Author(s): Passera, S., Kankaanranta, A. Louhiala Salminen, L.

Title: Diagrams in contracts: Fostering understanding in global business communication

Year: 2017

Version: Author accepted / Post print version

Please cite the original version:


Rights: © 2017 IEEE. Reprinted with permission.

"In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of Aalto University’s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink."

This publication is included in the electronic version of the article dissertation:
Passera, Stefania. Beyond the wall of contract text - Visualizing contracts to foster understanding and collaboration within and across organizations. Aalto University publication series DOCTORAL DISSERTATIONS, 134/2017.
Diagrams in Contracts: Fostering Understanding in Global Business Communication

Stefania Passera, Anne Kankaanranta, Leena Louhiala-Salminen

ABSTRACT

Research problem: Business-to-business contracts are complex communication artifacts, often considered “legal stuff” and the exclusive domain of lawyers. However, many other stakeholders without a legal background are involved in the negotiation, drafting, approval and implementation of contracts, and their contributions are essential for successful business relationships. How can we ensure that all stakeholders in the global business context – whatever their native language or professional background – easily and accurately understand contract documents? This study suggests that integrating diagrams in contracts can result in faster and more accurate comprehension, for both native and non-native speakers of English. Literature Review: We focused on the following research topics: 1) how to integrate text and visuals in order to create more effective instructions, since we conceptualize contracts as a type of business instructions; 2) cognitive load theory, as it may help explaining why contracts are so hard to understand and why text-visuals integration may ameliorate their understandability; 3) cognitive styles, as individual differences may affect how individuals process verbal and visual information, thus allowing to explore the limitations of our suggested approach; 4) the English lingua franca spoken by business professionals in international settings, their needs and challenges, and how pragmatic approaches are needed to ensure successful communication. Methodology: We conducted an experiment with 122 contract experts from 24 countries. The research participants were asked to complete a series of comprehension tasks on a contract, which was provided in either a traditional, text-only version or in a version that included diagrams as complements to the text. In addition to measuring answering speed and accuracy, we asked the participants to provide information about their educational background, mother tongue, perceived mental effort in task completion, as well as to fill out the
Object-Spatial Imagery and Verbal Questionnaire to assess their cognitive style. **Conclusions:** We found that integrating diagrams into contracts support faster and more accurate comprehension of contracts; unexpectedly, legal background and different cognitive styles do not interact with this main effect. We also discovered that both native and non-native speakers of English benefit from the presence of diagrams in terms of accuracy, but that this effect is particularly strong for non-native speakers. The implication of the study is that adding diagrams to contracts can help global communicators to understand such documents more quickly and accurately. The need for well-designed contracts may open new opportunities for professional writers and information designers. Future research may also go beyond experimental evaluations: by observing this new genre of contracts *in vivo*, it would be possible to shed light on how contract visualizations would be perceived and interpreted in a global communication environment.

**INTRODUCTION**

Think about contracts: long, complex, demanding documents written in “legalese”. Not really an archetype of effective, clear and easy professional communication. Misunderstandings may lead to litigations, which damage businesses in terms of wasted resources and reputation. For instance, BellAliant and Rogers Telecom, two Canadian firms, went through a bitter 18-month dispute, worth one million Canadian dollars, on whether BellAliant was allowed to terminate a contract by just giving notice, or whether the contract was iron-clad until the end of the agreed period. The source of confusion boiled down to a single comma in the English version of the agreement, which suggested the first interpretation was the correct one, while the French version of the agreement (Canada is a bilingual country) supported the latter interpretation [1]. Plain language scholars have been addressing the shortcomings of traditional legal writing for years. Even attorneys and legal scholars have been repeatedly warning that legalese is a waste of money for businesses and
governments – in the form of claims, administrative and compliance costs, and time lost in writing, reading, and managing confusing documents [2]–[4].

It is also erroneous to think of contracts only as “legal stuff”, the exclusive, specialized domain of lawyers [23]. In today’s global, networked economy contracts represent key business instruments for companies to synchronize, collaborate and undertake new ventures. They are used in all spheres of business, from such narrow aspects as the working conditions of a certain employee, to large corporate wide issues such as mergers and acquisitions. Moreover, after signature, contracts need to be implemented by responsible teams, who need to carefully follow contracts as instructions and translate them into concrete actions. Thus, we see contracts as a business communication genre. Managers and executives, in addition to lawyers, are key stakeholders of contracts, and should be able to understand what contracts say and how they impact day-to-day business.

The challenge is even bigger in international business, as parties do not share the same language and may lack common interpretive frames. The highly formal register of English “legalese”, which is modeled around the English used by its legally-trained native speakers (NS), is the standard language of international contracts. However, this is not the real-life ‘English’ used in international business, which is better characterized as a *lingua franca*, typically used by non-native speakers (NNS) of English – to get the work done. As showed in the BellAliant vs. Rogers example, the long and syntactically complex sentences typical of legalese may be ambiguous even for NS managers. In light of the concept of global communicative competence (GCC) [6][7] – which, drawing on the lingua franca aspect of English, stresses pragmatic communication skills, rather than flawless native proficiency – we propose that the threat of ambiguity to international contracts must be addressed by going beyond the use of textual language. After all, the legal principle of freedom
of contract does not limit in any way contract formats, style and aspect (even a verbal contract is valid, although it may be hard to enforce it).

In this study we show how the comprehension of agreements among international contract professionals can be enhanced by presenting information in a more visual fashion: contracts complemented with explanatory diagrams, and presented in clearly structured layouts. The argument in favor of visualizations as a way to clarify the meaning of complex textual information is based on these areas of earlier research: the benefits of integrating texts and visuals in instructional texts (and how contracts can be intended as a particular type of instructions); cognitive load theory and the effects of individual cognitive styles on learning; the specific needs of professional NNSs of English for pragmatic solutions and practices that can improve their overall communicative competence.

The overall problem of the present study can be formulated as follows: can a visually-enhanced style of communication help professionals to understand complex contractual information? This leads to our first two research questions:

(1) To what extent can professionals’ understanding of contracts be enhanced by employing diagrams in such documents?

(2) To what extent can diagrams in contracts reduce the time taken to understand these documents?

If we find that diagrams indeed support contract comprehension, we want to investigate whether their presence equally benefits individuals with different cognitive styles, language abilities, and knowledge of specific jargon. For instance, could a visual approach increase the communicative
competence of NNSs dealing with English legalese, and reduce the gap with NSs? This leads to two further research questions:

(3) Which individual characteristics (e.g. cognitive style, educational background, age, language…) may affect the extent to which a professional benefits from diagrams in contract documents?

(4) In light of the possible effects of individual characteristics, what are the implications of using diagrams in contracts in an international business setting?

The remainder of the article is structured as follows. Firstly, in the Literature Review we describe our theoretical orientation, the selection process for the literature, and then present the identified key topics. At the end of section, eight hypotheses are suggested. Then, we introduce the methodology of the study, after which the results are presented and discussed. Lastly, we address limitations and suggestions for future research.

**LITERATURE REVIEW**

Insights from past research motivate our interest in complementing international contracts with explanatory diagrams as a way to make them more understandable. The reviewed studies contribute to formulate the hypotheses of our study, and identify the relevant variables to investigate. This section begins with an introduction of our theoretical orientation, followed by sections focusing on visual representations in instructions, Cognitive Load Theory and instructional design, cognitive styles and the challenges of English legalese within the frame of Global Communicative Competence.

**Theoretical Orientation**
This study is multidisciplinary in nature, combining theories of visual communication, cognition and language. Although the study focuses on business contracts, it looks at them from the perspective of design and communication, rather than from the tradition of legal studies. Communication effectiveness in a global business setting is the main concern, and we zoom in on how and why complementing texts with diagrams enhances information search and comprehension.

Traditionally, the efforts on clarifying and simplifying contracts have aimed at “lifting the fog of legalese” [5]. Rethinking legal and bureaucratic language has been (and still is) the main battle of the proponents of Plain Language – "language that reflects the interests and needs of the reader and consumer rather than the legal, bureaucratic, or technological interests of the writer or of the organization that the writer represents" [8, p.7]. The focus on visual communication is thus a novelty in contracts, even though a number of plain language authors have acknowledged the key role of information design in communicating content clearly [9][10]. A small, but growing, body of legal scholarship has been proposing the use of visual representations – for instance diagrams [92][93], photographs [93], comics [94], icons and document design [8][77] – as means to increase the understandability of contracts [8][23][26][81][91]-[95] and other legal documents [25]. One issue in such literature, however, is that several studies do not rely on actual user testing [23][26][94]. The studies that provide some quantification of the comprehension benefits of visual communication do so in an insufficiently rigorous manner, e.g. too small samples [81][92][95], and insufficient statistical analysis [81][95]. A theoretical explanation why and to what extent visual communication supports contract comprehension is also missing.

Our concern with user-centeredness and effective communication led us to deepen our understanding of the individual and the context in which s/he communicates. Although contracts encode complex information, understanding is not only dependent on the intrinsic ease or difficulty of the document: the cognitive resources available to the readers should also be accounted for. For
this reason, we looked into cognitive theories about learning and comprehension, and into individual cognitive variables affecting the processing of verbal and visual information. We focus in particular on cognitive load theory, which acknowledges the importance of information presentation in promoting or hindering understanding, and on individual cognitive styles, i.e. how each individual prefers to process, encode and express information. Additionally, we reviewed literature on communicative competence in global settings, and, since our study focused on the comprehension needs of contract experts operating in an international setting, we paid particular attention to the research streams of English as (Business) Lingua Franca and English for Specific Purposes.

The reviewed literature also affected our methodological choices: we designed the study as an experiment, but paired it with a self-reported questionnaire. The experimental setup follows the tradition of evaluating and comparing design solutions in their ability to satisfy users’ needs, and improve comprehension performance; the questionnaire allows us to introduce individual cognitive and language abilities as variables in the study, and understand how they interact with the presentation format of the contract, and affect comprehension.

Selection of Literature for the Review

To address the questions and concerns of this article, we opted for a convenience review of past studies focusing thematically on 1) how to integrate text and visuals in order to create more effective instructions (we conceptualize contracts as a type of business instructions) 2) cognitive load theory, as it may help explaining why contracts are so hard to understand and why text-visuals integration may ameliorate their understandability; 3) cognitive styles, as individual differences may affect how individuals process verbal and visual information, thus allowing us to explore the limitations of our suggested approach; 4) the English lingua franca spoken by business
professionals in international settings, their needs and challenges, and how pragmatic approaches are needed to ensure successful communication.

