Designing Proficiency-Oriented Performance Tasks for the 21st-Century Workplace Written Communication: An Evidence-Centered Design Approach

Ahmet Dursun[a]
Corresponding author
[a]*University of Chicago*
5811 S. Ellis Ave. Suite 211
Chicago, Illinois, USA, 60637
+1(773)-834-0503
adursun@uchicago.edu

Jennifer K. Morris[b]

[b]*British Side Educational Services*
Barbaros Bulvarı, No: 91
Beşiktaş, İstanbul, Turkey, 34353
jennifer.morris@britishside.com

Aylin Ünaldı[c]
[c]*University of Huddersfield*
School of Education and Professional Development
Lockside, Queensgate, Huddersfield HD1 3DH, United Kingdom
A.Unaldi@hud.ac.uk

Abstract

CEPA® Written Communication Assessment™ is a computer-based, criterion-referenced, and proficiency-oriented test designed to measure non-native speakers' workplace English proficiency in writing. In its creation, developers used the evidence-centered design (ECD) approach (Mislevy & Haertel, 2006) to define and operationalize the construct of writing in the context of modern workplace communication. The first phase of ECD, Domain Analysis, yielded a comprehensive definition of 21st-century workplace written communication, revisions to CEFR language-proficiency descriptors, and an adaptation of the GRASPS framework (Wiggins & McTighe, 2008). These discoveries served the Domain Modeling phase by defining proficiency-oriented performance task design patterns; after which, a Conceptual Assessment Framework could be constructed. Due to the essential expansive detailing of each phase, this paper focuses only on the first three 'assessment conceptualization' phases of ECD. That is, the discussion of the last two 'assessment operationalization' phases, the Assessment Implementation and Delivery, is beyond the scope of the current paper. In keeping with the notion that assessment as argument is a cornerstone of test validation (Kane, 2006; 2012), this study reports on the use of ECD's first three phases, conceiving of assessment practices as evidentiary arguments, to inform the design and development of the CEPA Written Communication Assessment.

Keywords: *Evidence-centered design, written workplace communication, workplace writing construct, proficiency-oriented tasks, performance-based tasks*

1. Introduction

On a daily basis, companies and organizations the world over face similar challenges when it comes to decisions of selecting, hiring, promoting, and placing their employees based on a range of their abilities and skill sets. Foreign language skills (in this case, proficiency in written English) are no exception, and the assessment process is not always as straightforward, nor as accurate, as simply reviewing formal education degrees, references letters, institutional scores, and cover letters. This type of evaluation is further complicated by the oscillating nature of a second language, specifically its progression or attrition over time, its embedded-ness in modern technology, and the sociolinguistic competencies (i.e. workplace *and* cross-cultural communication knowledge and strategies) which inevitably accompany it. Thus, retooled assessments are needed to reveal the language capabilities and needs of the candidates as they pertain to their immediate level of English within the functioning, communicative environment of a real-life, modern workplace.

These challenges prompted the realization of the CEPA® (Communicative English Proficiency Assessment) research and development group in 2016, which would establish the foundation for a new and enriched type of online English language assessment in Turkey for adults in the business world ("CEPA", 2019). This current study, which was and still is, the process of this working group, attempts to explain how ECD (Evidence-Centered Design), with its first three phases of Domain Analysis, Domain Modeling, and Conceptual Assessment Framework, guided the design and basic structure of the computer-based proficiency-oriented performance tasks of CEPA, specifically the writing assessment section. Furthermore, we argue that by following the conceptualization phases outlined in ECD, this study was able to identify, define and re-define the key multi-faceted and dynamic writing skills to be measured for 21_{st}-century workplace written communication.

2. CEPA® assessment and mandate-driven evidence-centered design

CEPA® has been designed in response to the mandate handed down by British Side Educational Services, a private entity, to develop a screening assessment that could measure a non-native English speaker's workplace communicative English proficiency, and subsequently, their ability to function in a real-world workplace. The need for this assessment was due to a niche market in Turkey with unmet demand. More specifically, the mandate required CEPA to be a criterion-referenced proficiency-oriented test that measures test takers' ability within language contexts and situations representative of a workplace environment, yet another concern and request of the target market. It also required that the tasks and their results should be aligned with the Common European Framework of Reference for Languages (CEFR) guidelines for more efficient interpretation of scoring by Human Resources departments. CEPA Written Communication Assessment™ is one of four stand-alone components in the overall CEPA. The mandate for the CEPA writing section, which is the focus of this paper, required that it measures, precisely, 21_{st}-century workplace written communication through performance-based tasks, as would be the real attributes of the work for any potential employee in the real-world workplace.

The group's initial goal was to create an assessment model that could address the specifications described in the mandate, and that could more efficiently and accurately assess a potential candidate's current communicative English language level in writing, using the most current theories and approaches in the assessment field. In attempting this task, it became clear that the nature of the challenge itself must first be addressed. How would the assessment model be developed? And why? As a result, in order to identify and define a construct for 21st-century workplace written communication and inform the process of task design, it was necessary to examine broader questions concerning the real-world demands of written English specifically, in the Turkish (non-native English speaking) workplace context.

- How and for what purpose is written English language being used (together with current technological developments) in the non-native English workplace, specifically Turkey, at the time of this investigation?
- How can the strategies and abilities called for by authentic business and socioculturally sensitive communication be characterized?
- How can the particular features of design, used in online assessments, aid in the assessment itself?
- How can online screening assessment tasks be designed to take into account the constructs to be measured and the reporting that should result?

Several avenues could have been taken to address these questions, but the direction of the research and the pilot assessments were made in view of the mandate for the assessment-design as aforementioned. Thus, a development path guided by evidence-centered design (ECD), as described by Mislevy, Steinberg, and Almond (2003) and Mislevy (2011), was selected because it could provide tangible concepts (i.e. English language writing skills in real-life 21st century workplaces in Turkey) and a high level of detail that could be used to inform the design decisions (i.e. which skills were in more or less demand and how those could be represented in assessment tasks). More importantly, ECD can help create a solid argument for validation of test score interpretation and use by establishing explicit links between design decisions and scores obtained from the test (Chapelle et al., 2018). For CEPA, the ECD process was used to define and operationalize the construct of writing in the context of workplace communication leading to the design and development of the CEPA Written Communication Assessment, the outline for which is further exhibited in Figure 1.

As shown in the figure, ECD consists of five layers. As the first stage in assessment design, domain analysis organizes the information that will provide the grounding for constructing assessments. It requires analyzing and gathering fundamental information from the target language use domain. For the CEPA project, one of the concrete goals in this phase was to identify and define a construct for 21st-century workplace written communication. The layer that follows is domain modeling in which the test developers "organize the information and relationships discovered in domain analysis into the shape of assessment arguments" (Mislevy & Haertel, 2006, p.8). For the CEPA project, this entailed the formation of task patterns (i.e. sets of task characteristics presented as arguments to represent the tasks in the target language use domain in the assessment) that could measure the construct defined in the domain analysis phase. The next layer is the conceptual assessment framework (CAF), which lays out the technical specifications (i.e. a blueprint) both for the complete assessment and

individual task types. For CEPA, this phase yielded test and task specifications based on the arguments created in the previous two phases. The assessment implementation layer "concerns constructing and preparing all of the operational elements specified in the CAF", including, "authoring tasks, producing test forms, finalizing rubrics or automated scoring rules" (Mislevy & Haertel, 2006, p. 20). The final layer is the assessment delivery, "where students interact with tasks, their performances are evaluated, and feedback and reports are produced" (Mislevy & Haertel, 2006, p. 20). For the CEPA project, these combined phases resulted in the creation of a tailored assessment engine that was able to deliver the tasks that were designed and developed while also operationalizing the complex nature of both the test and task specifications. The last two phases of ECD will not be covered in this paper, but they are still presented in this section to outline the whole ECD process and to understand how CEPA, through ECD, theoretically progresses.
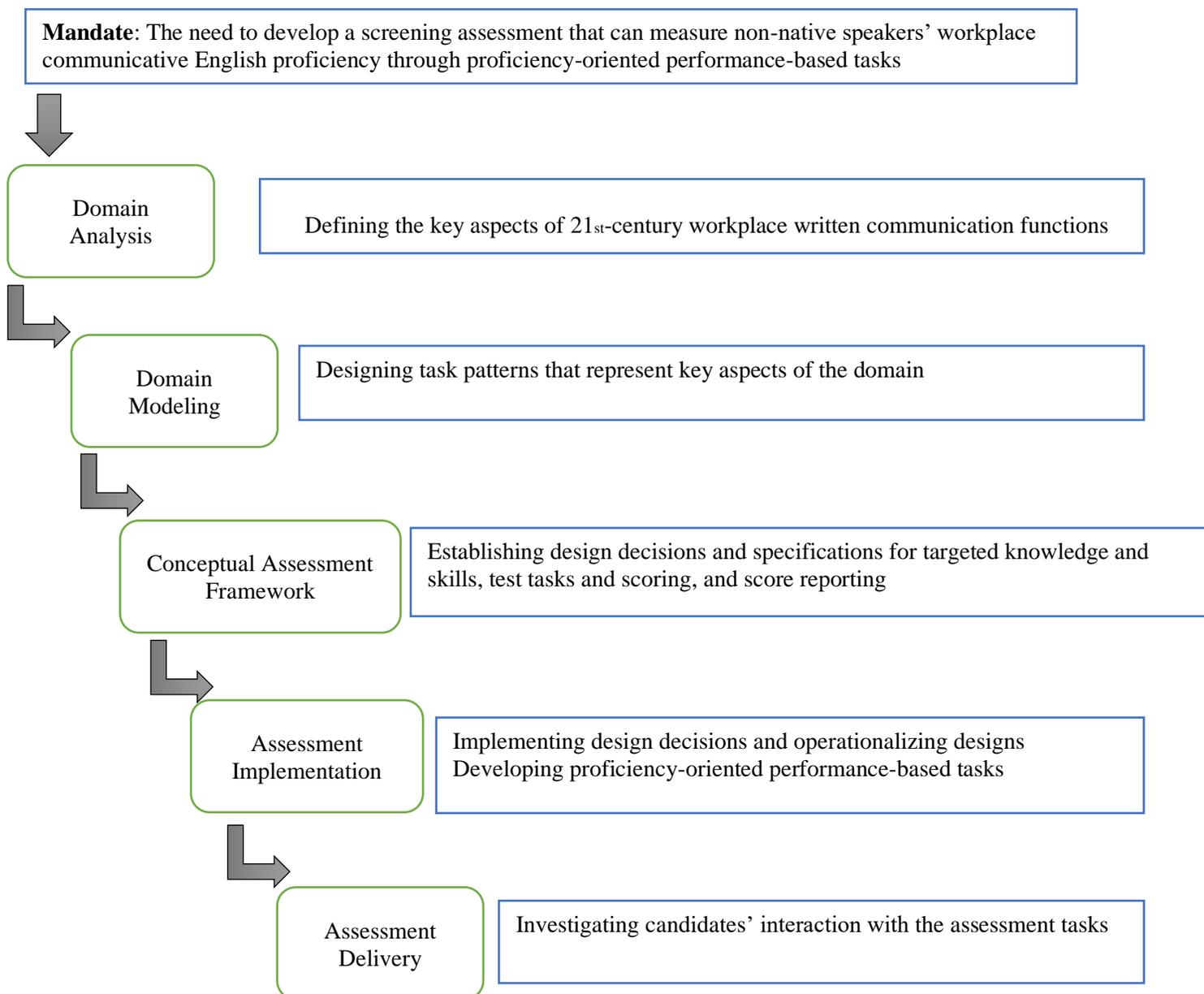
Mandate: The need to develop a screening assessment that can measure non-native speakers' workplace communicative English proficiency through proficiency-oriented performance-based tasks

