In order to qualify as a recalibration response, it must be approved for use by the Scoring Content Specialist or other designated leadership staff. Depending on contract specifics, recalibration papers may also either be part of the approved rangefinding set or be approved by the client.



#### **Voiding Scorer Work**

When scorers meet or exceed accuracy standards, they will continue to have access to student responses and may continue to score. If scorers fall below the established accuracy threshold, they will be retrained and Scoring leadership will determine whether a scorer is allowed to resume scoring.

The scoring system allows Cognia to void a scorer's work. If a scorer fails to maintain accuracy standards, his or her work for the impacted time frame will be invalidated, and the affected student responses will be routed to other qualified scorers for re-scoring.

#### **Crisis and Alert Responses**

Scorers are trained to identify crisis or alert responses. These include responses which indicate that a student may present a danger to themselves or others, the student or another child is in danger, there are indications of sexual or physical abuse, and/or other criteria as specified by the client.

As soon as a crisis or alert response is identified, the Scoring Content Specialist will notify the Scoring Project Manager who may reach out to the Program Manager. Student demographic information and copies of the student response are posted to designated client staff members.

#### **Scorer Monitoring Reports**

To monitor the accuracy, consistency, and pace of scoring, the scoring system generates a variety of reports to allow scoring leadership to monitor all aspects of a complex assessment program. These reports show both the overall performance of the scoring project as well as immediate and real-time scorer level data and provide the opportunity to monitor an individual, the group, and the overall project.

STLs and Scoring Supervisors have access to a select number of reports which aids them in monitoring and ensuring quality scoring. Scoring Content Specialists and scoring management have access to all quality and production reports in the scoring system. Clients will also have access to a variety of quality and production reports in the scoring system, including interpretive guides, when applicable.

The following is a summary of the most commonly used reports in iScore, Cognia's proprietary scoring system:

- The Read-Behind Summary Report shows the total number of read-behind responses
  conducted per scorer and shows the number and percentage of responses that were in
  exact, adjacent, and discrepant agreement between the scorer and the STL. The report
  also provides an overall statistical summary of all scorers working on the item. The report
  has both a daily and a cumulative option.
- The **Double-Blind Summary Report** shows the total number of double-blind responses read by a scorer and will note the number and percentages of exact, adjacent, and discrepant scores. The report also provides an overall statistical summary of all scorers working on the item. The report has both a daily and cumulative option.
- The Daily Embedded Summary Report shows the total number of validity responses read by a scorer and will note the number and percentages of exact, adjacent, and discrepant scores.
- The Qualification Statistics Report lists each scorer by name and ID#, identifies which
  qualification sets each scorer has taken and the respective pass or fail status for each
  set.



- The Summary Report shows each item and the total number of student responses to be scored for each item. During ongoing scoring, it also shows the number of responses that have already been scored for each item and the number of double-blind scores provided.
- The **Score Point Distribution Report** shows the total number of student responses per assigned score point. The report offers both a daily and a cumulative option.
- The Compilation Report shows, for each scorer, the total number of responses scored, the number of read-behind responses and the number of scored recalibration responses (both individually and combined), and the percentage of exact, adjacent, and discrepant scores assigned in comparison to read-behinds and recalibration responses.

#### **Distributed Scoring**

Cognia has implemented a distributed scoring model that provides our clients with accurate, reliable, and timely results. Our distributed scoring model adheres to the same requirements as Cognia's center-based scoring model. The following security features are implemented to support the secure nature of distributed scoring:

- Two-Factor Authentication login protocol which prevents unauthorized users from gaining access to the scoring system and materials.
- The scoring system and materials are housed within a secure scoring kiosk which disables any print and download functions.

The communication process between scoring leadership and scorers is managed via a communication tool (e.g., Zoom, MS Teams, Skype) to support regular face-to-face check-ins. All scoring associates are required to utilize a webcam to maintain direct communication and facilitate positive identification.

#### Cognia Facilities

Cognia currently maintains facilities in Portsmouth, NH; and Alpharetta, GA. Cognia reserves the right to decide on the appropriateness of their utilization depending on any potentially existing health risks to its employees and/or the suitability for use of these facilities.

These facilities are locked, and admission is limited to authorized staff. Access is monitored by a security system that only admits staff with an electronic access card. This card also serves as Cognia identification card which must be worn at all times while in the building.