Visual representations in instructions

A pragmatic approach to effective communication in contracts leads us to conceptualize contracts as instructions on how to correctly deliver a given promise to the other party. This metaphor is not completely new, as in fact legal scholars have previously proposed contracts as blueprints for [business] performance [24, p.92], user’s guides and instruction manuals [25, p.21] and visible scripts for the parties to follow [26, p.57]. Taking inspiration from this metaphor, we explore research on multimodal instructions, which has previously tackled the problem of supporting users’ comprehension by integrating texts and visuals. Most research on instructions focuses on how to effectively provide guidance to use appliances [27], software [28][29], to perform procedures [30][31], and to solve abstract/logical problems [32][33]. Even though this constitutes a key difference from contracts – where the to-do actions are more abstract (e.g. when and how to pay the supplier; how to calculate liquidated damages for a delay in product delivery) – we believe that the key results of this research area could inform and be applied to our area of inquiry.

Firstly, visuals in instructions – such as screenshots, technical illustrations, photos and diagrams– offer visual relief from dense pages full of text [29], which users are not enticed to read [34]. Moreover, visualizations reduce complexity, for instance by pre-selecting and giving salience to the most important details needed to accomplish the goal [35]; making more explicit relationships which need to be inferred from the text [27]; presenting procedures in an easy-to-follow step-by-step approach [35]; illustrating more saliently alternative routes and conditions, as well as conditional information (if X then Y) [35][36]; by offering a visual reference that helps users to verify that they are proceeding correctly and recover from errors [37].
Secondly, visual representations facilitate readers in creating a mental model of themselves in action, executing the procedure [38]. In respect of successful mental model creation, research has shown the importance of the presentation order in which instructions should be carried out [39] and the type of perspective – e.g. user-centric or system-centric – to be adopted in visualizations [30][40]. While mental models can be created also from purely textual information, this requires more effort and induces a heavier cognitive load compared to a multimodal approach [41]: visual representations allows for a more direct construction of mental models, since external and internal representations share a quasi-spatial nature (e.g. [27][42])

However, this does not mean that visual representations alone are sufficient to ensure comprehension. Cumulative evidence shows that the most effective instructions combine text and visuals [41], either with explanatory visuals accompanying text [43][44], or as visually structured diagrams which include text, such as flowcharts, decision trees, and tables [35][45][46]. Poe Alexander [47] summarized the benefits of integrating text and visuals as faster task completion, less comprehension mistakes, faster learning to use a device, and more positive user attitudes towards the instructions. Several scholars have also pointed out how visual elements play an indispensable rhetorical role in relation to text: it better organizes and structures the content [48], it guides readers’ attention to precise topics and arguments [49], and it affects interpretation of the overall document [50] by providing hints about its tone, credibility, function and relevance. [51].

Lastly, several studies indicate that understanding and applying instructions is a cognitively intense activity, which imposes high demands on the working memory [41], as readers need to explore the document and find relevant information, understand it, integrate it with previous and situational knowledge, and finally apply the instruction. Spatially integrating texts and visuals allow reducing
the cognitive load by creating a degree of redundancy, which is beneficial to comprehension [40], and by reducing the need to split attention between sources of information located in different places [52].

In the next subsection we take a closer look at cognitive load theory, which is concerned with the learning of complex information and how learners often struggle to process the many elements needed for understanding, because of the capacity limits of the working memory system.

**Cognitive Load Theory and instructional design**

Cognitive load theory (CLT) [53][54] postulates that too complex and/or abundant information overload our cognitive systems, as capacity limitations prevent adequate processing of all inputs, and, ultimately, learning. CLT is based on three assumptions: human working memory (WM) possesses different processing systems for visual and verbal information (*dual channel assumption*; e.g. [56][57]), each channel has a limited processing capacity (*limited capacity assumption*; e.g. [53][56]), and meaningful learning requires substantial cognitive activities in scanning, organizing and integrating information from different channels (*active processing assumption*; e.g.[58]). Additionally, as Schnotz and Banner [33] point out, comprehension is an active, goal-oriented process: humans construct multiple mental representations with the goal of accomplishing a specific task (e.g. being able to perform a calculation; solving a problem; behaving accordingly to a rule…). Different ways to present the same content affect how we construct it mentally, and these mental representations may facilitate or hinder problem resolution.

The superior performance associated with the integration of visuals and texts in instructions, compared to texts or visuals alone, can be explained in light of CLT. For example, a long running text may feel difficult to understand because it is processed through one channel only, which has
limited capacity; if the same information is presented as a diagram information processing is routed to different channels, as text conveys separate bits of meaning, while the visual communicates which relationship exists between these bits. The cost of interpretation of such relationships is offloaded to the external representation, as they simply need to be “read-off” instead of being inferred effortfully [59].

In the case of contracts, one might argue that they are difficult because they carry intrinsically complex messages; however, information presentation matters. CLT conceptualizes cognitive load as being composed by three separate elements, which are differently affected by changes in how the information is designed [53]: intrinsic, extraneous and germane load (summarized in Figure 1).

**FIGURE 1**

The three components of cognitive load, according to [53].

<table>
<thead>
<tr>
<th>Total Cognitive Load</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Intrinsic Cognitive Load</td>
</tr>
<tr>
<td>- Processing inherent to the complexity of the content</td>
</tr>
<tr>
<td>- Cannot be affected by instructional designs</td>
</tr>
</tbody>
</table>

| 2. Extraneous Cognitive Load |
| - Processing imposed by how to content is presented |
| - Should be decreased, because prevents learning |
| - Can be affected by instructional designs |

| 3. Germane Cognitive Load |
| - Processing dedicated to schema formation and automation |
| - Should be increased, because leads to learning |
| - Can be affected by instructional designs |

*Intrinsic cognitive load* derives from the inherent complexity of the information, and it mainly depends on whether the different elements to be processed can be attended individually or need to be attended all together [60]. For example, performing a series of simple, unrelated numerical
addition does not cause overload, while solving a literal equation does, as it requires knowing where and how to apply several algebraic rules, which need to be used in concert. Intrinsic load usually cannot be affected by instructional means. *Extraneous cognitive load* is caused by the manner in which information is presented, and thus its reduction through different designs has traditionally been the focus of instructional designers. Intrinsic and extraneous cognitive loads are additive and both impose demands to WM, and extraneous load interferes with learning, especially when information is intrinsically complex [53][60]. *Germane cognitive load* is the processing devoted to schema formation and automation, that is, the cognitive processing which actually positively contributes to understanding and learning. Usually there is an opposite relationship between extraneous and germane cognitive load [61]: for example, devoting much mental effort to understand something unclearly written will not result in increased learning (extraneous load), but thoroughly processing and actively working through an example will (germane load) [62][63].

Comprehension errors and slow reading/task completion are typical symptoms of cognitive overload [64]. However, in order to exactly assess and compare the efficiency of alternative instructional designs in supporting learning, [65] and [66] suggest the need to take into consideration both performance and mental effort: this is because learners may maintain a good performance in cognitively demanding tasks by simply making a bigger mental effort. Taking in consideration both performance and mental effort allows to distinguish between desirable and undesirable instructional designs: low-efficiency instructions will require high mental effort and still yield poor performance, while high-efficiency instructions will yield good performance with little mental effort [66].

These considerations open the possibility to improve the efficiency of contracts – in terms of learning and comprehension – by changing how they are designed. In the previous subsection, we
saw how effective instructions integrate texts and visuals. There is considerable evidence for these results also in educational psychology literature, which show how diagrams – which integrate visual representations with texts – support learning (e.g. [32][67][68]). However, other studies suggest that the effects of visual presentation modes in enhancing learning are strong for novices, but tend to disappear and even reverse as expertise grows (e.g. [68][70]–[73]); further, some studies conclude that only high visual-spatial learners can fully benefit from integrating visual and verbal materials [32][74]. Our study needs thus to assess whether adding diagrams in contracts result in higher efficiency instructions, and whether the hypothesized benefits are limited to only some groups of readers.

**Cognitive styles**

In addition to cognitive load theory, individual cognitive styles should be taken into consideration when optimizing information for learning and understanding. ‘Cognitive style’ is a term used to describe the consistent way in which an individual acquires and processes information (e.g. [75]–[76]). According to Ausburn and Ausburn, [76] cognitive styles are stable characteristics, resistant to change by training and in no way correlated with intellectual ability: if a learner struggles with learning, the problem might be in how the information is presented, and a different approach to conceptualizing, encoding and displaying information can result in successful learning. It is thus recommendable (and more inclusive) to provide alternative complementary representations. For instance, an individual scoring low on verbal ability might struggle to interpret a complex text and integrate its parts; if we were to complement this text with a diagram, and were the learner to possess high scores on visual cognitive styles, the problematic information processing would be supplanted by an appropriate instructional design, thus allowing understanding [76].
While several labels and tools to describe and assess cognitive styles exist, it has been suggested [78]–[80] that they can be all grouped under two main cognitive style dimensions: the wholist-analytic style dimension – whether an individual processes information as wholes or in parts – and the verbal-imagery style dimension – whether an individual better processes and represents information as words or images – [79]. Given the focus of this paper, we will concentrate on the verbal-imagery style dimension.

Recent developments on the study of the verbal-imagery cognitive style have shown how this is not, as previously believed, a bipolar continuous construct, according to which individuals are categorized as either verbalizers or visualizers [69][82]. In fact, three relatively independent dimensions exist: object-imagery (ability in building and processing vivid, colorful, detailed images of individual objects); spatial-imagery (ability in representing and processing schematic images, spatial relationship between objects, movement and spatial transformations); and verbal (ability in verbal, logical, sequential and analytical reasoning) [69][75][83]. From this follows that an individual with a distinctively favorite cognitive style may or may not struggle with other types of information representations, as a high score on one dimension does not imply low scores on the other two.

In the same way in which the savvy global communicator has to be sensitive to different values and cultures, a higher appreciation of different cognitive styles may lead to more effective communication – given that contract may already be intrinsically complex, and even more so for a non-native speaker of legalese or English. In the next section we will lastly introduce the lingua franca spoken by global business professionals, their practical communication needs, and how to support such needs through a pragmatic approach.
The challenges of English legalese within the frame of Global Communicative Competence

In the past twenty years, English has become the shared language of most international organizations, both at the macro level of the organization as “the corporate language”, and the micro level of the individual professionals working within the organization (e.g. [11]–[14]). For this reason, it is necessary to see what this extremely fast spread of English to international companies, and hence contracting, might mean from the point of view of communicative competence.

The majority of the international professionals who use English as a shared language are NNSs, making use of a resource labeled *English as a Lingua Franca* (ELF) or more specifically, *English as Business Lingua Franca* (BELF). This resource-based view of language differs from, e.g. *English as a native* language (ENL), *English as a second* language (ESL, e.g. Singaporean English) or *English as a foreign* language (EFL) (e.g. [15]), which all reflect the conceptualization of language as an independent entity and a pre-existing system.