Domain Analysis

Defining the key aspects of 21st-century workplace written communication functions

Domain Modeling

Designing task patterns that represent key aspects of the domain

Conceptual Assessment Framework

Establishing design decisions and specifications for targeted knowledge and skills, test tasks and scoring, and score reporting

Assessment Implementation

Implementing design decisions and operationalizing designs
Developing proficiency-oriented performance-based tasks

Assessment Delivery

Investigating candidates' interaction with the assessment tasks

Figure 1. ECD process as a framework to define and operationalize the construct of writing in the context of workplace communication

3. Domain analysis

As previously mentioned, the ECD framework, as the first phase, requires a detailed analysis of the target language use domain for which the test is developed, and the scores used. The identification and analysis of the various components in a given domain depends on the particular test development project (Chapelle et al., 2018). With this understanding, the components (i.e. inputs) that were investigated and analyzed for the CEPA Written Communication Assessment project are explained. Figure 2 shows an overall schematic diagram of the needs derived from the mandate (i.e. inputs to domain analysis), the components that were identified and investigated (i.e. domain analysis), and the results from domain analysis.

3.1. Inputs to domain analysis

The mandate indicated the need to develop a proficiency-oriented performance-based screening test to be used by the corporations, specifically in Turkey, for employability purposes. The mandate revealed three inputs for the domain analysis: (1) the need to screen candidates for corporations requiring workplace English proficiency for employability, (2) the need to define 21_{st}-century workplace written communication knowledge, skills, and abilities, and (3) the need to design performance-based assessment tasks.

The activities of the domain analysis were carried out through three relevant components that would each provide evidence of what it takes to communicate in a 21_{st}-century workplace environment. The first component was the stakeholders, including human resources staff, and other key personnel, both in domestic and international corporations in Turkey. These people were included because they were the ones who created the demand, and therefore it was important to investigate their responses as to how and why English written communication is *actually* being used in their workplaces. The second component was identified as professional workplace research reports as well as prominent popular business magazines so that we could determine the most recent and applicable discourse topics from a non-language-teaching perspective and gain a better understanding of what those in the business world value in written workplace communication. The third component was the business English/English for workplace communication textbooks and curricula. This final component was included because we wanted to capture the fundamental skills and knowledge being taught by, and the relevant learning materials provided in, major textbooks, which were in use in English language schools in Turkey and beyond. Next, we describe the research methodology for the domain analysis study, including the data collection and analysis tools that were used to investigate and analyze each of these components.

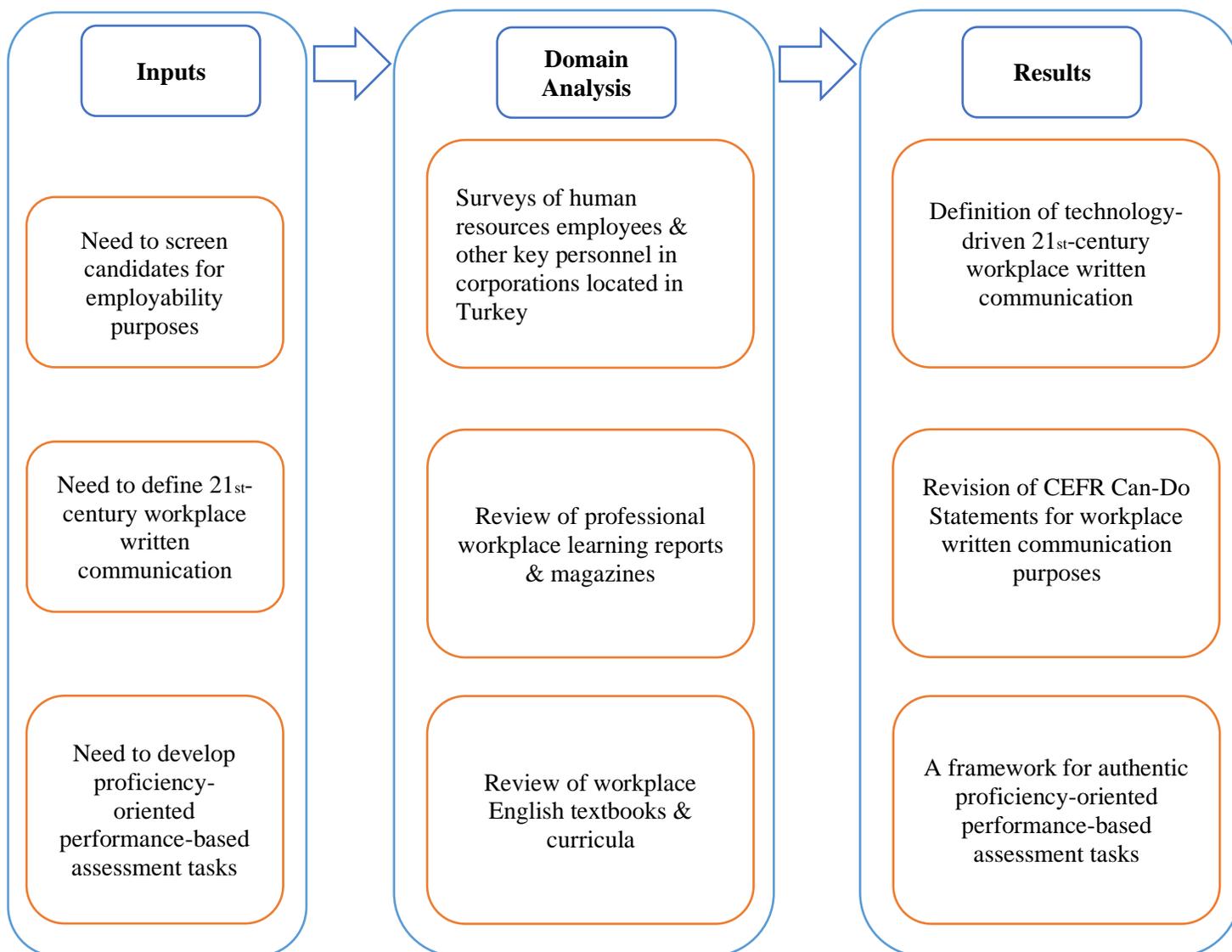| Inputs	Domain Analysis	Results
Need to screen candidates for employability purposes	Surveys of human resources employees & other key personnel in corporations located in Turkey	Definition of technology-driven 21st-century workplace written communication
Need to define 21st-century workplace written communication	Review of professional workplace learning reports & magazines	Revision of CEFR Can-Do Statements for workplace written communication purposes
Need to develop proficiency-oriented performance-based assessment tasks	Review of workplace English textbooks & curricula	A framework for authentic proficiency-oriented performance-based assessment tasks

Figure 2. Schematic diagram of inputs, activities, and results for domain analysis. Portions of this figure were adapted from Chapelle et, al. (2018)

3.2. Methodology for domain analysis research

The domain analysis study was conducted to help the CEPA research and development team obtain data to explore and define the work skills and situations that were central for workplace written communication, vital knowledge representations and language functions in written English communication, and key task features and performances needed to function within this domain. In this regard, the domain analysis study was exploratory, with no pre-defined claims or hypotheses, and implemented a qualitative methodology. As a result, we

came up with the broad research questions that utilize systematic and explicit methods to identify, select, and critically analyze each of the components in the domain analysis.

1. What are the key general and rhetorical purposes, genres, and audiences of English written communication in the modern workday environment of global corporations?
2. What are the main (written) communication goals and conventions (styles, trends, mediums, platforms) valued and used in the 21$_{st}$-century workplace?
3. What are the most prominent written English skills and functions being taught in current business/workplace textbooks?

The study participants and materials, instruments, and procedures to answer these questions are explained next.

3.2.1. Participants and materials

The participants (i.e. stakeholders in this case) were 21 talent acquisition managers who work in the human resources department of either global or local corporations located in Turkey. Among the companies represented, seven were local and located only in Turkey but operated beyond Turkey. The rest were from companies that have offices globally, including ones in Turkey. These talent acquisition managers were responsible for the discovery of, screening, and recruitment of employees for their companies. In some cases, they were also managing the process of finding and delivering the necessary English language services to their current employees.

In addition to the participants, two groups of materials were used in the study. The first group consisted of "LinkedIn Workplace Learning Reports" from 2017 to 2019 ("LinkedIn Learning Resource Center", 2020) and popular business world magazines from 2017 to 2019, including Forbes®, Fortune®, and Harvard Business Review®. The LinkedIn reports were chosen because of their comprehensive look at the defining talent development in today's dynamic world of business. All the popular magazines chosen for the CEPA domain analysis were targeted primarily based on their non-domain specific, globally applicable content and their contemporary nature. Two of the three selections (Forbes and Fortune magazines) can be further delineated into subsections of prominence by leading worldwide subscription numbers and editions of business sources, which are also published in Turkey (and in the Turkish language). Forbes, in 2019, held an adult readership in the U.S. of 6.6+ million and has 39 global editions, and similarly, Fortune holds an adult readership of approximately 3 million in the U.S. with multiple global editions (MRI-Simmons, 2019). The third selection, The Harvard Business Review (HBR), was chosen for its supplementary perspective and its rich value of being an esteemed, academically founded, and sourced magazine, although it garners a lesser amount of circulation.

The second group of materials included the English for business/workplace textbooks. These were "The Market Leader Extra" from Pearson/Longman, "Business One: One" by Oxford University Publishing, and "Business Builder" resource pack by McMillan. These textbooks were selected because they were defined in the mandate handed down by the British Side

Educational Services as the three most commonly used resources by English language schools in Turkey.

3.2.2. Instruments

The instruments used in the domain analysis study included (a) an online written needs analysis survey that talent acquisition managers completed, (b) analysis table to record data from professional workplace learning reports and popular magazines, and (c) an analysis table to summarize the key language skills and functions taught by business/workplace English textbooks. The needs analysis survey completed by the talent acquisition managers included a total of 14 selected-response and open-ended questions that captured the frequency, scope, and purposes of overall English language use as well as the description of forms, methods, and purposes of English written communication being used in their companies. It also captured the key written communication skills their companies were seeking in prospective candidates.