#### Addendum: Non-Mutual Non-Disclosure Agreement

This Confidentiality and Non-Disclosure ("Agreement") is made on **«Effective\_Date»**, by and between Cognia<sup>™</sup>, Inc., with a physical address of 9115 Westside Parkway, Alpharetta, Georgia 30009, a 501(c)(3) non-profit organization incorporated under the laws of the State of Georgia, United States of America, and **«Name»**, with a principal address of **«Address1»**, **«City»**, **«State» «PostalCode»**, and taken together, known as ("the Parties").

WHEREAS, "Name" intends to offer services such as but not limited to; scoring and/or distributed scoring for Cognia through a temporary agency service arrangement with such services performed either in facilities arranged by Cognia or location(s) identified by temporary agency agreement with "Name" (the "Transaction"); and

WHEREAS, the Parties may disclose certain confidential and proprietary information to each other for the purpose of evaluating the Transaction, and the Parties mutually agree to enter into a confidential relationship with respect to the disclosure by one or each (the "Disclosing Party") to the other (the "Recipient") of such proprietary and confidential information; and

NOW, THEREFORE, the Parties, intending to be legally bound, agree as follows:

Definition of Confidential Information. For purposes of this Agreement, "Confidential Information" means (1) any and all information, data, design, memoranda, models, prototypes, equipment and/or other material, of a confidential, non-public or proprietary nature, including, without limitation, information relating to or regarding the products or services developed or being developed by the Disclosing Party. information regarding intellectual property (including ideas that may be subject to patent, trade mark, service mark or trade secret protection) and other rights, techniques, research, development, samples, marketing, sales, know-how, operations, distribution, strategy, services, applications, promotions, advertising, costs, prices, business plans, financial statements, software, source code, and firmware and process information and such information relating to the Disclosing Party's existing and prospective invention, business partners, and customers, (2) documents and information that are marked or designated with a word or symbol indicating that the document or information should be considered confidential, such as "Confidential", "Proprietary", or "Privileged", (3) documents and information that the Disclosing Party informs the Recipient, either in writing or orally, are confidential, and (4) information that is a trade secret or the confidential or proprietary information of a third party, which is obtained from the Disclosing Party, irrespective of whether it is in tangible or intangible form, irrespective of whether it was communicated orally, in writing or on any other record bearing media and irrespective of whether it was marked or designated as confidential in connection with the disclosure.

Notwithstanding the foregoing, the term "Confidential Information" does not include information which: was in the public domain prior to the Recipient's receipt of same from the Disclosing Party, or which subsequently becomes part of the public domain by publication or otherwise, other than by the wrongful act of the Recipient; information which the Recipient can show by reasonable proof was in its possession prior to the Recipient's receipt of same from the Disclosing Party and which was not acquired directly or indirectly from the Disclosing Party; information which is independently developed by the Recipient without reference to or reliance upon the Confidential Information of the disclosing party and without breach of this Agreement; or that the Parties agree in writing is not proprietary or confidential.

<u>Confidentiality</u>. Recipient agrees to treat as confidential all Confidential Information provided to it by Disclosing Party or Disclosing Party's representatives, whether disclosed before or after the date of this Agreement. In no event, including the breach of this Agreement or any other agreement between the Parties, shall either Party allow the disclosure of any Confidential Information disclosed to it by the Disclosing Party except as permitted under the terms of this Agreement or with the prior written consent of the Disclosing Party. The Parties shall take commercially reasonable steps to prevent the unauthorized disclosure, use, dissemination, or publication of the Confidential Information and shall protect such Confidential Information to the same extent that it protects its own confidential and



proprietary information, but in no event using less than a reasonable standard of care. This Agreement shall be binding on all directors, officers, stockholders, members, managers, employees, agents, representatives, successors and assigns of the Recipient (collectively, "Agents"), and Recipient shall take commercially reasonable steps to assure that its Agents to whom Confidential Information is disclosed maintain the confidential nature of the Confidential Information. Recipient shall immediately notify the Disclosing Party upon discovery of any loss or unauthorized disclosure of the Confidential Information of the Disclosing Party.

<u>Use</u>. Recipient agrees that the Confidential Information shall be used solely for purposes of the Transaction and in connection with any transaction entered into by the Parties. Recipient shall not disclose any Confidential Information to any other party. Recipient further agrees that it is prohibited from using the Confidential Information for its competitive advantage, or to further its own business, professional or economic position. Neither the execution of this Agreement nor the transmission of any Confidential Information by the Disclosing Party to the Recipient shall constitute a conveyance or transfer to the Recipient of any right, title, interest or license in the Confidential Information.