ELF research looks at language differently: ELF is not “owned” by anyone, nor directly linked to a particular cultural context (e.g. [16]); instead the ELF approach assumes a constructionist view that emphasizes the particular ELF situation and its participants. ELF communicators tend to avoid complex lexis or structures, and they incorporate their own, native communicative practices and situation specific requirements into the situation at hand [17].

The conceptualization of ELF as an independent linguistic resource also entails that ELF speakers are not assumed to imitate ENL as closely as possible, but rather identify themselves as speakers in their own right (e.g. [18][19]). Neither is ELF considered deficient in comparison to ENL albeit rather different; nor is resorting to other languages considered during interaction regarded as evidence of a gap in the knowledge of English but rather a bilingual pragmatic resource (e.g. [20]).
Drawing from ELF research, but contextualizing their study in international business, Louhiala-Salminen et al. [21] identified BELF (Business ELF) to emphasize the significance of the domain – the “B” – with its goal-oriented nature and shared business fundamentals. When the employees of globally operating companies attend to their daily practices – including contracting –, they thus draw from their (1) BELF resource, but also need (2) business knowhow and (3) multicultural competence to do their job. Although ‘English’ in BELF is not grammatically perfect, the BELF competence involves the use of highly specialized, shared terms, concepts and genres to be able to adapt to the business situation at hand. It also includes the element of strategic skills; as shared understanding cannot be taken for granted, successful BELF communicators ask questions, repeat utterances and use more than one channel or mode to achieve shared understanding. Finally, clarity, directness, and politeness are essential.

As was pointed out earlier, although legal English causes challenges for both native and non-native speakers of English because of its highly specialized, formal register, it is particularly demanding for non-native speakers (e.g. [22]). The needs of NNSs are addressed in the paradigm of English for Specific Purposes (ESP) focusing on the pedagogy of teaching specialized language, in this case legal English. Although the English needed is defined as ‘legal’, it is still conceptualized as the English of the NS, whose language use the learner is supposed to imitate. The difference to the ELF/BELF paradigm is distinct: in ELF/BELF the language user is a professional using the language at work, whereas in the ESP paradigm, the focus is on learning a language.

Lawyers can be seen as “native speakers of legalese”, but they are far from being the only ones involved with contracts. As suggested by Haapio [23, p. 54–55], several contract users exist at the same time, in different contexts of use, with different goals and needs. On the one hand, traditional
legal scholarship focuses on judges, arbitrators, litigating lawyers and authorities, and the prototypical scenario of use is the court: the main interest is to predict how courts might interpret contracts, so contract drafting is seen as a communication and persuasion activity happening between legal professionals. This is the paradigm where contracts written in “native legalese” are the unchallenged norm, and where communication might break down if the correct grammar, registers, rhetorical devices and norms are not followed. On the other hand, Haapio identifies a managerial-legal paradigm, focusing on the needs of in-house counsels, managers (sales, procurement, risk management), financial officers and engineers, where the main driver is to reach business objectives through commercial relationships, proactively minimizing legal and business risks and always balancing them with rewards. These contract users, who may as well be English native speakers, are the “non-native speakers of legalese”, for whom communicative success can be assessed according to criteria similar to those relevant to Global Communicative Competence [6]: multicultural (in this case, also multidisciplinary) competence, or the awareness and sensitivity of the communicator towards how different cultures/disciplines get things done; specific business knowhow; and BELF competence.

In conclusion, findings on BELF and GCC can be used as a lens to look at any context where imposing a native English language model (e.g. legalese) on non-native speakers (e.g. internationally operating managers) not only causes difficulties in comprehension, but does not serve the practical goals of the professionals, i.e. getting their work done.

**Conclusions from Literature Review and Hypotheses**

The earlier studies discussed in the Literature Review above suggest a way forward to answer to our research questions. Diagrams and other strategies to integrate texts and visuals offer a way to provide effective instruction to individuals tasked to accurately follow a procedure – a scenario akin
to using contract in day-to-day business. The difficulty in understanding contracts accurately and quickly, moreover, hints at a problem of cognitive overload, in particular the presence of extraneous load, as typically contracts do not follow any best practice in document design and have been criticized by plain language supporters for their use of legalese. Changing the design of the information is a way to reduce cognitive load, and support better understanding. We also established that the risks associated with ineffective contract communication are particularly high in international settings, as the English Lingua Franca used by non-native speakers differs from the convoluted register of native language-informed English legalese. Reframing and redesigning contracts as instructions, in our view, is a solution well in line with the pragmatic approach to communication needed by global communicators. However, given the novelty of this approach in contracts, we need to discover the extent of benefits that may be brought forward by integrating text and visuals, and how this solution may also accommodate individual characteristics, such as cognitive styles, language and professional proficiency – of native and non-native contract professionals alike. Below, we formulated a series of testable hypotheses that will help us find answers.

Based on the consistent evidence indicating that instructions integrating visuals and texts support better comprehension, and on the predictions of CLT (errors and slow reading/answering are symptoms of cognitive load; part of the cognitive load – the extraneous one – can be reduced through design solutions, arguably reducing errors and time taken), we put forth two initial hypotheses:

**H1. Contracts including diagrams allow for faster information finding than traditional, text-only contracts.**
H2. *Contracts including diagrams allow for more accurate comprehension than traditional, text-only contracts.*

Although improved and rapid comprehension is the ultimate practical goal for contract readers, this alone cannot tell us whether different designs actually reduce cognitive load. When performing a difficult task, people may still achieve good performance by investing even more mental effort as a reaction to an increase in cognitive load [65]. For a more meaningful comparison of different communication solutions, we need to assess how mentally efficient they are – meaning how much mental effort they require in order to achieve a certain level of performance. Estimating efficiency allows us to determine whether one document requires more mental effort than the other (meaning, it imposes a higher cognitive load), and what are the performance outcomes of such mental effort investment.

H3. *The mental efficiency of the version of the contract including diagrams is higher than that of the text-only version, meaning that participants require less mental effort to obtain an equal level of performance*

An appreciation for individual and different ways to process, learn and represent information – as shown in studies on cognitive styles – might warrant a more fine-tuned assessment of the benefits that visualizations might bring to comprehension. As described in Appendix 1, the ‘visual contract’ used in our study includes explanatory diagrams, which complement rather than substitute the text. We know that spatial-visualizers learners benefit more from multimodal instructions [74], and we assume that verbalizers may easily process either version, given the high content of text. Predominantly object-visualizers may comparatively improve their performance with a more visual contract, but since contract understanding requires analytic skills, they may still have a
comparatively weaker performance. However, since verbal and visual skills are not mutually exclusive, we need to study all possible interactions of cognitive styles and treatment in an explorative manner, as it is difficult to predict their final effects on performance, and whether a specific class of learners is the fastest and most accurate. Thus:

**H4a. Cognitive styles interact with experimental treatment, and with each other:**

- verbalizers are equally fast with either contract, and are fastest with the textual contract;
- spatial-visualizers are the fastest with contracts including diagrams;
- object-visualizers become faster with contracts including diagrams than textual ones, but remain the slowest performing group.

**H4b. Cognitive styles interact with experimental treatment, and with each other:**

- verbalizers are equally accurate with either contract, and are most accurate with the textual contract; spatial-visualizers are the most accurate with contracts including diagrams;
- object-visualizers become more accurate with contracts including diagrams than textual ones, but remain the least accurate group.

Lastly, we need to discriminate the effects of visualization for native and non-native speakers of English, the globally shared language used in international contracting. Arguably, a NS (especially if she has a legal background) should be more proficient in understanding contracts than a NNS. Studies on CLT have consistently demonstrated an “expertise reversal effect”, in which more knowledgeable readers do not benefit from additional instructional aids, and actually experience higher cognitive load because of the presence of redundant, more explicit information (e.g. [68][70]–[73]). Although NNSs using English – or rather BELF – in their everyday work may be highly competent global communicators, English legalese is not within that type of pragmatic
competence. This would mean that as non-experts they are more likely to benefit from more explicitly presented information.

**H5a.** *Language interacts with experimental treatment:* native speakers of English are faster and more accurate than nonnative speakers with the traditional version of the contract (text-only), while nonnative speakers perform faster and more accurately with the version of the contract including diagrams.

**H5b.** *Profession-specific knowledge interacts with experimental treatment:* subjects with a legal background are faster and more accurate with the traditional version of the contract (text-only), while subjects without legal background perform faster and more accurately with the version of the contracts including diagrams.

**H5c.** *Language interacts with profession-specific knowledge and experimental treatment:* native speakers of English with a legal background are the fastest and most accurate with the traditional version of the contract (text-only), while nonnative speakers without a legal background are fastest and most accurate with the version of the contract including diagrams.

**METHODOLOGY**

In this section we explain how we designed our study – an experiment, complemented with a questionnaire aimed at collecting control variables – and analyzed the results. Firstly, we introduce our choice of a research methodology and how we sought out participants. Then we review how we have collected data, administered the study, and analyzed the data. Lastly, we close the section by presenting how we assured the reliability and validity of the study.
Choice of a research methodology

In line with the research traditions analyzed in our literature review, we opted for an experimental design, integrating it with a self-reported questionnaire. This combination of methodologies allowed us to collect objective comprehension performance data in concert with the use of a given contract version (experiment), and to integrate the results with individual psychometric and self-reported data. The experiment and the questionnaire were previously piloted during two smaller-scale studies [2 references omitted for peer review] and their design and measures were discussed with several researchers not involved in the study.

Participants

We sought to recruit participants who:

1) Currently held a position in which they routinely wrote, negotiated, or implemented international contracts
2) Were currently involved in international deals
3) Represented both sell-side or buy-side of business transactions,
4) Represented both native or non-native English speakers,
5) Represented different organizational functions involved with contracts (commercial and contract management, sales, procurement, legal)
6) Had a university degree, so as to avoid low-literacy as a confounding factor.

The participants were thus recruited among the members of the IACCM (International Association for Contract and Commercial Management), a global, industry-independent professional association promoting best practices and standards in contract and commercial management. At the time of the
study, IACCM had 34,000 members from 158 countries, working in private and public organizations alike. Collaborating with IACCM allowed us to access a highly diverse community, without the need to individually contact a great variety of organizations worldwide.

IACCM sent out our request for participants as part of their newsletter, and we recruited 122 participants. Even though the newsletter is sent to all IACCM members, their webinars (which last approximately 1 hour, just like our study) are most typically attended by 200 people. Since 200 people was, most likely, best participation rate we could aspire too, we calculate response rate on its basis (62% response rate). As an incentive, the volunteers were offered the possibility to receive their detailed, individual results of the experiment, how they compared with the overall average, and an assessment of their cognitive style.