The analysis table to record the data in the reports and magazines included defining the type and scope of the data from each resource and captured the description of workplace learning environment, instances of workplace written communication knowledge areas, skills or strategies, and finally observations of written communication conventions (e.g. style, trends, medium, platforms). The analysis table to record the key language skills and functions taught by business/workplace English textbooks included the key written language skills and functions, contexts, and written communication forms being taught in each textbook at different proficiency levels ranging from "Elementary" to "Advanced" as defined by the publishers.

3.2.3. Procedures

In order to answer the first research question, an anonymous needs analysis survey was sent to human resources decision-makers (i.e. talent acquisition managers) of global corporations in the network of British Side Educational Services in Turkey. Qualtrics® online survey platform was used to deliver the survey. All the questions were translated into Turkish and piloted with two human resources personnel working for the British Side Educational Services before it was sent out to the study participants. The data analysis and reporting functions of the Qualtrics platform were used to generate the preliminary results. Then, to report the descriptive information and ratios, the data were transferred to MS Excel for analysis.

For the second research question, following the magazine selections themselves, the table of contents for each issue of each magazine, from 2016 onwards were mined for recurring subjects and areas of focus in the business world. Next, individual articles were extracted for their more detailed compositions (i.e. goals, strategies, infographics, concepts, arguments, etc.). In doing this, the most current and authentic real-world demands and issues of a global workplace were recorded in the analysis table. As for the annual LinkedIn Workplace Learning Reports, each yearly report was scanned for information regarding the overall learning environment of global corporations, any instances or mention of written

communication strategies/skills/knowledge and the valuing of such, and finally any written communication conventions (trends, styles, medium, platforms, etc.) being used in the global workplace were recorded in the analysis table. As these reports were not intended for conveying or teaching written communication strategies, most of the data on written communication were extracted by inference. It is, therefore, important to note that these reports were used as, and are indicative of a very vague, big-picture snapshot of the corporate global learning and development world as well as some broad view and valuing (implicitly) of real written communication strategies/skills/knowledge and how they are transmitted. An unintended but welcome byproduct of scanning these reports was also gaining insight into global learning and development movements, justifying CEPA's online assessment platforms.

For the third question, a multi-level analysis table was first used to record the relevant information in each category from "The Market Leader Extra", "Business One : One", and "Business Builder" textbooks and resource packs independently. The information was recorded for each proficiency level (Elementary, Pre-intermediate, Intermediate, Upper-intermediate, and Advanced respectively) separately. "Contents" or "Scope and Sequence" sections of each book, which provide a summary of the unit objectives, key functions, and skills for each unit, were analyzed first. Then, the whole book was scanned for any additional information that could be left out in the initial phase, and the tables were updated respectively. In the end, the three independent tables were combined to generate a comprehensive results summary with the key information that was recurring in all three books. While we acknowledge that different naming conventions have been proposed and used by different publishers, the naming convention developed in "The Market Leader Extra" series was used to organize the records.

3.2.4. Results

Research Question 1: What are the key general and rhetorical purposes, genres, and audiences of English written communication in the modern workday environment of global corporations?

The data to respond to the first research question in the domain analysis came from the needs analysis survey that talent acquisition managers completed. They were asked to list and describe the most common rhetorical purposes (labeled as 'purpose' in the survey) and genres (labeled as 'forms' in the survey) occurring in their workplaces. The various audiences, sometimes multiple types of audiences, were assumed by the type of rhetorical purpose given. Before presenting these results, it is important to note that participating companies indicated in the survey that the frequency of English language writing used in their workplace communication is either "a great deal" or "a lot". Only 4 out of 21 companies indicated it was only moderately used. Similarly, they indicated that having employees competent in English written communication is either extremely important or very important to them. In terms of the written communication purposes, the top three rhetorical purposes included drafting formal reports, making presentations, and scheduling for various purposes. This is followed by written negotiations and deciding options. For a full summary, see Figure 3.
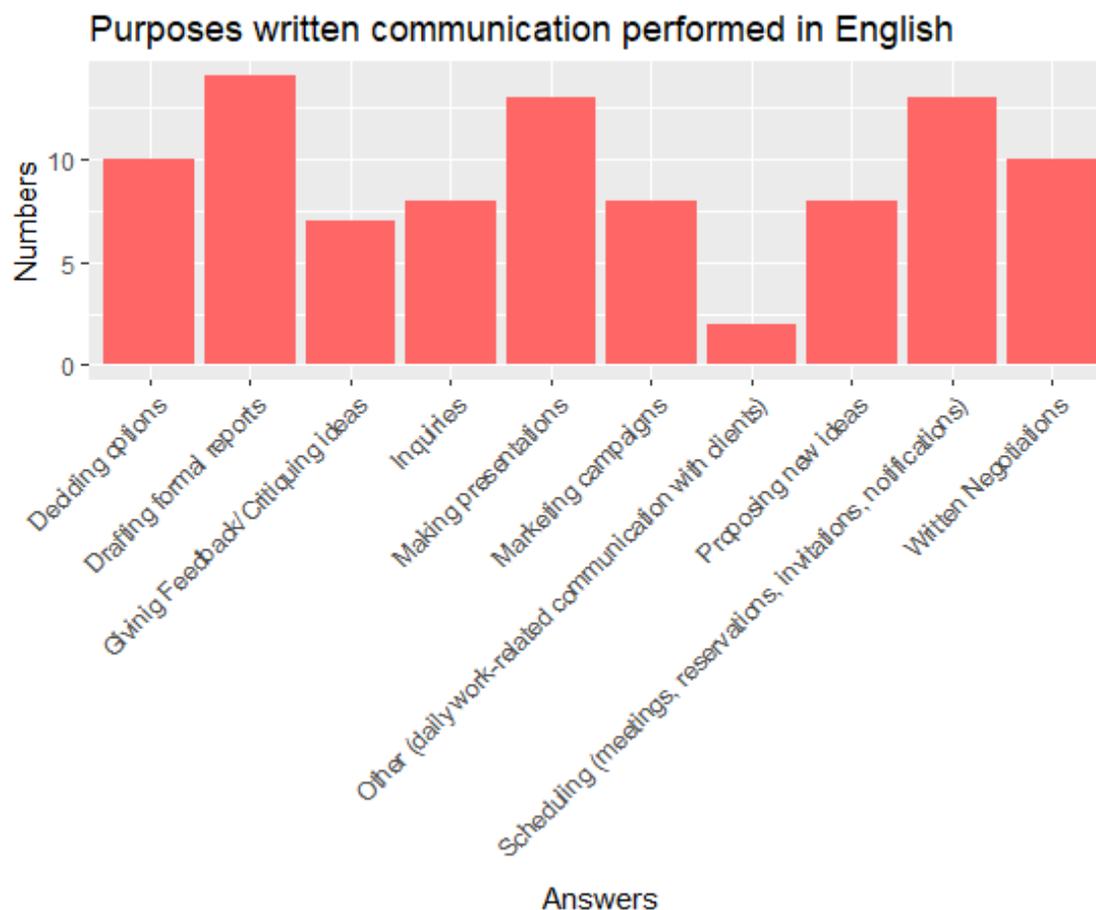
Figure 3. The most common rhetorical purposes of written communication (the vertical column 'numbers' indicates the total number of companies indicating that rhetorical purpose as one of the reasons their employees communicate in written English)

Concerning the most common genres of written communication, emails were the most cited. That is, eighteen out of 21 companies indicated emails as a form of written communication that their employees perform the most often, which was not surprising due to the indispensable nature of email communication in the business world. This was followed by writing reports and instant messages. Instant messaging was an unexpected writing genre that appeared in this list. However, considering the instant messaging platforms recently being introduced (e.g. Slack®) for team communication and collaborations in modern workplaces and the practicality it provides, this form of communication might be even more prevalent in the near future. This discovery, in particular, might create a new challenge for test designers and developers due to its immediate interactive form of communication. Filling in forms and typing text messages were also among the most used genres. Figure 4 shows a full summary of written communication genres performed by employees in English.
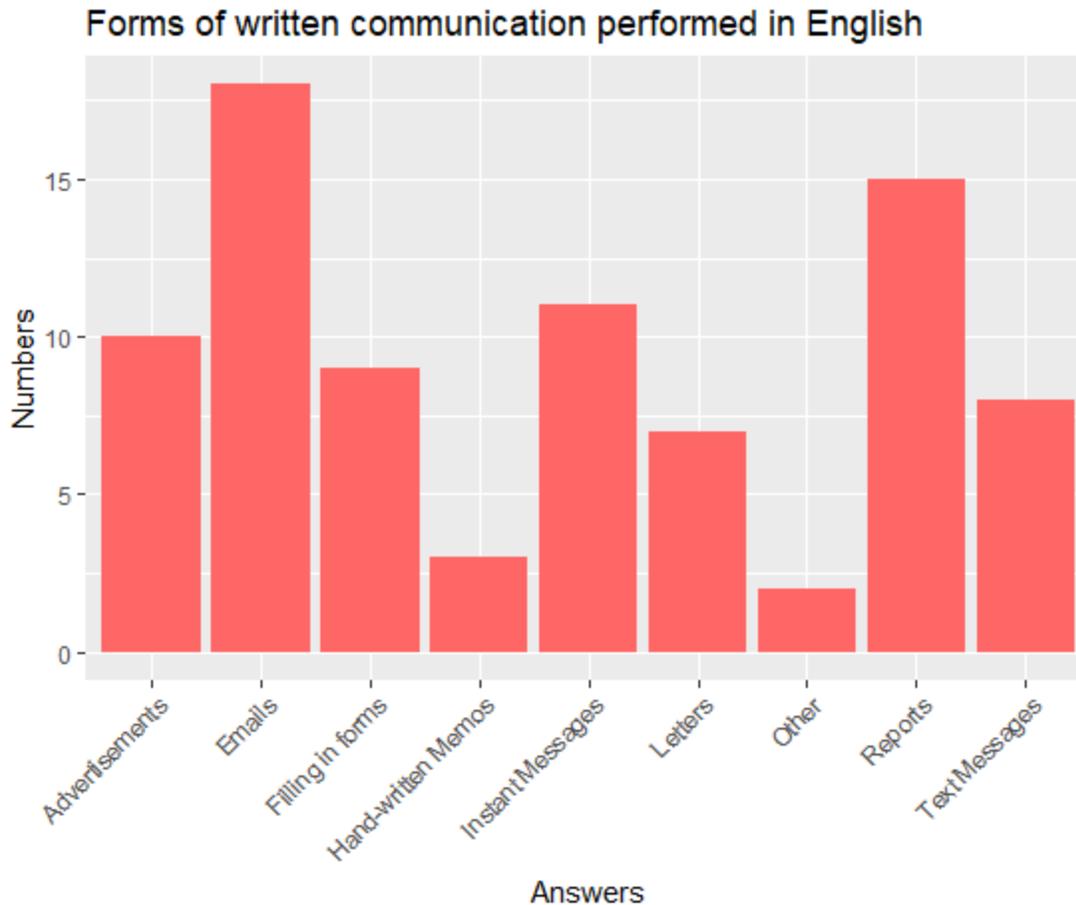
Figure 4. The most common genres of written communication (the vertical column 'numbers' indicates the total number of companies indicating that genre as used by their employees to communicate in written English)

