Extending the Sentence Verification Technique to Tables and Node-Link Diagrams

Mark A. Livingston(✉), Derek Brock, Tucker Maney, and Dennis Perzanowski

Naval Research Laboratory, 4555 Overlook Ave SW, Washington, DC 20735, USA
{mark.livingston, derek.brock, tucker.maney, dennis.perzanowski}@nrl.navy.mil

Abstract. In addition to prose, situation reports used by various organizations often present supporting information in "visual" formats that pose unique challenges for assessing readers' comprehension. Two of the more common information categories seen in these reports are text-based tables and node-link diagrams. To better understand what readers attend to in these formats, we adapted a proven method for assessing prose comprehension, known as the Sentence Verification Technique (SVT), to these ancillary materials and conducted an exploratory reading study with format, aspects of the SVT, and contextual information as independent variables. Except for tables, error rates were comparatively uniform. Assessments of prose were significantly faster than assessments of diagrammatic information, which in turn were significantly faster than table assessments. The latter also took longer when posed without contextual details. We conclude that the SVT can be successfully adapted for information in ancillary formats and discuss further research issues for this endeavor.

Keywords: Reading comprehension · Sentence Verification Technique (SVT) Text tables · Node-link diagrams · Quantitative evaluation

1 Introduction

A situation report (SITREP) is a standardized information or status report that is commonly used by military forces [1] and international organizations [2, 3] for operational purposes. SITREPs present much of their information in narrative or prose form, but many also make use of ancillary tables and diagrams. The display format and/or non-linguistic elements of these latter modes of information generally require readers to employ correspondingly different perceptual and cognitive skills [4].

To gain a better sense of what makes tabular and diagrammatically depicted information easier or harder to read, understand, and retain, we extended a proven procedure for measuring prose comprehension, known as the Sentence Verification Technique (SVT) [5], to these "visual" information formats. We then applied this work to a simulated military SITREP written for non-specialists and conducted an exploratory human performance study to evaluate the merits of this approach and

© Springer International Publishing AG, part of Springer Nature (outside the USA) 2019
S. Nazir et al. (Eds.): AHFE 2018, AISC 785, pp. 1–13, 2019.
https://doi.org/10.1007/978-3-319-93882-0_21

examine the comparative effects of different ways of "quoting" information from a larger visual context.

2 Related Work

The SVT ([5]; all quotes in this paragraph are taken from this reference) was introduced as a method for measuring a reader's comprehension of a prose passage or message. It is based in part on the premise that "the memorial representation of this comprehended message is thought to be in a form which preserves the meaning of the message but not its surface structure." This premise motivates various text alterations for querying comprehension. The SVT procedure entails first presenting a passage to read, giving the reader as much time as needed to peruse and understand the information, and then removing the written material. Subsequently, readers are given a series of sentences based on the written source and are asked to verify whether each sentence is "old" information (i.e., information that was stated in what they read) or "new" information (information that disagrees with or was otherwise not represented in the passage). The SVT defines four types of queries that authors of an SVT test must derive from individual sentences in the prose passage. An *original* query is simply a verbatim copy of any sentence in the reading passage; clearly, this is "old" information. A *paraphrase* query is a sentence "in which as many words as possible were changed without altering the meaning or the syntactical structure of the original sentence." This also constitutes "old" information. Conversely, a *meaning change* query is a sentence "in which one word in the original sentence was altered such that the meaning of the sentence was changed." This change would thus make the sentence "new" information. A *distractor* query is "consistent with the general theme of the passage," but unrelated to the content of any original sentence (making it "new" information). Furthermore, a distractor should have "the same length, syntactical structure, and difficulty level as the original sentence."

The SVT paradigm has already been adapted to measure comprehension in other information display settings. Royer [6] describes its use with speech, which follows naturally for listening to sententially structured spoken material. In our own work, the SVT was recruited to assess listeners' comprehension of rate-accelerated speech. Brock et al. [7, 8] found that the content of digitally accelerated speech from a single talker was comprehended at similar rates to non-accelerated speech and significantly better than that of concurrent talkers. Our present extension of the SVT to tables and node-link diagrams represents an initial attempt to adapt it to media in which propositions and cues are conveyed through means other than sentences. Royer and Cunningham [9] weighed the potential of this idea against the challenge of creating appropriate visual passages and queries.