**How data was collected**

Here, we describe the methods though we employed to collect quantitative data for our study. This section starts with a description of the Experimental Task, followed by Instrument, where we explain the questionnaire we used to collect control variables.

**Experimental Task**

The experimental task consisted of answering six comprehension questions about a contract, randomly assigned to the participants in its traditional, text-only version or in a visually enriched version which featured explanatory diagrams, clearer layout and typography.

The contract used in the experiment was a B2B agreement on the purchase of machinery and equipment. We obtained the permission from a European metal and engineering company ([details omitted for peer review](#)) to use the anonymized version of their B2B machinery and equipment
purchase agreement, as we were aware that they had asked a professional graphic designer to help them develop a new, visualized version of it. The visual version and the older, text-only version of the agreement had exactly the same text, and differed only in the way in which it was presented. The order and numbering of the clauses was also the same. Sensitive details such as the names of the contracting parties, prices, places and scope of delivery were fictionalized.

In this agreement, diagrams are presented below the relevant clause they seek to clarify. The diagrams can be categorized mostly as timelines or charts representing a process. All diagrams contain textual labels, repeating terms from the clause next by, so as to connect texts and diagrams through a strategy of repetition [27]. The layout of the agreement is a single column, with clause headings on the left-hand side to facilitate their search. Compared to the text-only version, there are wider margins and more white space on the page. In addition to the diagrams, key information (e.g. deadlines, prices) is displayed in boldface in the text of the agreement, so as to give it more salience. Excerpts from the visualized contracts are shown in Appendix 1.

Comprehension performance on the experimental task was measured as follows:

*Answering Speed.* The sum of the time taken to answer the 6 comprehension questions. The time taken for each answer was measured individually, and then summed up.

*Answering Accuracy.* The sum of correct answers. Each correct answer was assigned one point. Partially correct answers were graded with 0.25, 0.5 or 0.75 points, depending on the magnitude of the mistakes and imprecisions (see grading rationale in Appendix 3), and thus the scores vary between 0 (all wrong) and 6 (all correct). In order to avoid bias, two colleagues not involved in this study and blind to the experimental condition in which the answers were given were asked to grade
the accuracy of the answers. The graders agreed on 98% of the answers, and were asked to agree and re-assess together the answers that they had graded differently.

**Instrument**

Before the experiment, the participants had to fill in a questionnaire, thus providing data on the following variables (see details in Table 1): 1) age 2) gender 3) educational background 4) legal background 5) nationality 6) level of English language proficiency 7) native language 8) current professional role 9) years of experience in working with contracts 10) industry 11) object-spatial-verbal cognitive style 12) mental effort. While most of these were control variables, we will explain ‘Native language’, ‘legal background’, ‘object-spatial-verbal cognitive style’ and ‘mental effort’ in more detail, as they were instrumental in our analysis and hypothesis testing.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Age</td>
<td>Continuous variable.</td>
<td>Control</td>
</tr>
<tr>
<td>2) Gender</td>
<td>Categorical variable: female (1); male (0).</td>
<td>Control</td>
</tr>
<tr>
<td>3) Educational background</td>
<td>Categorical variable: Engineering, manufacturing, construction; Law; Business and administration; Life sciences, physics, math, computer science; Education; Arts; Humanities; Social sciences; Other.</td>
<td>Background information</td>
</tr>
<tr>
<td>4) Legal background</td>
<td>Categorical variable: yes (1); no (0) (variable obtained from &quot;Educational background&quot;).</td>
<td>Analysis of main and interaction effects</td>
</tr>
<tr>
<td>5) Nationality</td>
<td>Categorical variable: US; UK; Canada; Australia; Germany; Italy; Spain; Finland; Norway; Ireland; Czech Republic; Denmark; Poland; France; Hungary; Netherlands; Ukraine; Croatia; China; India; Pakistan; Senegal; South Africa; Nigeria.</td>
<td>Background information</td>
</tr>
<tr>
<td>6) Level of English language proficiency</td>
<td>Categorical variable: No proficiency (0); Elementary proficiency - (1); Elementary proficiency (2); Elementary proficiency + (3); Limited working proficiency (4); Limited working proficiency + (5); Professional working proficiency (6); Professional working proficiency + (7); Full professional proficiency (8); (9); Full professional proficiency + (10); Native or bilingual proficiency (11).</td>
<td>Background information</td>
</tr>
<tr>
<td>7) Native language</td>
<td>Categorical variable: English (1); Other (0).</td>
<td>Analysis of main and interaction effects</td>
</tr>
<tr>
<td>8) Current professional role</td>
<td>Categorical variable: Contract management; Legal; Sales/Commercial; Procurement/Strategic sourcing; Combination sales/sourcing; IT; HR/Recruitment/Placement; Finance; Operations; Academic; Other.</td>
<td>Background information</td>
</tr>
<tr>
<td>9) Years of experience in working with contracts</td>
<td>Continuous variable.</td>
<td>Control</td>
</tr>
<tr>
<td>10) Industry</td>
<td>Categorical variable: IT/Telecoms; Oil/Gas/Minerals/Utilities; Banking/Insurance/Finance; Aerospace/Defense; Technology; Services/Outsourcing/Consulting; Transportation/Logistics; Automotive; Engineering/Constructions; Electronics; Healthcare/Pharma/Chemicals; Public sector; Manufacturing/Processing; Retail; Other.</td>
<td>Background information</td>
</tr>
<tr>
<td>11 a) Object-visualizer cognitive style</td>
<td>Continuous variable: value ranging from 0 to 5; obtained as average of the 15 “O” items in the OSIVQ questionnaire.</td>
<td>Analysis of main and interaction effects</td>
</tr>
<tr>
<td>11 b) Spatial-visualizer cognitive style</td>
<td>Continuous variable: value ranging from 0 to 5; obtained as average of the 15 “S” items in the OSIVQ questionnaire.</td>
<td>Analysis of main and interaction effects</td>
</tr>
<tr>
<td>11 c) Verbalizer cognitive style</td>
<td>Continuous variable: value ranging from 0 to 5; obtained as average of the 15 “V” items in the OSIVQ questionnaire.</td>
<td>Analysis of main and interaction effects</td>
</tr>
<tr>
<td>12) Mental effort</td>
<td>Categorical ordinal variable ranging between: very, very low mental effort (1) and very, very high mental effort (9)</td>
<td>Analysis of main and interaction effects</td>
</tr>
</tbody>
</table>

*English language proficiency / native language.* As it would have been impossible, because of time constraints, to objectively test the participants for English language proficiency, we asked participants for a self-evaluation of their language skills on a rating scale from 0 (No Proficiency) to 11 (Native or Bilingual Proficiency), in order to reflect the number and type of proficiency levels as measured by the ILR Scale [84]. We clearly instructed the participants to mark 11 *only* if English was actually their mother tongue, and in this way we were able to discriminate native speakers from
non-native speakers in our analysis. Information on the English proficiency of NNSs was kept as background information.

Legal background. Based on the educational background of the participants, we created a dummy variable to single out everyone with a legal background. As, in theory, participants with legal knowledge should be able to skim through and interpret legal texts with more ease, being “legalese native speakers”, we needed to analyze whether this variable had a simple or interaction effect on either answering speed or accuracy.

Object-spatial-verbal cognitive style. The verbal-visual cognitive style of the participants (the individual preference in relying on visual imagery versus verbal strategies in thought processes) was measured according to the Object-Spatial Imagery and Verbal Questionnaire (OSIVQ) developed by Kozhenikov and Blazhenkova [75]. The OSIVQ is an efficient self-report instrument composed of 45 items in the format of a 5-point Likert scale, and provides three separate scores: object-imagery skills (ability in building and processing vivid, colorful, detailed images of individual objects); spatial-imagery skills (ability in representing and processing schematic images, spatial relationship between objects, movement and spatial transformations); and verbal skills (ability in verbal, logical, sequential and analytical reasoning). The questionnaire is easy to administer and, as shown in [75], it displays a relatively good internal reliability (Cronbach’s .74 for the verbal scale, .83 for the object scale and .79 for the spatial scale).

The OSIVQ is superior to previous instruments assessing individual differences in the ability to process visual and verbal information (e.g. IDQ – Individual Differences Questionnaire [57], VVQ – Verbalizer–Visualizer Questionnaire [82]) as it corrects some misconceptions. Firstly, the visual-verbal dimension is not a bipolar construct, and different processing abilities need to be evaluated
individually and separately [85]. Secondly, neuroscientific evidence shows that visual information is processed in two different imagery sub-systems, the object imagery system and the spatial imagery system [86], resulting in two distinct visual cognitive styles rather than one [75].

*Mental effort.* Although a self-assessed measure of perceived mental effort may appear questionable, especially on a unidimensional scale, several studies on CLT had employed this simple and quick method (for a review, see [66]), as it has been demonstrated that these scales are as sensitive and reliable as more intrusive and complex methods [87]. Moreover, given the virtual setup of our experiment, it would have been impossible to use physiological measures (e.g. heart rate and pupillary responses). Following [65], we chose a 9-point scale, ranging from ‘very, very low mental effort’ (1) to ‘very, very high mental effort’ (9). Mental effort was used as a control variable in our regression models, and to calculate mental efficiency (see “How data was analyzed”).

**Process for Administering the Study**

The participants had the chance of getting instructions on the study and on its procedure by attending one of three webinars, organized at different times to accommodate the time zone of our international participants, and ensure a wide participation. One of the authors was available online for the duration of the experiment in order to answer any further questions from the participants in regards to the procedure. The webinars were also recorded for the participants who could not attend live (2% of the participants), and we gave them the possibility to ask further questions via email from one of the authors. The link to the survey was shared with participants during the webinar.

Upon accessing the webpage for the experiment and the survey, a script randomly assigned the participants into control and treatment groups. Firstly, the participants filled in the questionnaire collecting the control variables. Then, they proceeded to the experimental setup, and had up to 4
minutes to get acquainted with the assigned version of the document. After that, each of the 6 comprehension questions were presented one at a time. After having carefully read the question, the participants had to click a button to start the time countdown and gain access to the document in order to search for the information needed to provide the correct answer. After completing their answer, the participants stopped the timer, and could proceed to the next comprehension question. After stopping the timer, the answers could not be modified anymore.

The participants were given 7 minutes at most for each comprehension question. This was done for two reasons: firstly, to limit the length of the research session (experiment plus questionnaire) to about one hour; secondly, because our questions were quite practical and expressed operational doubts, and did not require legal interpretation (Appendix 2): if contracts are to be seen as instructions or “blueprint for performance” [24], it must be possible for the readers to find information and understand in a relatively short time what needs to be done in business-as-usual situations. If this is not possible, the contract is not an effective communication tool.