We also asked the participants to rank written communication performed in their workplaces both in terms of frequency and importance. Emails were ranked in the first place by almost 76% of respondents, second by 19% of respondents, and third by 4 % of respondents as the most important and frequent genre of written communication. Instant messages, reports, and filling forms were ranked as the most important and frequent genre of written communication only by 4% each. Reports were ranked as the second most important and frequent genre of communication by 40% of the respondents. This is followed by instant messaging (25% of respondents). On the other hand, instant messaging/text messaging was ranked as the third most frequent and important genre of communication. It was interesting to see hand-written memos not being ranked either by importance or frequency. Figure 5 shows the rankings of written communications performed in English in terms of both frequency and importance.
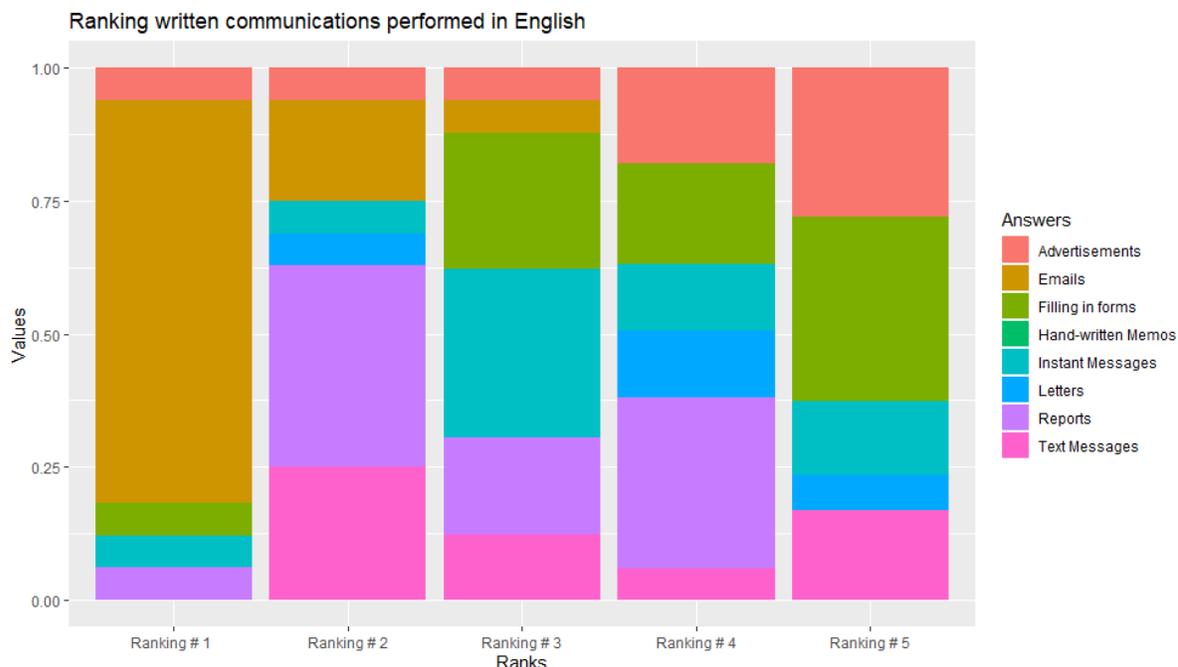
Figure 5. The rankings of written communication genres performed in English in terms of both frequency and importance

Research Question 2: What are the main (written) communication goals and conventions (styles, trends, mediums, platforms) valued and used in the 21st-century workplace?

The data to respond to the second research question came from the critical analysis of workplace research reports and popular magazines. As previously mentioned, these reports were not intended for conveying or teaching written communication strategies. It is, therefore, essential to note that these reports were used as, and are indicative of a very vague, big-picture snapshot of the corporate global learning and development world and some broad view and valuing (implicitly identified) of real written communication goals (e.g. knowledge areas, skills, strategies) and conventions (i.e. how they are transmitted).

According to the LinkedIn Workplace Learning Reports from 2016 to 2018, most of the main (written) communication goals were closely connected to leadership, people management, and collaboration. It should be noted that all of the reports championed adept communication as one of the most important skills in the workplace. Though mostly nonspecific, there were a few examples (soft skills and hard skills) narrowing down the kinds of communication skills that corporations were seeking to acquire or refine. The soft skills, to name a few, were listed as creativity, persuasion, analytical reasoning, and collaboration. The hard skills were much too specific to be of particular use but are noted in Table 1. Conventions found around these communication goals are best summarized as multi-channeled (most of which are located online) and fluid or fluctuating, often dependent on internal or external audiences.

As for the popular global business magazines (The Harvard Business Review®, Fortune®, and Forbes®), many of the main themes and topics throughout these magazines support the connections stated in the Linkedin Reports, especially with emphases placed on leadership,

people management, and collaboration. There are many articles in each magazine (and in each issue) that highlight the importance and give advice on how to more effectively communicate as a leader/manager, as a subordinate, in teams, with colleagues, and with outside clients or partners. And again, most of this communication is happening more and more over digital and online platforms (e.g. company team chat applications, email, social media, online collaborative workspaces, and conferencing platforms). A summary of communication goals and conventions (styles, trends, mediums, platforms) in the 21st-century workplace is also given in Table 1.

Research Question 3: What are the most prominent written English skills and functions being taught in current business/workplace textbooks?

The data to respond to the third research question came from the critical analysis of business/workplace textbooks. The results revealed many language skills and functions specific to written communication at different proficiency levels. At the lower levels, these included relatively smooth and straightforward communicative functions such as making introductions, asking or answering simple questions, making arrangements, or exchanging information. However, presenting an idea in writing or negotiating a situation with no complications were also among key functions at the pre-intermediate level. As the level went up, the functions became more complex by embedding a more formal register or a broader purpose to the functions taught at lower levels (e.g. networking in formal registers instead of simple introductions). They also became complex in terms of the situation itself. That is, the situations included a complication such as negotiating a solution, resolving a conflict, taking part in decisions, and presenting arguments.

In terms of the contexts, the breadth increased as the levels went up. At lower levels, the contexts included job fairs, phone conversations, business correspondence, meetings, business meals and parties, and, rarely, business deals. At the higher levels, it included customer relations and complaints, formal presentations, business contracts, and terms and conditions. Conversely, the genres of written communication stayed mostly consistent across all proficiency levels. For instance, writing emails, action minutes, reports, or letters existed at all levels, though the goal and scope changed. Writing summaries, formal letters, presentation slides appeared only at higher levels. It is important to note that none of these textbooks included text messaging or any other forms of instant messaging as part of their main objectives of written communication forms. The most important skills, functions, contexts, and forms of written communication in business English textbooks were synthesized and are presented in Table 2.

Table 1. Summary of communication goals and conventions (styles, trends, mediums, platforms) in the 21st-century workplace

Primary Source	Type and Scope of Data	Description of workplace learning environment	Instances of workplace written communication goals (knowledge areas, skills, strategies)	Observations of written communication conventions (styles, trends, mediums, platforms)
LinkedIn Reports: Annual Workplace Learning Report ("LinkedIn Learning Resource Center", 2020)	**2017** – Survey of 500 L&D (Learning and Development) professionals with profiles on LinkedIn from the U.S. and Canada, who influence or are decision-makers for their companies' L&D budgets	• L&D growing in demand and value • Highly focused on Leadership/People Management • 3 most important skills: coaching, leadership communication, and team collaboration	• Focus on "leadership communication" (no specific communication type) • Developing a [written] communication plan for your company	• An example of <u>written message dialogue</u> demonstrating how to communicate effectively as an L&D consultant (consultancy skills) • Link to "best practice" <u>consultancy questions</u>
	2018 – Survey of 4,000 professionals (1,200 talent development professionals, 2,200 employees, 200 executives, and 400 people managers) with profiles on LinkedIn from around the world (North America, Europe, and Asia)	• Short shelf-life of skills and focus on developing employees "ahead of the curve" • High demand from employees for development, but not making time for it • Steep rise of online learning	• Communication is the 2nd most important skill for employees to learn from L&D programs • Execs believe the most important skills are soft skills: leadership, communication, and collaboration	• N/A
	2019 – Survey of more than 1,200 talent developers (Learning and Development professionals and HR professionals with L&D development responsibilities) and over 2,100 workplace learners globally with additional data from	• Fewer budget constrains for online learning • A shift from instructor-led and in-person training to online solutions • Tailored training (mix and match content)	• Most in-demand soft skills: creativity, persuasion, analytical reasoning, collaboration, flexible approach, and time management • Most in demand (and relevant to written communication) hard skills: video production (subtitles), people	• Engage learning in company by using a *creative* multi-channel approach: <u>email, intranet posts, signage, videos, and influencer campaigns</u> • Branding L&D department with informal approaches, gaining

	LinkedIn Learning platform data from 2017 to 2018	• Talent developers prioritize identifying, assessing, and closing skill gaps • Mobile learning is rising • Consistent learning globally	management (evaluations, reports, communication), and translation	recognition with cultural <u>tropes like the infamous 'got milk' ad</u> to grab attention • Creating an organizational culture with <u>upbeat internal emails that include working jokes and quirky quotes</u> for reader enjoyment

Business Magazines (Forbes®, Fortune®, and Harvard Business Review®)	Editorials, articles, and infographics from a range of highly subscribed to, general-business themed, and globally circulated online subscription magazines from 2016 to 2019	• Dynamic, future-oriented focus, and encompassing a common global business culture
		• Mobile, highly interconnected interests/outcomes, and real-time communication
		• Centered on leadership, innovation, creativity, and culture-building
		• Technology, automation and a turn toward soft skills concern/focus
			• Best practices and advice in conducting:
			  - appraisals/assessments/reviews
			  - reports/presentations
			  - discussions/meetings
			  - hiring/training
			  - incentivization/regulation
			  - negotiations/communication
			
			…of or between:
			  - departments/teams /employees
			  - managers/executives
			  - companies/company policies and practices
			  - products/services
			  - customers/shareholders
			
			• Identifying and weighing the effects of business-world trends
				Formal and professional-level tonality with some kinds of external communication; some implication that written communication is more easily kept as a "permanent" record of communication
				