<u>Term</u>. This Agreement shall be in effect for a period of three (3) years from the latter-dated signature below. The obligations contained herein shall survive until the earlier of (a) an exception to what is Confidential Information set forth in Section 1 is met, or (b) one (1) year after the expiration of this Agreement; provided, however, each Party's trade secrets shall be subject to those obligations herein and survive until they are no longer a trade secret.

Remedies. Because of the unique nature of the Confidential Information, Recipient agrees that breach of this Agreement will result in the irreparable harm to the Disclosing Party. Therefore, in addition to any and all other remedies available at law or in equity, the Disclosing Party shall be entitled to injunctive or equivalent relief enjoining the breach of this Agreement, without the necessity of posting bond or other surety. In the event of a breach of this Agreement by the Recipient, the Recipient agrees to pay reasonable fees incurred by the Disclosing Party to protect its rights under this Agreement including, without limitation, attorneys' fees and other costs to bring any lawsuit, action, or proceeding necessary to protect the Disclosing Party's rights. These remedies in addition to any rights by temporary agency related to employment law or dismissal for cause.

Governing Law; Venue. This Agreement shall be governed, interpreted, and/or construed in accordance with the laws of the State of Georgia without giving effect to choice of laws principles that require the application of the law, regulation or rule of a different state. Recipient and Disclosing Party hereby agree that any legal proceeding involving a dispute between Disclosing Party and Recipient concerning any aspect of this Agreement shall be brought solely in a State court located within the State of Georgia or the United States District Court for Georgia.

Return or Destruction of Confidential Information. After the performance of the services relating to the Transaction, Recipient agrees to destroy all Confidential Information and all documents containing Confidential Information Securely or Return to Cognia all Confidential Information held in the parties' position immediately (including any copies, notes, or abstracts, in any media).

<u>Amendment and Assignment</u>. This Agreement may be amended only upon mutual written agreement by the Disclosing Party and the Recipient. This Agreement and the rights and obligations contained herein are not assignable. Nothing in this Agreement obligates the parties to enter into the Transaction.

Severability. In case any provisions (or portions thereof) contained in this Agreement shall, for any reason, be held invalid, illegal or unenforceable in any respect, such invalidity, illegality or unenforceability shall not affect the other provisions of this Agreement, and this Agreement shall be construed as if such invalid, illegal or unenforceable provision had never been contained herein. If, moreover, any one or more of the provisions contained in this Agreement shall for any reason be held to be excessively broad as to duration, geographical scope, activity or subject, it shall be construed by limiting and reducing it, so as to be enforceable to the extent compatible with the applicable law as it shall then appear.



<u>Notices</u>. All notices or reports or secure return of materials permitted or required under this Agreement will be in writing and will be delivered by electronic mail or by certified or registered mail, return receipt requested, and will be deemed given upon personal delivery, five (5) days after deposit in the mail, or upon acknowledgment of receipt of electronic transmission. Notices will be sent to the addresses set forth at the end of this Agreement or such other address as either Party may specify in writing.

<u>Entire Agreement</u>. This Agreement is the final, complete and exclusive agreement of the Parties with respect to the subject matters hereof and supersedes and merges all prior discussions between the Parties with respect to such matters.

<u>Counterparts; Signatures</u>. This Agreement may be executed by one party as identified in the first paragraph, which shall be deemed an original for all purposes and all of which will constitute a single instrument. Facsimile signatures shall be deemed original and binding signatures.

<u>Survival</u>. All duties and obligations with regard to the protection of Confidential Information shall survive any termination of the discussions relating to the Transaction.

Parties hereby accept the terms and obligations set forth in this Agreement.

IN WITNESS WHEREOF, the parties, intending to be legally bound, hereto have executed this Agreement made effective as of the day and year set forth above.

By:	«Name»
Signature:	
Print Name:	Click or tap here to enter text.
Title:	Click or tap here to enter text.
Date:	Click or tap here to enter text.

[Non-Mutual Confidentiality and Non-Disclosure Agreement – Signature Page]

Email Legal@cognia.org



#### ADDRESS FOR RETURN OF MATERIALS:

Cognia

9115 Westside Parkway

Alpharetta, GA 30009