After the experiment, the participants were asked to self-assess the mental effort they experienced in completing the tasks.

How data was analyzed

In order to test H1 and H2, relating to speed and accuracy, we analyzed the comprehension performance scores obtained in the experiment. The difference in mean answering speed in the two experimental groups was analyzed with a Welch t-test, an adaptation of the classical t-test to be alternatively used when the two samples have unequal variances, as in this case. The difference in answering accuracy, instead, had to be analyzed through the non-parametric Mann-Whitney U-test, because the scores of the experimental group were non-normally distributed, thus violating one of
the assumptions required for a t-test. The Mann-Whitney U-test could be interpreted as a test of medians, because its two main assumptions were respected: the two samples had equal variance, as assessed through a Levene’s test, and their distributions were similar in shape (as assessed through visual inspection of the plotted data-points).

In order to test H3, we followed the methodology described in [65] on how to calculate the mental efficiency of alternative instructional conditions. The method puts individual performance in the context of related subjectively experienced cognitive load, and viceversa [65]. This procedure was deemed appropriate as it provides an overall score for each document alternative, helping to determine which one requires less mental effort and yields better performance.

The first step of the procedure is to standardize answering speed, answering accuracy, and mental effort scores into z-scores:

\[
z = \frac{(\text{individual score} - M)}{SD}
\]

Performance is then obtained as the mean of \(z_{\text{Answering Speed}}\) and \(z_{\text{Answering Accuracy}}\). Mean z-scores of performance and mental effort constitute the Cartesian coordinates to plot a result point for each condition (x axis = performance; y axis = mental effort). Efficiency (E) of each condition is finally calculated as the distance of the plotted points from the bisecting diagonal:

\[
E = \frac{Z_{\text{Performance}} - Z_{\text{Mental Effort}}}{\sqrt{2}}
\]

The statistical difference between Efficiency scores is then assessed through a t-test, as the data was normally distributed, had homogeneous variance, and samples were independent.
In order to test H4a, H4b, H5a, H5b and H5c, related to the interaction of the cognitive style and language with experimental treatment, we employed stepwise linear regression and 2-way ANOVA.

Firstly, we examined the interactions of objective performance measures (speed and accuracy) by different verbal skills (OSIVQ verbal skill score, native language and legal background) and by different visual skills (OSIVQ object-imagery skill and spatial-imagery skill scores). All variables involved in the interaction terms were centered to avoid multicollinearity between these terms and their individual components [88, pp. 37-38]. Regression analysis also helped us to prove that the experimental manipulation was the main cause for changes in comprehension performance, and thus discard other variables as alternative explanations for differences in performance between groups, ensuring internal validity. Linear regression was appropriate as it allowed us to include all control variables – categorical, continuous, and interaction terms – in the same analysis. Since our study sample was too small to generate sufficient predictive power if all variables and interaction terms would have been included at once, we chose a stepwise method based on probability of the F value. We included the variables in the regression model one at a time, starting with those with lowest p-values. Variables already in the equation would be eliminated if their p-value became not significant (higher than 0.05) because of the inclusion of a new variable.

We used a 2-way ANOVA to investigate further the only interaction term that resulted statistically significant in the regressions (native language × experimental treatment). ANOVA was a suitable follow-up method because the interaction term we discovered was between categorical independent variables on one continuous dependent variable, and we could more precisely assess the main effect of each independent variable and how they interacted. To conclude, we also analyzed the simple main effects through univariate post-hoc tests and Bonferroni-adjusted pairwise comparisons. The
post-hoc tests reveal whether there are simple main effects at each combination of variable levels (e.g. whether “treatment” has a significant effect on NSs and NNSs), and provides an effect size through the F-value. The pairwise comparisons establish whether the mean difference between two variable levels within a group is statistically significant (e.g. whether the mean score difference of NSs and NNSs within the “visual” treatment group is significant).

All statistical tests and analyses were performed with the statistical software SPSS.

**Assuring Reliability and Validity**

In this section we illustrate the efforts we took to assure reliability and validity.

Ecological validity:

- We used an authentic agreement, for which existed both text-only and visually-enhanced versions
- We asked experts from IACCM to check whether the type of contract we chose, its language and structure, and our comprehension tasks would be plausible and “business as usual” for the average contract professional
- We ran the experiment with actual contract professionals, rather than using master students as proxies.

Population validity:

- We sought to increase the cultural and professional diversity of the participants as much as possible by 1) recruiting participants through IACCM, an international association, rather than through individual national associations (some countries do not even have national associations of contract professionals) 2) organizing explanatory sessions at different times, so as to eliminate hurdles to participation due to unfavorable time zones; 3) avoiding to preselect participants based on specific criteria. However, we acknowledge that our
participants are volunteers, and may also be considered a convenience sample, since, in order to facilitate recruiting, we sought to recruit the members of an international association, rather than trying to recruiting participants directly from individual firms randomly chosen from all around the world.

Reliability of the instruments

- We sought to use scales and procedures documented and validated in previous studies.
- In grading answering accuracy, we followed an established rationale (Appendix 3) and checked inter-rater reliability

RESULTS

In this section we report detailed information about the participant sample, as well as the analyses performed to test our hypotheses.

Who participated in the study

The sample was composed of 122 participants. 59 of them were women, 31 had a legal background, and 71 had English as their mother tongue and were thus NSs. Among the NNSs the average self-assessed proficiency in English was 8.570 on a 11-points scale ($SD=1.375$). The average age of the participants was 44.460 years old ($SD = 10.12$), and in average they have been working with contracts for 12.611 years ($SD = 9.032$). Most of participants were from North America ($n = 56$) and Europe ($n = 45$). The most represented professions were contract manager ($n = 71$) and lawyer/in-house counsel ($n = 16$). The most represented industries were IT ($n = 36$), services/outsourcing/consulting ($n = 16$), oil/gas/utilities ($n = 15$) and technology ($n = 11$). In terms of cognitive styles, we had 53 predominant object visualizers, 44 predominant verbalizers, 28 predominant spatial visualizers and 3 individuals with mixed cognitive styles (their highest score
was the same for 2 or 3 cognitive styles). The average scores of the whole sample were 3.370 for object imagery, 2.950 for spatial imagery ability, and 3.230 for verbal ability. 67 people were randomly allocated to the control group (answering to the comprehension questions with the text-only, original contract), and 55 to the treatment group (answering to the comprehension questions with the redesigned visually enhanced contract).

**Hypotheses testing**

**H1.** *Contracts including diagrams allow for faster information finding than traditional, text-only contracts.*

Experimental data on the answering speed of the participants confirmed H1: the treatment group, using the visually enhanced version of the contract, was significantly faster than the control group in finding the relevant information and solving the comprehension tasks. Comparing the sums of answering times of the two groups, in average the treatment group was 182.760 seconds faster (treatment group mean = 897.070 ± 305.750; control group mean = 1079.830 ± 396.440). This difference is statistically significant, as assessed through a two-tailed Welch’s t-test ($t = -2.940, p = 0.004$). Figure 2 shows the mean of the time taken to answer each question and the mean of the sum of all individual scores.

**H2.** *Contracts including diagrams allow for more accurate comprehension than traditional, text-only contracts.*
Experimental data on the answering accuracy confirmed H2: the treatment group was significantly more accurate than the control group, with respectively a mean score of 4.350 correct answers versus only 3.010 correct answers. This difference is statistically significant, as assessed through a two-tailed Mann-Whitney U-test ($U = 2835, p < 0.001$). Figure 3 shows the mean accuracy scores for each question and as a sum of all individual scores.

**FIGURE 2**

Mean answering speed for each comprehension question (left), and total mean answering speed (right). Results of Welch’s t-test (table at the bottom).
FIGURE 3

Mean answering accuracy for each comprehension question, where 0 indicates a wrong answer and 1 a correct answer (left). Total mean answering accuracy, where 0 indicates that participants’ answers were all wrong, and 6 that they were all right (right). Results of Mann-Whitney U-test (table at the bottom).

H3. The mental efficiency of the version of the contract including diagrams is higher than that of the text-only version, meaning that participants require less mental effort to obtain an equal level of performance.
H3 is strongly confirmed as the participants in the treatment group not only experienced less mental effort, but also achieved higher performance. Given the formula presented earlier, the visual contract’s Efficiency (0.282) is higher than the textual contract’s (-0.232), and the difference is statistically significant ($t = 3.399, p = 0.001$). Moreover, the visual contract’s score indicates high-efficiency, while the textual contract’s indicates low efficiency (Figure 4). We can also exclude that performance in the visual group was caused by the motivation of putting extra-effort to succeed.

Since the intrinsic cognitive load was arguably the same in each group (because all participants had the same contract content and the same tasks), we can infer that the overall decrease in mental effort was caused by a reduction in extraneous cognitive load, which is the load imposed by inefficient ways of communicating the content and providing a problem representations.
**FIGURE 4**

Graph of Efficiency (E) for Visual and Textual conditions.

**H4a. Cognitive styles interact with experimental treatment, and with each other:**

- verbalizers are equally fast with either contract, and are fastest with the textual contract;
- spatial-visualizers are the fastest with contracts including diagrams;
- object-visualizers become faster with contracts including diagrams than textual ones, but remain the slowest performing group.

And

**H4b. Cognitive styles interact with experimental treatment, and with each other:**

- verbalizers are equally accurate with either contract, and are most accurate with the textual contract;
spatial-visualizers are the most accurate with contracts including diagrams; object-visualizers become more accurate with contracts including diagrams than textual ones, but remain the least accurate group.

The results from the regression analyses show that cognitive styles do not interact with experimental treatment, alone or in concert, and do not have a statistically significant effect in predicting answering accuracy or speed, neither alone or in interaction with other variables (Tables 3 and 4).

Looking at the regression model explaining answering speed (Table 3), the only statistically significant coefficient is experimental treatment ($\beta = -0.271, p = 0.030$), and the model statistically significantly predicts answering speed ($R^2 = 0.073, F (1, 119) = 9.404, p = 0.030$).

If we look at the regression model explaining answering accuracy (Table 4), treatment is also statistically very significant ($\beta = 0.468, p < 0.001$). The model is statistically significant in predicting answering accuracy ($R^2 = 0.270, F (3, 117) = 14.438, p < 0.001$), and features two more significant coefficients – mother tongue, alone and in interaction with experimental treatment – which we will illustrate in the next section.