				A noticeable transition into new mediums (mobile and augmented reality) with an emphasis on informal language within internal team-building/culture-building communication and external marketing/branding (when targeting specific audiences *sociocultural nuances)

Table 2. Summary of key skills, functions, contexts, and genres of written communication in business English textbooks

Level	Key Written Language Skills and Functions	Contexts	Genres of Written Communication
Elementary	• Introducing yourself and others • Asking and answering simple questions • Making bookings & checking arrangements • Making decisions • Making arrangements • Describing a product	• Introductions at a job fair or work • Writing or responding to customers • Workspace/office related issues	• E-mail • Action minutes • Presentation • Letter
Pre-intermediate	• Writing notes • Making contact • Exchanging information • Presenting your company/product • Negotiating an arrangement	• Phone conversations • Meetings • Advertisements/Sales • Company meals/parties	• Notes • E-mail • Proposal • Letter • Report
Intermediate	• Managing meetings • Writing action minutes • Presenting a new idea • Taking parts in decisions • Negotiate a deal/ a new contract/conflict	• Job interviews • Business meetings and follow-ups • Business correspondence • Introductions and networking • Business deals/contracts	• E-mail • Action Minutes • Summary • Letter • Report
Upper-intermediate	• Making or presenting recommendations • Solving problems/Responding to situations with complication • Resolving a conflict • Reaching agreements	• Business meetings and follow-ups • Customer complaints • Customer relations	• E-mail • Action Minutes • Guidelines • Report • Letter • Summary
Advanced	• Networking in formal and informal registers • Making arguments/presenting claims • Writing presentation slides • Managing situations/questions with complications	• Formal networking • Formal Presentations • Business Contracts • Sales and agreements • Customer relations • Business terms and conditions	• Formal letter • E-mail • Report • Press Release • Summary • Presentation slides

3.3. Outcomes of the domain analysis

In evidence-centered design, the intended outcomes of the domain analysis help the test designers to gain insight into and focus on (i.e. define) the essentials, including central work and situations, the vital knowledge representations and language functions, and key task features and performances in the target domain. This domain definition thus "presage the entities and structures that appear in subsequent layers" (Mislevy & Haertel, 2006, p. 7) in order to design the assessment tasks.

In the CEPA project, the domain analysis helped us understand the central skill requirements and situations of written communication in global corporations as well as vital knowledge representations, language functions, and key task features and performances needed to function within this domain, all of which led to our understanding of a global definition of technology-driven 21$_{st}$-century workplace written communication. Subsequently, we also critically analyzed the CEFR Can-Do statements (in this case, for writing) in the light of domain analysis outcomes and revised them to represent key language knowledge, skills, and functions of written communication in global corporations. This was an essential step we had to take since the mandate required us to align the assessment with the CEFR guidelines even though they were not explicitly designed for the target domain under investigation. Finally, we adapted a framework to define key characteristics of authentic proficiency-oriented performance-based assessment tasks. In the following subsections, we describe and expand on each of these outcomes.

3.3.1. **Definition of technology-driven 21$_{st}$-century workplace written communication**

According to the domain analysis, we could define the technology-driven 21$_{st}$-century workplace written communication as business correspondence involving both real-time and delayed unidirectional or bidirectional communicative acts either between the members of the same corporate/team or other business partners/clients/potential clients, mainly through technology-mediated communication channels, in order to work on tasks situated within socioculturally dynamic contexts. The domain analysis also indicated that communication sometimes spins off from a multi-modal input and sometimes without such an input.

The written communication acts in this regard are action-oriented since all of the themes, situations, and tasks identified in the domain analysis require the use of measurable and demonstrable language functions (written communicative competences), which lead to real-life consequences in the workplace environment. Additionally, whether it is real-time or delayed, unidirectional or bidirectional, the situations and tasks involve one or multiple audiences at different levels of formality and hierarchy. To this end, written communicative competence in this domain reflects an interrelated integration of linguistic, sociolinguistic, and pragmatic competence. Similarly, this necessitates the utilization of a proficiency perspective. Here, we also define and use the term 'proficiency assessments' as ones used to assess a test taker's level of language ability without respect to a particular curriculum (Carr, 2011). In this regard, proficiency focuses on the ability to functionally use a language in real-life settings and indicates what a test taker can do with the language in the real world which typically involves assessing more than one narrow aspect of language.

Moreover, in terms of laying out the reasons for written communication, the themes or topics are so numerous that it is impractical to come up with a comprehensive yet unified list of all business correspondence. In other words, each business has its own specific context and language that might necessitate the use of particular vocabulary or jargon. This, as a result, constitutes a critical challenge in terms of representing authentic situations in the description of this domain. However, it was possible to identify the overall purposes of the correspondence. These include (a) information presentation, (b) information exchange or business inquiries, (c) sales and advertisements, (d) customer relations and success, (e)

negotiation, (f) resolving a conflict, and (g) networking. This outcome highlights the need of bridging general language and language for specific purposes (very specific purposes in some cases), as evident in the literature (see O'Sullivan, 2006; 2012 for a detailed analysis of the issue).

3.3.2. Revision of CEFR can-do statements for workplace written communication purposes

The description of the domain suggested revision of the writing Can-Do statements in order to achieve the re-alignment required by the mandate. This kind of adaptation is in keeping with the intentions of the CEFR designers:

The construction of a comprehensive, transparent and coherent framework for language learning and teaching does not imply the imposition of one single uniform system. On the contrary, the framework should be open and flexible, so that it can be applied, with such adaptations as prove necessary, to particular situations (*Common European Framework of Reference for Languages*, 2001).

Accordingly, the changes were made in two broad categories. First, we added or sometimes replaced the general language knowledge, skills or abilities with domain-specific ones identified by the domain analysis. Second, we broke down the Can-Do statements into a more analytical set of categories at each main level (A1 to C2) of CEFR. These categories included overall written communication, content and coherence, grammatical accuracy, vocabulary range, orthographic control, sociocultural competency, and workplace communication strategies in addition to a global description. Table 3 exemplifies the revised Can-Do statements in writing based on CEFR's A1 level.

Table 3. Revised writing Can-Do statements in CEFR A1

Written Communication Competencies	Can Do Statement at A1 Written Communication Proficiency
Overall Written Communication	Can write simple and short texts (brief emails, short memos, text messages) conveying essential information to someone as part of a business transaction or in response to a simple and straightforward request or question
Content & Coherence	Can write very short texts in a logical and organized manner (with some amount of deviation), which would be understood by a very sympathetic audience
Grammatical Structure & Accuracy	Can use very simple words, short sentences or common phrases but is prone to making many errors in these structures and in attempts at more complex constructions
Vocabulary Range & Choice	Can use only a very limited range of very simple, high-frequency vocabulary (e.g. greetings, numbers), alphabet, subject pronouns, dates, times, occupations, the most common tangible objects (e.g. chair, mobile phone, door), and the most common verbs (e.g. to be, to have, to go)
Orthographic Control	Can write very simple words, short sentences or common phrases, but these may include many errors in capitalization, hand-written characters, spelling, and punctuation

Sociocultural Competency	Demonstrates very little familiarity of, or sensitivity to, relevant sociocultural subject matters and circumstances (e.g. daily practices, context and situational awareness, potential group cultural practices or observances, and potential group stigmas)
Workplace Communication Strategies	Demonstrates very little familiarity or understanding of written formal workplace communication patterns (e.g. appropriate tone and register) and strategic approaches (e.g. use of examples and providing support for ideas)

In addition to the seven categories represented in the table, we also came up with a global description of written communication proficiency at each CEFR level. For the A1, this reads as follows:

The candidate at A1 has very basic written communication skills, which may contain several orthographic or grammatical errors that greatly hinder the writer's ability to convey information. The candidate can write in very basic phrases aimed at satisfying immediate and tangible workplace needs. They may require extended periods of time and re-referencing of the instructions or questions to produce written English. In formal written tasks, coherence may diminish significantly.

The revision of CEFR was a critical undertaking that yielded a more specific, and therefore more useful framework for the CEPA project. The revised framework included, for instance, sociocultural competency (specific to the diverse, but unique groups of global corporations) as well as workplace communication strategies, which were found to be vital in carrying out day to day written communication tasks.

3.3.3. A framework for authentic proficiency-oriented performance-based assessment tasks that accounts for sociocultural competency

The analysis of written communication in the domain of global corporations revealed that communication is, for the most part, action-oriented and task-based. It also revealed the indispensable role of sociocultural competency (an awareness and sensitivity to a group/context) and workplace communication strategies (support of argument, providing examples, re-phrasing, etc.) in this domain. This highlighted a need to have a framework that defines the key characteristics of and establishes clear criteria for authentic proficiency-oriented performance-based tasks. These tasks are distinguished from other types of assessments by their particular features in that they (a) present test takers with a problem: a real-world goal, set within a realistic and relevant context of challenges and possibilities, (b) push test-takers to develop a concrete product or performance for an intended audience (real or simulated), and (c) set evaluation criteria and performance standards that are appropriate for the task (Wiggins & McTighe, 2008).

In search of such a model, we decided to adapt t*he goal*, *role*, *audience*, *situation*, *product*, *standards* (GRASPS) model by Wiggins and McTighe (2008), which appeared in the context of the general education field. As the name indicates, the model sets six principles to ensure that the task is proficiency-oriented and performance-based so that it can elicit the target function(s) being measured in the task. Figure 6 shows the schematic diagram of the GRASPS model, laying out criteria to design performance-based assessment tasks.

This model was selected for a few reasons: (a) it defines and sets criteria for authentic proficiency-oriented and performance-based tasks (b) it defines clear parameters to contextualize the tasks, (c) it shifts the task completion criteria from quantity-based standards (e.g. minimum number of words or paragraphs) to quality-based, real-world language use criteria (i.e. whatever it takes to communicate the goal or complete the task), and (d) it allows measuring sociocultural competency and workplace communication strategies in a task-based assessment model by specifying the communicative context, the specific audience and the relationship to that audience, and a clear purpose towards a realistic and relevant goal. This last reason was particularly important as previous research indicated some key challenges relevant to measuring L2 pragmatics (e.g. Norris, 2016; Timpe-Laughlin et al., 2015).

In summary, three outcomes were derived from the domain analysis: we could define the necessary parameters for technology-driven 21st-century workplace written communication, which partially yield the following two outcomes; that is the need for revision of CEFR Can-Do statements and developing a model for proficiency-oriented performance-based assessment tasks. These outcomes constituted the inputs for the next layer, domain modeling in the ECD.