Mosenthal and Kirsch [10] developed a methodology that scores document readability in part by including the complexity of the structure of tables and graphs. For example, a table with multiple columns is harder to read than a single column list. Intersected lists (tables that require row and column headers), nested lists (multiple layers of headers), and the number of labels a table requires further increase the difficulty of reading a document and, so, its readability score. Equivalencies between

data charts (pie chart, bar graph, line graph, and time line) and various table properties were also defined, but empirical and/or applied evaluations of their method were not discussed. Kosslyn [11] provided guidelines and examples of organizing and labeling charts and node-link diagrams. Kosslyn's work was derived from principles and findings in cognitive psychology, but no formal assessment of difficulty or studies of the guidelines were reported.

3 Testing Comprehension of Tabular and Diagrammatic Information

3.1 Fictional Scenario

The non-specialist SITREP composed for this study outlines a fictional military scenario that delineates an area of responsibility (AOR) assigned to a Marine Corps battalion. The scenario includes the battalion's mission, routine assignments, the duties of its four companies, and a range of local tactical, institutional, and socio-political concerns. The document has an extended prose narrative, a geographical map of the AOR, two informational tables, and a relational node-link diagram. The first table textually outlines the geography, key institutions, and other matters of interest in the AOR with separate entries for urban and rural areas. The second table similarly lists operational locations, assignments, and duties for the companies in the battalion. The diagram depicts important demographic and political entities, in the AOR and the state of their activities and relationship(s) to each other.

3.2 Adapting Sentence Verification Technique Queries

One of our goals was to use the SVT to study readers' comparative understanding of the prose, table and diagrammatic portions of the SITREP. To adapt SVT queries to tables and diagrams we defined equivalent structural units in each format to a sentence of prose. This information was visually "quoted" and an alternative *paraphrase*, *meaning change*, and *distractor* was composed for each *original* quote per the SVT definitions for queries. Table "sentences" corresponded to single table entries (i.e., a cell) and only the cell text was altered. Diagram "sentences" entailed two nodes and an edge, the latter being the "verb," and both node labels and edge types could be altered depending on the query. To preserve the relevant coordinating information in each format, unaltered row and column headers and the diagram's legend were always respectively included in these queries.

Another aspect of tables and diagrams we felt warranted consideration is the cueing function visual layout may have in readers' coded understanding of information in these formats (cf., [12]). To study whether removing different degrees of this cue would affect performance, three levels of this factor, which we refer to as "contextual display format" (or more simply as "context") were devised. These levels are designated as *with context*, *minus context*, and *context free*. Figure 1a shows a diagram used in our study's training material along with examples of how our context definitions apply to diagram queries. Queries shown *with context* include grayed-out versions of the

complete diagram (Fig. 1b), which preserves the full layout, the information being queried, and the legend. However, grayed-out text is replaced by ellipses and grayed-out edges are replaced by undirected dotted lines. The next level (Fig. 1c), designated *minus context*, retains only the original layout of what is being queried and removes the remainder of the source diagram. *Context free* queries (Fig. 1d) are redrawn more compactly to show the information being queried as a straightforward assertion, independent of its original visual context. As a comparative baseline for these three contextual levels of visual queries, we also derived—as a fourth level of "context"—an equivalent set of linguistic queries for each visual quote (i.e., a lexical sentence composed to convey the visually quoted portion of a table or diagram and a corresponding *paraphrase*, *meaning change*, and *distractor* derived from this sentence). Respectively, these queries incorporated either a table cell's relevant coordinating information (its row and column headers) or the relevant definitional information from the diagram's legend and/or its title. Not to be confused with "prose queries"—queries specifically derived from quotes of the prose portion of the SITREP that are also included in the study (see below)—we refer to these lexical characterizations of queries derived from visually quoted tabular and diagrammatic information in the SITREP as *written* queries. Our study's design choices are not the only way to form queries consistent with the SVT rubric; for example, in our "table queries" and "diagram queries" respectively, the relevant headers or the legend could also have been altered per the SVT's rules for query types. We leave these alternatives for future work.

3.3 User Study

With our Institutional Review Board's approval, we recruited 32 volunteers (21 male, 11 female) from the scientific and clerical staff at our laboratory. Participants ranged in age from 23 to 68, with a mean of 42 and median of 38. Three were non-native speakers, but each had at least 25 years of experience speaking and writing English.