To conclude, we can confidently associate the cause of lower response times and increased comprehension accuracy to the design of the visualized contract, for visualizers and verbalizers alike. The results provide further support to accept H1 and H2, while H4a and H4b are rejected.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>s.d.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Answering speed</td>
<td>997.438</td>
<td>360.193</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Answering accuracy</td>
<td>3.606</td>
<td>1.431</td>
<td>0.059</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Experimental treatment</td>
<td>0.450</td>
<td>0.500</td>
<td>-0.254**</td>
<td>0.464**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Age</td>
<td>44.460</td>
<td>10.118</td>
<td>0.021</td>
<td>-0.041</td>
<td>-0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Gender</td>
<td>0.480</td>
<td>0.502</td>
<td>-0.015</td>
<td>-0.156</td>
<td>-0.086</td>
<td>-0.162</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Legal background</td>
<td>0.270</td>
<td>0.464</td>
<td>-0.028</td>
<td>-0.038</td>
<td>0.183*</td>
<td>-0.128</td>
<td>0.072</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Years of experience in</td>
<td>12.611</td>
<td>9.03</td>
<td>0.011</td>
<td>0.017</td>
<td>-0.048</td>
<td>0.674**</td>
<td>-0.299**</td>
<td>0.047</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>working with contracts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Mother tongue</td>
<td>0.580</td>
<td>0.495</td>
<td>0.161</td>
<td>0.151</td>
<td>-0.034</td>
<td>0.230**</td>
<td>-0.211*</td>
<td>-0.079</td>
<td>0.293**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Object-imagery skill score</td>
<td>3.375</td>
<td>0.595</td>
<td>-0.039</td>
<td>-0.148</td>
<td>-0.057</td>
<td>-0.093</td>
<td>0.273**</td>
<td>0.049</td>
<td>-0.145</td>
<td>-0.226**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Spatial-imagery skill</td>
<td>2.951</td>
<td>0.646</td>
<td>-0.001</td>
<td>0.069</td>
<td>-0.080</td>
<td>-0.065</td>
<td>-0.281**</td>
<td>-0.052</td>
<td>0.025</td>
<td>-0.178*</td>
<td>-0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Verbal skill score</td>
<td>3.230</td>
<td>0.499</td>
<td>-0.013</td>
<td>-0.022</td>
<td>-0.028</td>
<td>0.175</td>
<td>0.051</td>
<td>0.251**</td>
<td>0.151</td>
<td>0.161</td>
<td>-0.183</td>
<td>-0.190*</td>
<td></td>
</tr>
<tr>
<td>12. Mental effort</td>
<td>4.668</td>
<td>1.556</td>
<td>0.237**</td>
<td>-0.208*</td>
<td>-0.258*</td>
<td>-0.048</td>
<td>0.128</td>
<td>-0.046</td>
<td>-0.098</td>
<td>0.065</td>
<td>-0.153</td>
<td>0.076</td>
<td>-0.132</td>
</tr>
</tbody>
</table>

n = 121  *p < 0.05  **p < 0.01
TABLE 3
Results of the regression analysis (answering speed as the dependent variable)

<table>
<thead>
<tr>
<th>Variables included in the model</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental treatment</td>
<td>-0.271*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables excluded from the model, $p &gt; 0.05$</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.021</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.021</td>
</tr>
<tr>
<td>Legal background</td>
<td>0.012</td>
</tr>
<tr>
<td>Years of experience in working with contracts</td>
<td>-0.008</td>
</tr>
<tr>
<td>Mother tongue</td>
<td>0.136</td>
</tr>
<tr>
<td>Spatial-imagery skill score</td>
<td>-0.025</td>
</tr>
<tr>
<td>Object-imagery skill score</td>
<td>-0.089</td>
</tr>
<tr>
<td>Verbal skill score</td>
<td>-0.011</td>
</tr>
<tr>
<td>Mental effort</td>
<td>0.175</td>
</tr>
<tr>
<td>Spatial-imagery skill score × Experimental treatment</td>
<td>0.004</td>
</tr>
<tr>
<td>Object-imagery skill score × Experimental treatment</td>
<td>-0.002</td>
</tr>
<tr>
<td>Verbal skill score × Experimental treatment</td>
<td>0.114</td>
</tr>
<tr>
<td>Spatial-imagery skill score × Verbal skill score</td>
<td>0.018</td>
</tr>
<tr>
<td>Object-imagery skill score × Verbal skill score</td>
<td>-0.087</td>
</tr>
<tr>
<td>Object-imagery skill score × Spatial-imagery skill score</td>
<td>-0.065</td>
</tr>
<tr>
<td>Legal background × Mother tongue</td>
<td>0.086</td>
</tr>
<tr>
<td>Legal background × Experimental treatment</td>
<td>0.086</td>
</tr>
<tr>
<td>Mother tongue × Experimental treatment</td>
<td>0.084</td>
</tr>
<tr>
<td>Object-imagery skill score × Spatial-imagery skill score × Experimental treatment</td>
<td>-0.129</td>
</tr>
<tr>
<td>Object-imagery skill score × Verbal skill score × Experimental treatment</td>
<td>-0.066</td>
</tr>
<tr>
<td>Spatial-imagery skill score × Verbal skill score × Experimental treatment</td>
<td>-0.089</td>
</tr>
<tr>
<td>Object-imagery skill score × Spatial-imagery skill score × Verbal skill score</td>
<td>-0.137</td>
</tr>
<tr>
<td>Legal background × Mother tongue × Experimental treatment</td>
<td>-0.029</td>
</tr>
<tr>
<td>Object-imagery skill score × Spatial-imagery skill score × Verbal skill score × Experimental treatment</td>
<td>0.051</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Adjusted $R^2$</th>
<th>0.065</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>0.073</td>
</tr>
<tr>
<td>$F (1, 119)$</td>
<td>8.404</td>
</tr>
<tr>
<td>Significance of the model, $p$</td>
<td>0.003</td>
</tr>
</tbody>
</table>

$n = 121$ * $p = 0.003$


**TABLE 4**

Results of the regression analysis (answering accuracy as the dependent variable)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables included in the model</td>
<td></td>
</tr>
<tr>
<td>Experimental treatment</td>
<td>0.468**</td>
</tr>
<tr>
<td>Mother tongue</td>
<td>0.167***</td>
</tr>
<tr>
<td>Mother tongue × Experimental treatment</td>
<td>-0.179***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables excluded from the model, p &gt; 0.05</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.083</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.055</td>
</tr>
<tr>
<td>Legal background</td>
<td>-0.095</td>
</tr>
<tr>
<td>Years of experience in working with contracts</td>
<td>-0.012</td>
</tr>
<tr>
<td>Spatial-imagery skill score</td>
<td>0.139</td>
</tr>
<tr>
<td>Object-imagery skill score</td>
<td>-0.085</td>
</tr>
<tr>
<td>Verbal skill score</td>
<td>-0.023</td>
</tr>
<tr>
<td>Mental effort</td>
<td>-0.075</td>
</tr>
<tr>
<td>Spatial-imagery skill score × Experimental treatment</td>
<td>0.029</td>
</tr>
<tr>
<td>Object-imagery skill score × Experimental treatment</td>
<td>0.070</td>
</tr>
<tr>
<td>Verbal skill score × Experimental treatment</td>
<td>0.025</td>
</tr>
<tr>
<td>Spatial-imagery skill score × Verbal skill score</td>
<td>-0.005</td>
</tr>
<tr>
<td>Object-imagery skill score × Verbal skill score</td>
<td>-0.046</td>
</tr>
<tr>
<td>Object-imagery skill score × Spatial-imagery skill score</td>
<td>-0.015</td>
</tr>
<tr>
<td>Legal background × Mother tongue</td>
<td>0.016</td>
</tr>
<tr>
<td>Legal background × Experimental treatment</td>
<td>-0.092</td>
</tr>
<tr>
<td>Object-imagery skill score × Spatial-imagery skill score × Experimental treatment</td>
<td>-0.076</td>
</tr>
<tr>
<td>Object-imagery skill score × Verbal skill score × Experimental treatment</td>
<td>0.130</td>
</tr>
<tr>
<td>Spatial-imagery skill score × Verbal skill score × Experimental treatment</td>
<td>-0.155</td>
</tr>
<tr>
<td>Object-imagery skill score × Spatial-imagery skill score × Verbal skill score</td>
<td>0.074</td>
</tr>
<tr>
<td>Legal background × Mother tongue × Experimental treatment</td>
<td>0.052</td>
</tr>
<tr>
<td>Object-imagery skill score × Spatial-imagery skill score × Verbal skill score × Experimental treatment</td>
<td>-0.038</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.251</td>
</tr>
<tr>
<td>R²</td>
<td>0.270</td>
</tr>
<tr>
<td>F (3, 117)</td>
<td>14.438</td>
</tr>
<tr>
<td>Significance of the model, p</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

\( n = 121 \quad ** p < 0.01 \quad *** p < 0.05 \)

**H5a.** Language interacts with experimental treatment: native speakers of English are faster and more accurate than nonnative speakers with the traditional version of the contract (text-only), while nonnative speakers perform faster and more accurately with the version of the contract including diagrams.
As already illustrated above, the experimental treatment is the only factor explaining the answering speed scores, so H4 is partially rejected in regards to this dependent variable: experimental subjects do not differ in their answering speed because of their native language, but because of the contract version they use.

However, this is different in terms of answering accuracy: the results of the regression analysis show not only a statistically significant coefficient for experimental treatment, but also for mother tongue ($\beta = 0.167, p = 0.037$) and for the interaction effect of mother tongue $\times$ treatment ($\beta = -0.179, p = 0.026$). This means that answering accuracy is affected by both the version of the contract used in the comprehension task, and by whether the respondent is a native English speaker. The statistically significant interaction term suggests that the performance of both native and nonnative speakers is affected by the type of contract, but in a different measure. We further investigated this difference, and assessed whether an expertise reversal effect is present, through a 2-way ANOVA (Table 5), with post-hoc tests (Tables 6 and 8) and pairwise comparisons (Tables 7 and 9).

**TABLE 5**

2-way ANOVA: Test of Between-Groups Effects on Answering Accuracy

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Squares</th>
<th>F</th>
<th>p-value</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>59.945</td>
<td>1</td>
<td>59.945</td>
<td>39.410</td>
<td>&lt; 0.001</td>
<td>0.250</td>
</tr>
<tr>
<td>Mother Tongue</td>
<td>5.587</td>
<td>1</td>
<td>5.587</td>
<td>3.673</td>
<td>0.058</td>
<td>0.030</td>
</tr>
<tr>
<td>Treatment x Mother Tongue</td>
<td>8.041</td>
<td>1</td>
<td>8.041</td>
<td>5.286</td>
<td>0.023</td>
<td>0.043</td>
</tr>
<tr>
<td>Error</td>
<td>179.485</td>
<td>118</td>
<td>1.521</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>247.740</td>
<td>121</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In the control group, native speakers are in average significantly more accurate than nonnative speakers (mean difference = 0.959, 95% CI [0.350, 1.567], \( F (1, 118) = 9.737, p = 0.002 \)), but this superiority surprisingly disappears in the treatment group: native and nonnative speakers are virtually equally accurate, and there is no statistically significant difference in their accuracy (\( F (1, 118) = 0.67, p = 0.796 \)).