4. Domain modeling

The layer that follows the domain analysis in the ECD process is domain modeling in which the test developers "organize the information and relationships discovered in domain analysis into the shape of assessment arguments" (Mislevy & Haertel, 2006, p. 8). These arguments help test designers to "lay out what an assessment is meant to measure and how and why it will do so" (Mislevy & Haertel, 2006, p. 8). In this layer, the goal is to establish a connection between the knowledge, skills, and abilities discovered in the domain of interest and the situations, activities, and tasks that engender these in the form of assessment arguments. Therefore, the discoveries from the domain analysis serve as inputs in the domain modeling process of defining design patterns. Three aspects from the domain analysis served as rationales for the design patterns described in this section. These are: (a) definition of technology-driven 21st-century workplace written communication, (b) revision of CEFR Can-Do Statements for workplace written communication purposes, and (c) a framework for authentic proficiency-oriented performance-based assessment tasks.

Domain modeling can be developed through different supporting tools. One tool for domain modeling is to develop Toulmin diagrams for assessment arguments. A Toulmin diagram for assessment maps the assessment argument into a schema that lays out what assessments intend to measure and "...what behaviors or performances should reveal those constructs, and what tasks or situations should elicit those behaviors" (Messick, 1994, p. 16). Figure 7, adopted from Mislevy & Hartell (2006) and Mislevy (2011), is one of the representational forms that shows the implementation of domain modeling.

Explain what it means to successfully complete the task. Describe how their product or performance will be evaluated.

Explain what test takers' task is and what goal(s) they need to accomplish.

GOAL

STANDARDS (CRITERIA)

ROLE

ASSESSMENT TASK that ELICITS TARGET FUNCTIONS

Describe what test takers' role is in the task.

PRODUCT PERFORMANCE PURPOSE

AUDIENCE

SITUATION

Describe what test takers need to do (produce or perform something) using written English.

Explain the audience involved in the task and their relationship to the test takers, establishing the need to communicate.

Explain the (workplace) context that the tests takers will find themselves in as a non-native speaker of English.

Figure 6. Schematic diagram of the GRASPS model laying out criteria to design proficiency-oriented performance-based assessment tasks

The schema first starts with a claim on the top about examinees' knowledge and skills, which is based on the observation of their performance at the bottom. To be able to make a claim from the observed performance based on a reasoning, the claim should be supported "…through a warrant or rationale as to why examinees with particular knowledge or skill are likely to act in certain ways during the situation at hand" (Mislevy, 2011, p. 9). The kinds of data that support the claim through this warrant come from the domain analysis, which reveals the key knowledge, skills and, abilities as well as the situations, activities, and tasks that provoke these. The "alternative explanations" in the schema in Messick's (1989) terms are either about "construct underrepresentation" or "construct irrelevant demands". For instance, if the assessment task does not include a situation or environment that can provoke the targeted knowledge and skills, that means the intended construct is not well represented. Likewise, if the examinees need knowledge and skills beyond those targeted in order to complete the task successfully, then the assessment tasks create an extra burden for them, and that goes beyond what is intended to be measured. Both of these cases constitute a significant threat to validity (Mislevy, 2011).

Claim about examinees' particular knowledge and skills (the construct intended to be measured)

Unless — Alternative explanations (construct irrelevant factors threatening the validity)

Warrant for assessment argument supported by research — **Since**

So

Data from Domain Analysis on knowledge, skills, and abilities concerning the examinees	Data from Domain Analysis on situations, activities, and tasks that can provoke the intended constructs

Warrants for ascertaining salient aspects of examinees' performance — **Since**

Warrants for ascertaining salient aspects of assessment situation or task — **Since**

Examinees taking a test or completing assessment tasks

Figure 7. A Toulmin argument diagram for assessment argument

Next, the domain model for the CEPA Written Communication Assessment is presented using a Toulmin's argument diagram. The warrants used in the assessment argument come from the domain analysis study reported in the previous section. Figure 8 illustrates the domain modeling for the current project through Toulmin's argument diagram, as described above. Each part in the diagram thus includes the specific assessment arguments that need to be considered in designing tasks for the CEPA Written Communication Assessment.

Candidates can perform technology-driven written communication tasks demanded by 21st-century English workplace environment in global corporations

- The tasks must be situationally and interactionally authentic in a way that requires and elicits communicative functions (Douglas, 2000).
- The tasks must be language-use tasks, which require and allow candidates to use language to achieve a particular goal or objective (Bachman & Palmer, 1996).
- The task inputs need to bridge general language and language for specific purposes to avoid jargon or business-specific language (O'Sullivan, 2006; 2011).

Unless

Since

So

Alternative Explanations:

- Candidates' familiarity with proficiency-oriented performance-based (i.e. constructed response) assessment tasks
- The ability to use interactive communication technologies

- The tasks must elicit written communication competencies, including content and coherence, grammatical structure and accuracy, vocabulary range and choice, orthographic control, sociocultural competency, and workplace communication strategies integratively as defined by the domain analysis.
- The lower-level tasks must elicit simple and straightforward communicative functions such as making introductions, asking or answering simple questions, making arrangements, or e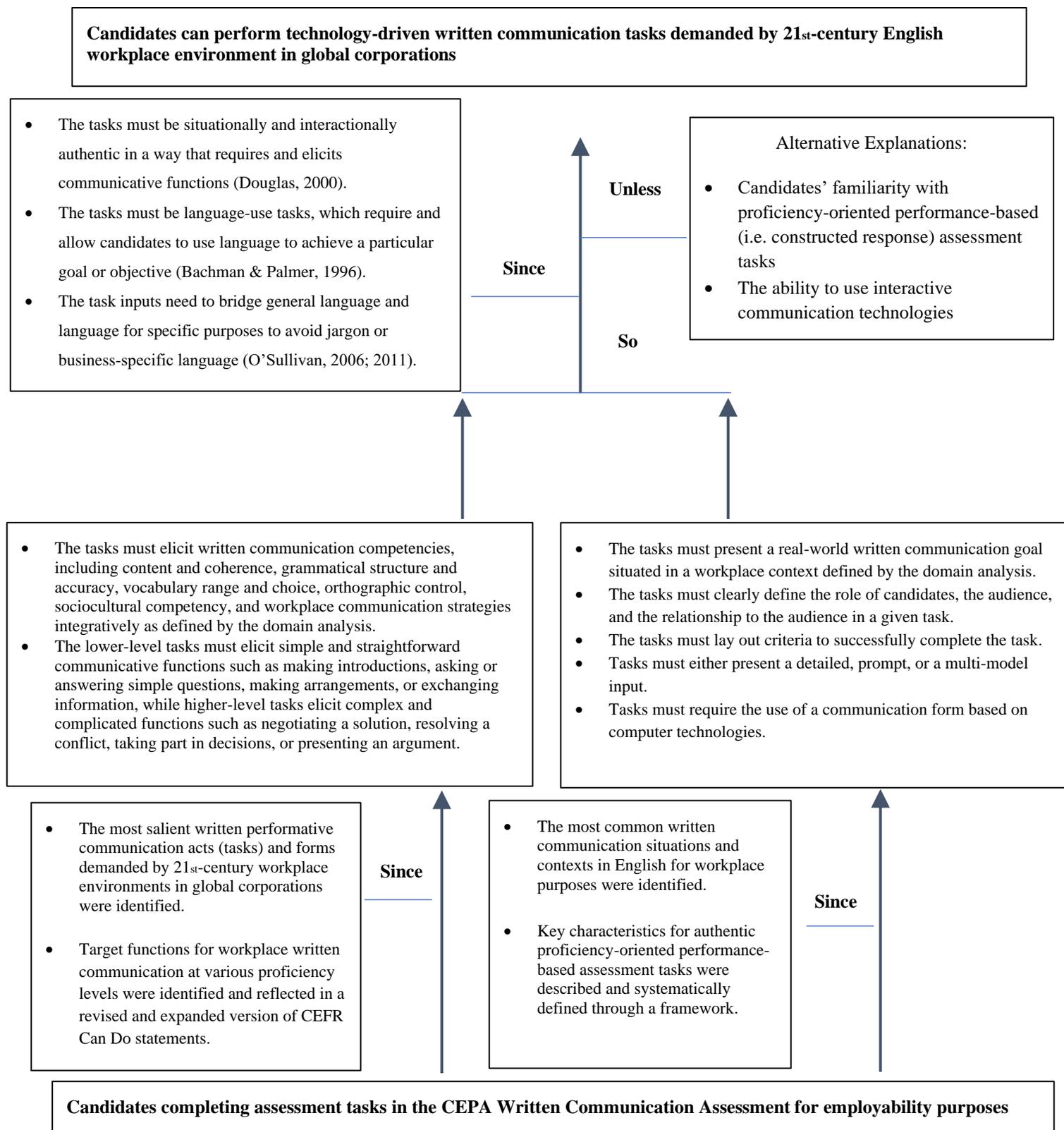xchanging information, while higher-level tasks elicit complex and complicated functions such as negotiating a solution, resolving a conflict, taking part in decisions, or presenting an argument.

- The tasks must present a real-world written communication goal situated in a workplace context defined by the domain analysis.
- The tasks must clearly define the role of candidates, the audience, and the relationship to the audience in a given task.
- The tasks must lay out criteria to successfully complete the task.
- Tasks must either present a detailed, prompt, or a multi-model input.
- Tasks must require the use of a communication form based on computer technologies.

- The most salient written performative communication acts (tasks) and forms demanded by 21st-century workplace environments in global corporations were identified.

- Target functions for workplace written communication at various proficiency levels were identified and reflected in a revised and expanded version of CEFR Can Do statements.

Since

- The most common written communication situations and contexts in English for workplace purposes were identified.

- Key characteristics for authentic proficiency-oriented performance-based assessment tasks were described and systematically defined through a framework.

Since

Candidates completing assessment tasks in the CEPA Written Communication Assessment for employability purposes

Figure 8. CEPA Written Communication Assessment domain modeling through Toulmin's argument diagram

The domain model for the CEPA starts with the claim on the top that once the candidates complete assessment tasks in the CEPA Written Communication Assessment for employability purposes, one will be able to make inferences about their ability to perform technology-driven written communication tasks demanded by 21st-century English workplace environment in global corporations; that is, the target domain that the assessment results intend to extrapolate. To be able to make that inference, first of all, one needs to be able to define what important knowledge and skills underlie the use of written English as a means of communication in the 21st-century workplace environment in global corporations. These were revealed by the domain analysis study reported above. Based on the outcomes of the study, two sets of arguments are presented.

The first set is pertinent to the data from domain analysis on knowledge, skills, and abilities (i.e. construct-related arguments). Here, it is argued that the tasks must elicit multiple written communication competencies or traits, including content and coherence, grammatical structure and accuracy, vocabulary range and choice, orthographic control, sociocultural competency, and workplace communication strategies integratively as defined by the domain analysis and reflected in revised CEFR Can-Do statements. An additional argument in this regard was made that the value of these traits must not be equally weighted; that is, language use traits must carry more value than the language-knowledge traits as a result of the utilization of a proficiency perspective in the project.