After giving their informed consent and completing a pre-study questionnaire on relevant skills, participants worked through an interactive training session that introduced the SVT and our extensions of it to visual quotes from tables and diagrams. The three levels of visual context shown in Fig. 1 (*with context*, *minus context*, and *context free*) and the corresponding *written* queries used as a comparative baseline for visual queries were also introduced at this point. Next, participants read a short prose passage on international energy use that included an informational table and a relational diagram. Then they practiced for the study by verifying a comprehensive series of queries on this material. Explanatory feedback was provided after each response, but only for these practice queries. The formal portion of the study with the simulated SITREP followed. Participants then completed a post-study questionnaire and were debriefed.

The design of the formal study was as follows. Twenty-four informational quotes were selected from the SITREP: eight were sentences drawn from its prose material, eight (in total) were cells taken equally from its two tables, and eight were relationships depicted in its node-link diagram; no quotes were drawn from the map of the AOR. Per the SVT, these quotes formed a base set of *original* queries for each mode of information (i.e., prose, table, and diagram). Next, a corresponding *paraphrase*, *meaning change*, and *distractor* query was developed for each quote. Each of the resulting 32

table and 32 diagram queries was then rendered in the four contextual display formats: *with context*, *minus context*, *context free*, and *written*. The study thus entailed 32 possible prose queries, 128 possible table queries, and 128 possible diagram queries. Participants verified only one query for each the 24 quotes (eight prose, eight table, and eight diagram queries), so coverage of the full set of 288 queries was accomplished through a counterbalanced set of 16 manipulations. Each manipulation was completed by two participants and featured two instances of each type of query in each information mode and, for table and diagram queries, two instances of each level of contextual display.

(a)

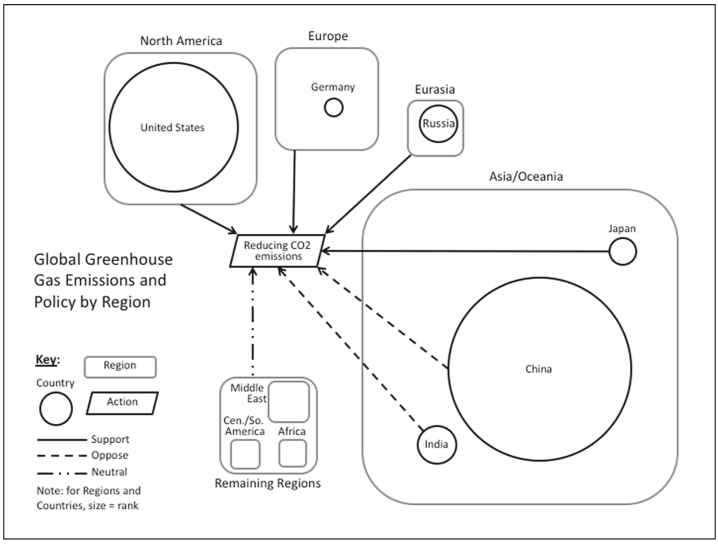
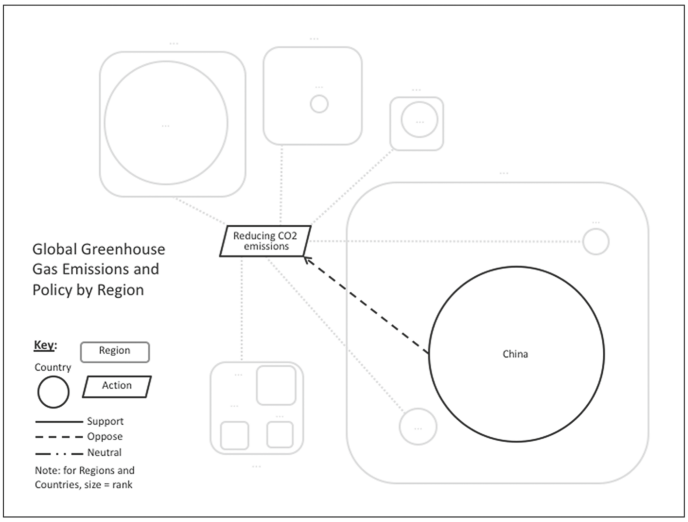
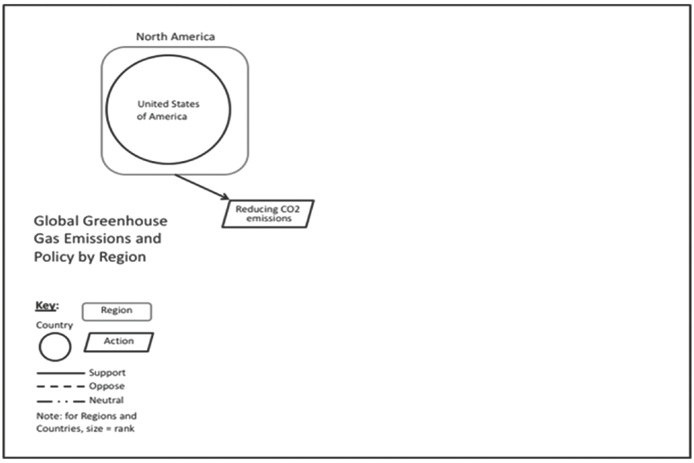