In fact, the average of correct answers of nonnative speakers (4.385) is even slightly higher than the average of native speakers (4.289) (Figure 5). However, this difference is statistically non-significant (Table 9): NNSs “catch up” on NS, performing equally accurately, but a full expertise reversal effect is not observed. This is because both groups significantly improve their accuracy with visual contracts, even though the effect is stronger for nonnative speakers (mean difference = 1.950, 95% CI [1.265, 2.635], \( F (1, 118) = 31.771, p < 0.001 \)) than for native speakers (mean difference = 0.905, 95% CI [0.320, 1.489], \( F (1, 118) = 9.397, p = 0.003 \)).

We can conclude that H4 is partially confirmed when considering answering accuracy. As postulated by the expertise reversal effect prediction, instructional visual aids are more effective for those with lower expertise (in this case, native language expertise). However, a full expertise reversal effect is not observed, because even native speakers benefit from visualization – yet just to a lesser extent than NNSs.
TABLE 6

Post-hoc univariate tests on simple main effect of ‘Treatment’ within ‘Mother Tongue’ groups

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Squares</th>
<th>F</th>
<th>p-value</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native speakers</td>
<td>14.293</td>
<td>1</td>
<td>14.293</td>
<td>9.397</td>
<td>0.003</td>
<td>0.074</td>
</tr>
<tr>
<td>Error</td>
<td>179.485</td>
<td>118</td>
<td>1.521</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonnative speakers</td>
<td>48.326</td>
<td>1</td>
<td>48.326</td>
<td>31.771</td>
<td>&lt; 0.001</td>
<td>0.212</td>
</tr>
<tr>
<td>Error</td>
<td>179.485</td>
<td>118</td>
<td>1.521</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 7

Pairwise comparisons: mean difference of “Treatment” within ‘Mother Tongue’ groups

<table>
<thead>
<tr>
<th>Mother-tongue Groups</th>
<th>Mean Score Treatment Group (visual)</th>
<th>Mean Score Control Group (textual)</th>
<th>Mean Difference</th>
<th>p-value</th>
<th>95% Confidence Interval for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
</tr>
<tr>
<td>Native Speakers</td>
<td>4.298</td>
<td>3.394</td>
<td>0.905</td>
<td>0.003</td>
<td>0.320</td>
</tr>
<tr>
<td>Nonnative Speakers</td>
<td>4.385</td>
<td>2.435</td>
<td>1.950</td>
<td>&lt; 0.001</td>
<td>1.265</td>
</tr>
</tbody>
</table>

TABLE 8

Post-hoc univariate tests on simple main effect of ‘Mother Tongue’ within ‘Treatment’ groups

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Squares</th>
<th>F</th>
<th>p-value</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment group (visual)</td>
<td>0.102</td>
<td>1</td>
<td>0.102</td>
<td>0.067</td>
<td>0.796</td>
<td>0.001</td>
</tr>
<tr>
<td>Error</td>
<td>179.485</td>
<td>118</td>
<td>1.521</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control group (textual)</td>
<td>14.811</td>
<td>1</td>
<td>14.811</td>
<td>9.737</td>
<td>0.002</td>
<td>0.076</td>
</tr>
<tr>
<td>Error</td>
<td>179.485</td>
<td>118</td>
<td>1.521</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 9

Pairwise comparisons: mean difference of ‘Mother Tongue’ within ‘Treatment’ groups

<table>
<thead>
<tr>
<th>Treatment Groups</th>
<th>Mean Score Native Speakers</th>
<th>Mean Score Nonnative speakers</th>
<th>Mean Difference</th>
<th>p-value</th>
<th>95% Confidence Interval for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment group (visual)</td>
<td>4.298</td>
<td>4.385</td>
<td>-0.087</td>
<td>0.769</td>
<td>[-0.757  0.577]</td>
</tr>
<tr>
<td>Control group (textual)</td>
<td>3.394</td>
<td>2.435</td>
<td>0.959</td>
<td>0.002</td>
<td>[0.350  1.567]</td>
</tr>
</tbody>
</table>

FIGURE 5

Mean Answering Accuracy for Native and Nonnative Speakers in the Control Group and in the Treatment Group.

H5b. Profession-specific knowledge interacts with experimental treatment: subjects with a legal background are faster and more accurate with the traditional version of the contract (text-only), while subjects without legal background perform faster and more accurately with the version of the contracts including diagrams.
And

**H5c.** *Language interacts with profession-specific knowledge and experimental treatment: native speakers of English with a legal background are the fastest and most accurate with the traditional version of the contract (text-only), while nonnative speakers without a legal background are fastest and most accurate with the version of the contract including diagrams.*

As illustrated in Tables 3 and 4, all regression coefficients involving the variable “Legal background” – alone or in interaction with other terms – are not significant in accounting for answering speed and accuracy results. Thus, H5b and H5c are rejected.

**CONCLUSIONS, LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH**

In this last section, we elaborate on our results, and their implications for practice and theory. We close the section by considering the limitations of the study, and by suggesting venues for future research.

**Conclusions**

In this study, we argued for integrating diagrams in B2B contracts, as a way to enhance comprehension speed, accuracy, and ultimately communication in an international business setting. Given the novelty of the approach, we set forth four research questions to explore the merit of this suggestion: to what extent this approach may be beneficial, whether individual characteristics may alter its effectiveness, and what would be the implications for international businesses. We discovered that adding diagrams to contracts indeed supports more accurate and faster
comprehension of their content, and the increase in accuracy is especially significant for non-native
speakers of English, and positive also for native speakers – thus reinforcing our belief that this
approach may be highly beneficial in international business, where English is the de facto lingua
franca. Moreover, we have not observed negative reversal effects caused by any individual
characteristic of respondents (e.g. cognitive styles, expertise, and professional background), leading
us to exclude that this approach may potentially backfire with specific user groups.

**Implications to practice.** The implications for global business are important, because through a
multimodal approach to contract design firms could ensure that professionals from different
backgrounds and cultures can more easily understand each other and collaborate. This mutual
understanding and collaboration, in return, could lead to more effective negotiations, prevent legal
troubles resulting from contract misinterpretation, and improved business relationships. Technical
and professional communicators, as well as information designers, could apply their skills in the
field of contracts, and help global firms conduct transactions more effectively and successfully.

Visualization skills can be seen as part of the strategic competence of the global communicator,
whose goal is to communicate effectively to “get the job done” [6]. In Louhiala-Salminen and
Kankaanranta’s Global Communication Competence framework [7], in fact, communication
proficiency goes beyond grammatical and discourse proficiency. One of the key abilities of
competent global communicators is the ability to continuously check for understanding (their own
and of their communication partners) through questions, repeated utterances and the use of more
than one channel or modality. We suggest that a competent global communicator should craft
messages in whichever way offers the best possibilities of constructing shared understanding:
proficiently using different media and modalities is a critical skill for professionals involved in
international business.
Implications to theory.

The results from the experiment show how a supplantation strategy [76] based on diagrams helps international managers in understanding contracts more accurately and more quickly. In particular, we discovered how diagrams can support non-native speakers of English in developing a more accurate understanding of a complex document written in English, the shared business language of today’s professional communicators. This effect will ultimately close the accuracy gap in regards to native speakers. These results are partially explained in terms of expertise reversal effect (e.g. [68], [70]–[73]): while NNSs benefit more than NSs from visual aids in contracts, the latter also see an improvement in their answering accuracy.

These results can be explained in light of cognitive load theory [53][54]: when information is already intrinsically complex and requires substantial processing, inefficient or confusing presentation formats needs to be minimized as much as possible; better instructional designs prevent extraneous cognitive load by offering a meaningful problem representation. When this problem representation offers a ready solution to the task at hand, the design is more mentally efficient: less cognitive resources need to be invested, and understanding will be better and faster.

However, we believe that the presence of visualizations is not enough to ensure comprehension [33][68], as for some tasks we found that there were no performance improvements. Diagrams must offer problem representations that are more explicit and readily applicable in relation to a given task (we will return to this point in the “Limitations” -subsection). There is no such thing as the right diagram, but only diagrams well-suited to answer specific questions and achieving specific goals.
These results also suggest avoiding pre-deterministic assumptions on what modes of presentation may work well for particular cognitive styles. Counterintuitively, individual cognitive styles did not have an effect on how the participants processed either version of the contract. Object and spatial visualizers were not at a particular disadvantage in understanding the traditional contract (it was equally difficult for verbalizers), and verbalizers did not benefit less from diagrams in the experimental version of the contract. The “fitness-for-purpose” of a design apparently trumps cognitive style in predicting the comprehension performance of an individual processing particularly complex information: dense prose will remain dense even for the best verbalizer, and an irrelevant diagram will not be helpful even for the most skilled spatial-visualizer.

**Limitations**

The current study is not without limitations. Firstly, while experiments provide a way to control for variables, they are carried out in an artificial setting. In real life, managers might overcome their difficulties in understanding a document by discussing and sharing knowledge with colleagues, so in addition to individual cognitive mechanisms there are social mechanisms at play in sensemaking, which our study does not consider.

Secondly, the participants to the experiments constituted a convenience sample: at best, the results can be generalized to the population of the members of IACCM Association. The members of this association are very active in learning contractual and commercial best practices, and all participants were volunteers. Arguably, the participants were interested from the start in trying out their contract comprehension skills, or having a chance to try out a visual contract. This may affect the results and participants’ characteristics in terms of non-response bias, as we do not know how the respondents may compare with non-respondents. There is good evidence that volunteers in psychological research
are, for instance, more educated and have higher IQs than non-volunteers [89]. If this factor was at play, it may have affected the representativeness of our sample and the mean overall performance results. Moreover, the fact that they were invited to participate to the study through the IACCM Association may have biased them in unconsciously considering the visual contract a superior solution from the very start. On the other hand, running the experiment without the help of the Association would have made it almost impossible to recruit such a varied and international sample, which was necessary to assess the possible benefits of our proposal for global communicators.

Thirdly, in order to further increase validity, the sample might have been bigger, or the experiment might have been replicated within different communities. Within the scope of this study, however, we felt confident to present only one experiment, because converging results have been obtained in other studies where visual communication has been used to improve speed and accuracy in contract interpretation (e.g. [91][92]).