It is noteworthy to mention, the specific categories of sociocultural competency and workplace communication strategies were selected, adapted, and included to address CEFR's call for competence "concerned with the knowledge and skills required to deal with the social dimension of language use" and the ways (i.e. strategies) in which the application of these competencies are employed (in the workplace context) or compensated for (CEFR, 2001, p. 118; p. 25). Sociocultural competency and workplace strategies categories often overlap with each other as both are relative to the same context, and by placing tasks in very precise situations with defined audiences, it is possible to elicit both. For example, when considering the construction of a task, it should have the potential to evoke sociolinguistic competency factors such as linguistic markers of social relations (e.g. professional address with some distance), politeness conventions (e.g. expressing gratitude or offering help), and even expressions of folk wisdom (e.g. "on the same page", "smooth sailing", "go the extra mile") as well as strategies (e.g. providing examples, support, or reasoning for choices to name only a few) to aid in demonstrating and clarifying these competencies, or lack thereof, with performance.

It is also argued that the lower-level proficiency tasks must elicit simple and straightforward communicative functions such as making introductions, asking or answering simple questions, making arrangements, or exchanging information. Whereas, higher-level ones must elicit complex and complicated functions such as negotiating a solution, resolving a conflict, taking part in decisions, or presenting an argument. These arguments are based on the finding that through the domain analysis the most salient written performative communication acts and forms demanded by 21st-century workplace environment in global corporations could be identified and that target functions to carry out these at various proficiency levels were reflected in a revised and expanded version of CEFR Can-Do

statements. For instance, one key implication of these arguments is that the CEPA Written Communication Assessment needs to be designed as a criterion-referenced test so that candidates' written communication skills are measured against the target functions (i.e. a set of criteria). Similarly, the rubrics and scoring procedures to evaluate candidates' performance must be designed according to the same criteria.

The second set of arguments, considering the data from domain analysis on situations, activities, and tasks, is that the tasks can provoke intended constructs (i.e. the task-related arguments). Here, multiple arguments are presented in that the tasks must: (a) present a real-word written communication goal situated in a workplace context defined by the domain analysis, (b) clearly define the role of candidates, the audience, and the relationship to the audience in a given task, (c) lay out criteria to successfully complete the task, (d) present either a detailed prompt or a multi-model input, and (e) require the use of a communication form based on computer technologies. These arguments are based on the findings and description, through the domain analysis, of the common written communication situations and contexts in English for workplace purposes and the key characteristics for authentic proficiency-oriented performance-based assessment tasks, which were all systematically defined in the form of a framework (i.e. GRASPS model).

Moving upward in the domain modeling diagram, the construct-related and task-related arguments are further developed through key findings and implications from relevant research. First of all, it is argued that the tasks must be situationally and interactionally authentic in a way that requires and elicits the target communicative functions (Douglas, 2000). The tasks must be language-use tasks which require and allow candidates to use language to achieve a particular goal or objective (Bachman & Palmer, 1996). The task inputs need to bridge general language and language for specific purposes to avoid jargon or an established content knowledge or expertise (O'Sullivan, 2006; 2012).

The final component in the domain model, as demonstrated in the Toulmin diagram, is "alternative explanations". These are the construct irrelevant variables that might threaten the validity of the claims being made in the domain model. In the current model, part of this is related to test-takers' familiarity with the task types and a mismatch between the written communication assessment tasks and those in the real-world 21st-century workplace environment. The other part is related to the actual ability of using interactive technologies. The domain analysis revealed that almost all of the communication forms are technology-driven, and thus an argument can be made that the task must require the use of a communication form based on computer technologies. However, great care must be taken so that the completion of the assessment tasks would not require a technical knowledge and experience of using a particular communication technology in order to complete the tasks successfully.

Overall, the assessment arguments that are laid out in the schema concerning the design and development of all the assessment tasks in the CEPA Written Communication Assessment define parameters for test design patterns. In other words, any assessment task to be developed for CEPA Written Communication Assessment needs to consider these arguments, which are built on the findings of the domain analysis. Since the CEPA system proposes

multiple assessment tasks targeting various proficiency levels, the assessment arguments here are left broad on purpose. The domain model presented here provides the fundamental assessment arguments and highlights the interconnections among aspects of examinees' capabilities, situations, and behaviors (Mislevy, 2011, p. 14) and thus use the assessment arguments to create a blueprint, or more commonly known as test specifications for the CEPA Written Communication Assessment in the next ECD layer, namely, the Conceptual Assessment Framework (CAF).

5. Conceptual assessment framework

In the CAF phase of the ECD process, "a blueprint is created considering the original mandate along with the constraints and logistics of the project and the results of the previous steps" (Chapelle et al., 2018, p. 19). In undertaking the CAF development, creating models for task design is essential. These include (a) student model, (b) task model, and (c) evidence model. Mislevy and Haertel (2006) described that, in the student model, specifications for what should be measured are developed "in terms of variables that reflect aspects of students' proficiencies" (Mislevy & Haertel, 2006, p. 10). In the task model, the specification needs to frame what the tasks should look like and clearly define "the environment in which students say, do, or make something to provide evidence. A key design decision is specifying the form in which students' performances will be captured" (Mislevy & Haertel, 2006, p. 10). In the evidence model, specifications must address "how one identifies and evaluates the salient aspects of student work, in terms of values of observable variables" (Mislevy & Haertel, 2006, p. 10). However, test designers need to make sure that the specifications in these three models are aligned so that a task can elicit the target functions it intends to measure and therefore what is inferred, as a result, about the candidates' performance in that regard is an accurate and valid interpretation of a person's language skills in the real-life context for which the tasks are developed.

In the CEPA project, CAF was developed using design arguments from domain modeling, which were built on domain analysis outcomes, as well as parameters defined by the original mandate. For the CEPA Written Communication Assessment, the student model needed to reflect candidates' written communication skills in the 21$_{st}$-century workplace environment and had to be interpreted in accordance and aligned with CEFR guidelines. As previously described, the CEFR Can-Do descriptors in writing were not explicitly designed to define workplace-written communication skills and strategies. We, thus, used the revised written communication functions, including sociocultural competency and workplace communication strategies, to elaborate the abilities underlying performance required by these guidelines. The written communication functions at the six proficiency-levels provided the detail required for developing and specifying what should be measured. These also provided the key patterns to define the task difficulty and the alignment with the corresponding CEFR level.

Specifications of CEPA Written communication task model required making design decisions to operationalize task-related arguments developed in the domain modeling. For each proficiency level, contexts, and topics, along with task type and format, and input that can be used in the task were defined. Moreover, to make sure that the tasks are proficiency-

oriented and performance-based, we integrated task specifications with the GRASPS framework. Thus, what candidates needed to produce at a given proficiency level was made clear. This was a vital procedure so that the forms where candidates perform could be captured, and as a result, an evidence model that could evaluate such a performance could be developed. Due to practicality concerns, the number of tasks was dropped from six to three. The first task, therefore, would capture the measurement component specified in the corresponding A1 and A2 levels, the task 2 in the B1 and B2 levels, and task 3 in C1 and C2 levels.

Specifications for the CEPA Written Assessment communication evidence model were developed in accordance with the construct-related arguments made in the domain modeling. The CEPA project utilized multi-trait analytical scoring procedures (in contrast to a holistic view) to identify the key traits that underlie the multiple written communication competencies targeted integratively in the candidates' products. These were designed to capture both language knowledge and use dimensions and included both language knowledge (grammatical accuracy, vocabulary range and choice, and mechanics) and language use (content and coherence, workplace communication, and sociocultural strategies) traits. Accordingly, these traits were not equally weighted; that is, language-use related traits were awarded more points than language-knowledge related traits. Finally, the total available points were distributed according to task difficulty so that the difference in performance could be reliably reflected in the score. The key specifications in the student model, task model, and evidence model were combined under the CEPA Assessment Design Matrix presented in Table 4.

For the CEPA project, the key outcome of the CAF was the development of task frames sketching task-level details in preparation to operationalize the specifications. The frames were kept general so that alternative tasks at the same difficulty level can be produced to elicit the corresponding written communication functions. These frames also played a key role in the creation of testing engine programs and databases during the assessment implementation stage of ECD. Figure 9 illustrates the task frame for Task 1 in the CEPA Written Communication Assessment.

Table 4. CEPA Written Communication Assessment design matrix based on ECD task models

Student Model Specifications What should be measured? Variables that reflect aspects of candidates' proficiencies		Task Model Specifications What the tasks should look like? The environment in which candidate perform the written communication task to provide evidence				Evidence Model Specifications How the tasks will be evaluated? Identification and evaluation of the salient aspects of candidate performance	
CEFR Level	**Key Written Communication Functions**	**Contexts/Topics**	**Possible Tasks**	**Task Type/Form**	**Input Type**	**Evaluation Criteria**	**Scoring**
A1	Can write simple and short texts (brief emails, short memos, text messages) conveying essential information to someone or in response to a simple and straightforward request or question as part of a business transaction	Introductions Writing or responding to a colleague or customer Meetings Informal or social gatherings	Leave a very simple email or instant message for a colleague or write short simple note	Email or online form with pre-specified fields or input	Audio or text	Multi-trait analytic rubrics and scoring procedures including: • Grammatical Accuracy • Vocabulary Range & Choice • Mechanics • Content & Coherence • Workplace Communication & Sociocultural Strategies	Total available: 100 points Task 1: 25 points weighted differently across 5 traits
			Write a simple routine request to a colleague	Email or online order form	Audio or text		
			Copy dates and places from notices on notice board/flyer/a meeting/voice message	Form/meeting request email	Audio or text		
A2	Can write short texts (emails, electronic letters, meeting notes, instant messages) conveying essential information about themselves, a simple business transaction/ procedure, or a concrete business matter.		Complete form related to personal information	Registration/personnel form	Text		
			Write a short, comprehensive note of request to a colleague or a known contact in another company.	Email/Instant Messaging	Text		
			Write a very short and simple narrative or description of a business transaction	Stand-alone paragraph or email	Video or flyer (of a location for description or event – meeting, for narrative)		