Fourthly, some variables may be more rigorously measured. We acknowledge that actual English language proficiency could have been more thoroughly assessed. Since a self-assessment without an objective test may not be reliable, we did not include self-assessed English proficiency as a control variable in our analysis. However, differences in English proficiency among NSSs – and in literacy among NSs – may yield more nuanced results. Moreover, we considered only the verbal-imagery dimension of cognitive style and found no evidence for their effect on contract comprehension. Future studies might need to additionally consider the wholist-analytical dimension of cognitive style [78]–[80]. Better controls for the culture of the participants could also be included, as visual representations are interpreted differently in “high-context” and “low-context” cultures [90].
Lastly, it is to be noted that the results of our study apply only to diagrammatic visualizations. The contracts used in the experiment did not include other types of visuals – e.g. photographs, illustrations, cartoons – which possess a more pictorial or metaphorical nature. Moreover, more attention should be paid to the type of diagram (or visual representation used) in relation to the task that participants need to achieve. More research is needed to discover what are the actual doubts and practical goals of contract professionals, as identifying those tasks is the first step to envision truly helpful diagrams. For example, in tasks 3 and 5 there was no significant difference in accuracy or speed across treatments because, we hypothesize, the diagrams did not provide a problem representation that was sufficiently helpful for the respondents and readily applicable. Question 3 (Appendix 2) asks to calculate the correct duration of the warranty period. The diagram (Appendix 1) does suggest how to perform the calculation correctly, but does not provide a ready-made answer and does not reframe the problem. The readers still need to apply a “tricky” rule. A better approach would have been to create a simple table proposing an easier-to-apply rule, as we suggest in Figure 6.

FIGURE 6

Suggestion on how to improve the “Warranty Period” diagram (cfr. Appendix 1).

<table>
<thead>
<tr>
<th>Condition</th>
<th>Warranty Period Expire</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>If Acceptance takes place</strong></td>
<td></td>
</tr>
<tr>
<td><strong>MORE THAN A YEAR</strong></td>
<td><strong>24 MONTHS</strong></td>
</tr>
<tr>
<td><strong>after Delivery</strong></td>
<td><strong>after Acceptance</strong></td>
</tr>
<tr>
<td><strong>EXACTLY ONE YEAR</strong></td>
<td><strong>36 MONTHS</strong></td>
</tr>
<tr>
<td><strong>after Delivery</strong></td>
<td><strong>after Acceptance</strong></td>
</tr>
<tr>
<td><strong>LESS THAN A YEAR</strong></td>
<td></td>
</tr>
<tr>
<td><strong>after Delivery</strong></td>
<td></td>
</tr>
</tbody>
</table>

Question 5 (Appendix 2) has very high correct response rates in either group. This is because, we discovered, it is customary that the Supplier carries the risk for the equipment during its testing. Most managers were probably familiar with this convention, and already possessed the relevant
knowledge to answer very quickly to the question. The diagram (Appendix 1), as well as the text, had thus a rather irrelevant effect, as it was not the main driver behind people’s good performance.

**Suggestions for future research**

This study establishes the usefulness of diagrams in simplifying contract interpretation and support effective global communication. We suggest that further research should go beyond experimental evaluation, by possibly observing *in vivo* how visual representations support comprehension and knowledge sharing. Such study could shed light on how global communicators might interpret the very presence of visualizations in contract: would the benefits in comprehension be accompanied by an increase in trust and collaborativeness? Or would people feel patronized? Would they fear images could be used insincerely, thus leading to lower mutual trust between the parties?

Other studies, looking at both experimental and observational data, could help us shed light on what diagrams more effectively foster competent contractual communication, and how and when they should be used in an international business setting. Is there such a thing as an archetypical, best way to communicate certain clauses, or it depends on the goal readers want to achieve with that information? Are certain types of representations consistently clearer and easier to process than others, or does their effectiveness depend on individual and cultural variables? Future research could take into consideration other types of visuals in addition to diagrams, for instance comics and photographs, which have been reported to be used in contracts [93][94], but have not been systematically evaluated.

Lastly, one possible avenue to take would be to investigate the relationship between cognitive load, information design and cognitive styles more thoroughly, as the current study counterintuitively showed a weak link between the verbal-imagery dimension of cognitive style and comprehension of
differently presented materials. This might reinforce the argument that individual cognitive styles ultimately do not significantly impact learning [55]; on the other hand, it may suggest a mechanism between cognitive load and cognitive styles, undiscovered in the current study, which would explain these results.

References


[52] M. Betrancourt and A. Bisseret, “Integrating textual and pictorial information via pop-up


APPENDIX 1: Excerpts from the visualized contract

Sample spread of the agreement.
This excerpt from “Test Runs” (clause + diagram) provides an answer to question 1.

In the event that the Equipment does not fulfill some of the guarantee values specified in Appendix 1 during the Test Run(s), the Supplier shall as soon as possible and not later than within one (1) month, at its own expense and at a time convenient to the Purchaser, remedy the deficiencies noted, after which a new Test Run shall be carried out. If the guarantee values are still not attained in this renewed Test Run, the Supplier shall, at its own expense, without delay and within a maximum period of two (2) months, effect the necessary improvements and modifications to the Equipment. If the said guarantee values are not attained in the subsequent Test Run, the Purchaser is primarily entitled to demand liquidated damages to be paid as prescribed in Appendix 8.

Test runs process

© 2013 Stefania Passera. Used with permission.

This excerpt from “Warranty” (clause + diagram) provides an answer to question 3.

19.1.2 The warranty period shall be twenty-four (24) months from the date of Provisional Acceptance as specified in clause 15 “Final Acceptance”, however at least thirty-six (36) months from the delivery of the Equipment in accordance with the applicable delivery term (hereinafter the “Warranty Period”).

© 2013 Stefania Passera. Used with permission.
18.4 Should the delivery of the Equipment or part thereof or any other performances of the Supplier hereunder be delayed from any deadline specified in Appendix 2 as a deadline subject to liquidated damages for any other reason than Force Majeure event or a reason solely attributable to the Purchaser, the Supplier shall be liable to pay liquidated damages for delay at two (2) percent of the Purchase Price (without value added tax) for each commencing week of delay, however, not exceeding 10% of the Purchase Price.

18.5 The above liquidated damages for delay shall become due upon demand. The Purchaser shall be entitled to deduct the amount of liquidated damages for delay from any unpaid invoice of the Supplier and any other monies owed by the Purchaser to the Supplier.

18.6 Liquidated damages shall be without prejudice to any other rights or remedies of the Purchaser under the Agreement or applicable legislation, including but not limited to the Purchaser’s right to cancel the Agreement with immediate effect.

18.7 The liquidated damages do not release the Supplier from its obligation to deliver the Equipment and/or documents.

18.8 The total amount of liquidated damages due to delay in delivery shall not exceed 20% of Purchase Price.
This excerpt from “Transfer of risk and ownership” (clause + diagram) provides an answer to question 5.

### 17. Transfer of risk and ownership

The title to the Equipment passes from the Supplier to the Purchaser in proportion as the payments are made and in any event on date of Provisional Acceptance at the latest. The risk of loss or damage to the Equipment passes from the Supplier to the Purchaser on the date of Takeover. However, the Supplier shall have the risk for the deterioration of and damage to the Equipment during the performance of the Test Runs.

#### Timeline 2: summary on the transfer of ownership, risks, and responsibilities

<table>
<thead>
<tr>
<th>Ownership to the Equipment (17)</th>
<th>Risk of loss and damage to the Equipment (17)</th>
<th>Risk of deterioration and damage of the Equipment (17)</th>
<th>Responsibility of repairing Equipment defects at own expense (20.1.1)</th>
<th>Responsibility of providing Equipment performance and availability at own expense (20.2.2)</th>
<th>Responsibility of providing available maintenance for the Equipment at own expense (20.3.1)</th>
<th>Responsibility of repairing Equipment latent defects at own expense (20.16)</th>
<th>Lapse of Warranty Period</th>
<th>5 years from Provisional Acceptance</th>
<th>10 years from Provisional Acceptance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belongs to the Supplier</td>
<td>Belongs to the the Purchaser</td>
<td>Belongs to the the Purchaser</td>
<td>Belongs to the the Purchaser</td>
<td>Belongs to the the Purchaser</td>
<td>Belongs to the the Purchaser</td>
<td>Belongs to the the Purchaser</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payment 1</td>
<td>Payment 2</td>
<td>Takeover</td>
<td>Payment 3</td>
<td>Test runs</td>
<td>Provisional Acceptance</td>
<td>Payment 4</td>
<td>Lapse of Warranty Period</td>
<td>5 years from Provisional Acceptance</td>
<td>10 years from Provisional Acceptance</td>
</tr>
</tbody>
</table>

© 2013 Stefania Passera. Used with permission
APPENDIX 2: Comprehension questions (© 2013 Stefania Passera).

Question 1 – What are the consequences if the Test Run of the Equipment fails for the second time?

Question 2 – Given that the date of provisional acceptance is 6th March 2012, which date marks the end of the availability measurement period?

Question 3 – Given that the date of the delivery of the Equipment is 15th July 2011, and the date of provisional acceptance is 6th March 2012, on what date does the Warranty Period expires?

Question 4 – The first batch of the Equipment delivery was delayed 6 weeks. The second batch of the Equipment delivery was delayed 3 weeks. The third batch of the Equipment delivery was again delayed 3 weeks. What is the total amount of liquidated damages that the Supplier is liable to pay to the Purchaser?

Question 5 – Which of the Parties shall bear the risk for deterioration and damage to the Equipment during the Test Runs?

Question 6 – A spare part, included in the original scope of supply, is taken into service and brought back to the Purchaser on 20th May 2015. Given that the date of provisional acceptance is 6th March 2012, on what date does the Warranty Period for such reconditioned part expires?
APPENDIX 3: Rationale for grading the accuracy of participants’ responses (© 2013 Stefania Passera).

**0.25 points** = The participant only mention the clause number where the correct answer can be found, or copy-pastes the relevant excerpt. However, s/he does not provide a correct, precise and clear response in own words (e.g. if answering correctly would require to provide a precise date, the date is not given)

**0.50 points** = The participant wrote down the relevant clause or rule in own words. However, s/he does not provide a correct, precise and clear response in own words (e.g. if answering correctly would require to provide a precise date, the date is not given)

**0.75 points** = The participant provided a correct, precise and clear response in own words, eventually referencing the clause or rule s/he applied. However, there is a minor imprecision (e.g. if answering correctly would require to provide a precise date such as 4\(^{th}\) of March 2016, the participant wrote 5\(^{th}\) of March 2016)

**1 point** = The answer is fully correct. The participant provided a correct, precise and clear response in own words, eventually referencing the clause or rule s/he applied (e.g. if answering correctly would require to provide a precise date, the date is not given)