B1	Can write medium-length texts (emails, electronic letters, detailed meeting notes) conveying descriptive information about others or themselves, familiar or predictable and concrete business matters.	Introductions and networking Business meetings Company or business meals	Write a letter on a topic related to personal experience and/or express opinion	Letter of application Letter of complaint/review	Text	Multi-trait analytic rubrics and scoring procedures including: • Grammatical Accuracy • Vocabulary Range & Choice • Mechanics • Content & Coherence • Workplace Communication & Sociocultural Strategies	Task 2: 35 points weighted differently across 5 traits
			Make accurate notes at a meeting	Meeting (action) minutes	Audio-visual		
			Write down some information at a business training session, (info should be more or less dictated)	Note taking	Audio		
B2	Can write medium-length to extended business texts (formal email, professional complaints, detailed meeting notes) in real-time or delayed, conveying a detailed descriptive information about others or themselves, dealing with concrete matters and familiar abstract concepts or solving problems/ responding to a request with complication	Business correspondence	Write to a venue to ask about the availability of services, for example facilities for the disabled or the provision of a special diet.	Email	Text – brochure, flyer or advertisement		
			Request/reply to routine requests for goods or services.	Email	Email - text		
			Summarize results of a survey displayed within a chart	Report	Text		
			Make notes of meeting / product demonstration for others' (not present) reference.	Note-taking	Video		
C1	Can write extended business texts (essay-length business communications, or reports, professional correspondence, detailed meeting notes) in real-time or delayed, conveying highly descriptive information about others or themselves, complex situations and opinions, concrete matters, and abstract concepts.	Business meetings Seminars Formal networking Formal Presentations Business Contracts Sales and agreements Customer relations	Write professional correspondence to an unknown contact using appropriate register and vocabulary.	Email/Fax	Text		Task 3: 40 points weighted differently across 5 traits
			Request professional services from an unknown contact.	Email	Brochures - text		
			Write a product review or a mission statement	Report/Review	Text		
C2	Can write any kind of business texts (full notes of meetings or seminars, reports, emails or letters on managing situations with complications, or presenting arguments or making claims) in real-time or delayed, with a high level of detail, expression, and accuracy.		Write a letter on any subject with good expression and accuracy.	Email	Text		
			Make full and accurate notes and continue to participate in a meeting or seminar	Note-taking	Audio		
			Make accurate and complete notes during the course of a lecture, seminar or tutorial.	Note-taking	Audio		

Optional Timer

Prompt/Scenario based on the GRASPS framework

To establish situational and interactional authenticity, the candidates must be able to:
- take notes as they read/listen to the prompt and,
- compose in an environment similar to email or instant messaging.

Text or audio input
Candidates will either read or listen to the prompt/instructions for the task here

Planning and composing panel with auto correction capacity and simulated email/instant messaging/electronic note-taking affordances

Built-in dictionary

Figure 9. CEPA Written Communication Assessment Task 1 environment and frame

6. Discussion

In order to identify and define the specific knowledge, skills, and abilities one must possess in order to function in the real-world language use context, a systematic analysis of language use in the target domain is needed. By following the conceptualization phases outlined in ECD, namely, the Domain Analysis, Domain Modeling, and Conceptual Assessment Framework, this study was able to identify, define and re-define the key multi-faceted and dynamic writing skills to be measured for 21st-century workplace written communication. The choice of using ECD as a framework was made to design a retooled assessment that measures written communication skills of candidates as it pertains to a test-taker's immediate level of English within the functioning, communicative environment of a real-life, modern workplace. The use of the ECD framework, while presenting many advantages, also came with some challenges as well.

One of the key challenges was the identification of the relevant components that needed to be analyzed in the target domain. ECD does not provide a prescriptive methodology or a list of components that should be analyzed. What needs to be investigated and how it should be analyzed depends on the particular test development project (Chapelle et al., 2018). That is, each project needs to define its own scope and breadth (i.e. the identification of relevant components and their analysis) when conducting a domain analysis, considering the specifications described by the mandate. Similarly, test designers/developers should evaluate the degree of relevance and importance of these components in terms of the target domain and the type of evidence they can gather from these components to describe the domain. Based on these parameters, in this particular project, for example, talent-acquisitions managers who work in the human resources department of either global or local corporations located in Turkey were selected because they were the stakeholders for the discovery of, screening, and recruitment of employees for their companies. In some cases, they were also the decision makers in the process of finding and delivering the necessary English language services to their current employees. Similarly, popular magazines chosen for the CEPA domain analysis were targeted primarily based on their non-domain specific, globally applicable content and their contemporary nature since the mandate established the need to define 21st-century workplace written communication knowledge, skills, and abilities. On the other hand, existing, commercially-available English-for-specific-purposes tests were excluded from the analysis, because the mandate as well as the nature of domain analysis (as it must be pertinent to the specific target language use domain) did not allow building an assessment model based on other tests which need a validity argument themselves. It is important to note that the scope and breadth of conducting a domain analysis also depends on the specificity of the target language use domain as well as the extent to which critical aspects of the target domain could be represented in the construct definition. In other words, more generic domains would require more ambitious domain analyses and test designs.

In terms of advantages, the ECD process helped us to establish a solid argument for validation of test score interpretation and use by establishing explicit links between design decisions and scores obtained from the test (Chapelle et al., 2018). For CEPA, the ECD process was used to define and operationalize the construct of writing in the context of workplace communication leading to the design and development of the CEPA Written Communication Assessment. For example, based on the domain analysis findings, we were able to expand the construct of 21$_{st}$-century workplace written communication by adding the sociocultural competency and workplace communication strategies. Since the 21$_{st}$-century workplace is now quite fully integrated into the global world via online connection, the need for understanding and sensitivity in group similarities/differences and the ability to communicate effectively across those junctions has become even more important. By identifying the most prevalent topics, general and rhetorical purposes, genres, audiences, and thus being able to set well-defined contexts, sociocultural competencies (i.e. how one approaches or responds to certain situations and different groups of people) and workplace communication strategies (i.e. aids to effective communication) can be better characterized and assessed. The ECD process allowed us to include and define these abstract concepts, which might be otherwise disregarded.

Moreover, the domain modeling included the argument that the tasks in CEPA must elicit written communication competencies, including content and coherence, grammatical structure and accuracy, vocabulary range and choice, orthographic control, sociocultural competency, and workplace communication strategies integratively. As a result of this argument, the *goal*, *role*, *audience*, *situation*, *product*, *standards* (GRASPS) model by Wiggins and McTighe (2008), which appeared in the context of the general education field was adapted, to the best of our knowledge, for first time to a language assessment context, in order to establish clear criteria for authentic proficiency-oriented performance-based tasks. Since previous research indicated some key challenges relevant to measuring L2 pragmatics (e.g. Norris, 2016; Timpe-Laughlin et al., 2015), it was found through our research that the GRASPS model can potentially address some of these issues, by allowing the measurement of sociocultural competency and workplace communication strategies in a task-based assessment model by specifying the communicative context, the specific audience and the relationship to that audience, and a clear purpose towards a realistic and relevant goal.

7. Conclusion

The CEPA Written Communication Assessment was initiated as a keystone project to create an online screening language assessment for non-native speakers' workplace communicative English proficiency in global corporations operating, for the most part, within Turkey, and as such, it was also intended, by using ECD, to gain a greater understanding of the nature of the challenge itself. The CEPA research and development team began with broad questions concerning the real-world demands of written English in the non-native Turkish workplace and were eventually guided by the mandate handed down from British Side Educational Services to select ECD because it could provide tangible concepts and a high level of detail that would, in turn, be used to inform design decisions.

This study explains, in great detail, how ECD, with its first three phases, guided the design of computer-based proficiency-oriented performance tasks that consider the constructs to be measured for 21st-century workplace written communication. ECD is a broad framework that may be applied to a variety of subject areas and contexts and by its very nature, as with any framework, has its limitations. Yet, by working within ECD, the CEPA developers were able to: (a) identify and define key assessment components/arguments, (b) shape and substantiate those arguments, (c) develop a blue print and comprehensive specifications for student, task, and evidence models, (d) operationalize these design decisions into assessment tasks, and (e) finally, delivered the tasks to candidates through tailored testing platform. All of these processes helped to establish CEPA Written Assessment and further the notion of assessment practices as evidentiary arguments in and of themselves. In keeping with the notion that assessment as argument is a cornerstone of test validation (Kane, 2006; 2012), this study reported on the use of ECD, its success as well as its limitations in implementation to inform the design and development of computer-based proficiency-oriented performance tasks that take into account the constructs to be measured for 21st-century workplace written communication.

References

Bachman, F. L. & Palmer, S. A. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press

Carr, N. (2011). *Designing and analyzing language tests*. Oxford, UK: Oxford University Press.

CEPA. Cepatest.org. (2019). Retrieved 6 February 2020, from https://cepatest.org/Web/.

Cambridge University Press. (2001). *Common European Framework of Reference for Languages (CEFR)*.

Chapelle, C., Schmidgall, J., Lopez, A., Blood, I., Wain, J., Cho, Y., Hutchison, A., Lee, H., & Dursun, A. (2018). Designing a prototype tablet-based learning-oriented assessment for middle school English learners: An evidence-centered design approach (Research Report No. RR-18-46). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12232

Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge: Cambridge University Press.

Kane, M. (2006). Validation. In R. Brennen, (Ed.), *Educational Measurement* (4th Edition), (pp. 17-64). Westport, CT: Greenwood Publishing.

Kane, M. T. (2012). Validating score interpretations and uses. *Language Testing*, 29(1), 3–17.

LinkedIn Learning Resource Center. Learning.linkedin.com. (2020). Retrieved 6 February 2020, from https://learning.linkedin.com/resources#all-topics/reports.

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd Edition), pp. 13–103). Washington, D.C.: American Council on Education.

Mislevy, R. (2011). Evidence-centered design for simulation-based assessment (CRESST Report 800). Los Angeles, CA: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Mislevy, R., & Haertel, G. (2006). Implications of Evidence-Centered Design for Educational Testing. *Educational Measurement: Issues and Practice*, 25(4), 6-20. https://doi.org/10.1111/j.1745-3992.2006.00075.x

Mislevy, R., Steinberg, L., & Almond, R. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives*, 1, 3–62. https://doi.org/10.1207/S15366359MEA0101_02

Norris, J. M. (2016). Current uses for task-based language assessment. *Annual Review of Applied Linguistics*, 36, 230–244.

O'Sullivan, B. (2006). *Issues in testing business English: the revision of the Cambridge Business English Certificates.* Cambridge: Cambridge University Press

O'Sullivan, B. (2012). Assessment issues in languages for specific purposes. *The Modern Language Journal*, 96(s1), 71-88.

MRI-Simmons. (2019). MRI-Simmons Spring 2019 Technical Guides (p. 29). MRI-Simmons. Retrieved from https://cdn2.hubspot.net/hubfs/2405078/MRI-Simmons/Technical%20Guides/MRI-Simmons%20Spring%202019%20Technical%20Guides.pdf

Timpe-Laughlin, V., Wain, J., & Schmidgall, J. (2015). Defining and operationalizing the construct of pragmatic competence: Review and recommendations (ETS Research Report No. RR-15-06). Princeton, NJ: Educational Testing Service. https://onlinelibrary.wiley.com/doi/pdf/10.1002/ets2.12053

Wiggins, G. & McTighe, J. (2008). *Understanding by Design Professional Development Workbook*. Alexandria, VA: Association for Supervision and Curriculum Development.