Fig. 1. Practice diagram included in the study's training materials and successive SVT query types respectively cast in the study's four levels of contextual display. (a) The entire diagram as it appeared in the information passage. (b) An *original* query shown *with context*. Note that the diagram's layout and its legend and title are preserved, but unrelated nodes and edges are grayed-out and extraneous labels are changed to ellipses. (c) A *paraphrase* query ("United States of America" paraphrases "United States") shown *minus context*. Again (cf. (b)), the source diagram's layout, legend, and title are preserved, but unrelated nodes, edges, and labels are removed. (d) A *meaning change* query (use of the "Support" edge contradicts the "Neutral" edge appearing in the source diagram (a)) shown *context free*. The legend and title are preserved but the source diagram's layout is dropped and the nodes and edges forming the query are compactly repositioned. (e) A *distractor* query (Korea is not in the source diagram (a)) in *written* form. The query is verbally derived from the information domain shown in the diagram. Although the diagram's coordinating legend and title are shown, unmodified, in visual queries (b), (c), and (d) —regardless of each query's level of "contextual formatting"—the title is the only relevant coordinating information referenced in *written* query (e). Bounding boxes shown in this figure indicate the extent of the source, (a), and each visual query, (b), (c), and (d), participants saw but were not shown in the study.

(b)



(c)



(d)

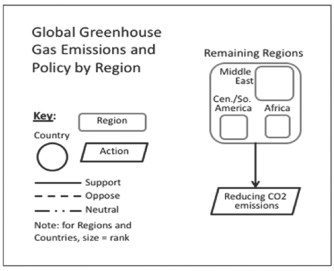


Fig. 1. (*continued*)

4 Results

Two of the study's chief concerns were: whether our SVT extensions would work for tables and diagrams and, if so, whether the SVT query types would yield reasonable levels of performance. Most participants completed the study within an hour; three participants took longer, primarily spending additional time on the practice material and reading the SITREP. Further details of our analysis (and design) can be found in [13].

Using repeated-measures analysis of variance (ANOVA), we conducted a series of analyses and focused on differences in error and response time due to three factors: SVT query type, source material mode, and contextual display format ("context"). We report test statistics, Greenhouse-Geisser corrections to *p*-values, generalized effect size (η_G^2), and post-hoc paired t-tests. $\alpha = 0.05$ is used for significance; values of *p* between 0.05 and 0.10 are noted as "marginally significant." Analyses of response time were conducted only with correct responses.

In the course of the analysis, we identified an intended *meaning change* corresponding to one of the prose quotes as erroneously formed. All eight participants who saw this query assessed it as "old" information. A careful review of the SITREP revealed that the *meaning change* arguably paraphrased a conflation of information in one of the tables. Consequently, this query was removed from the analysis.

4.1 Effect of SVT Query Type

Over the full SITREP, there was no main effect of SVT query type (*original*, *paraphrase*, *meaning change*, *distractor*) on error rate $F(3,93) = 1.188$, $p = 0.317$, $\eta_G^2 = 0.023$. There was a marginally significant main effect of SVT query type on the response time $F(3,93) = 2.622$, $p = 0.077$, $\eta_G^2 = 0.022$. The mean response time for *distractor* queries was lower than for *paraphrase* queries, $t(31) = 2.392$, $p = 0.023$, for *original* queries, $t(31) = 2.372$, $p = 0.024$, and for *meaning change* queries, $t(31) = 2.378$, $p = 0.024$. Tables 1 and 2 respectively report mean error rate and mean response time by query type and source.

Table 1. Error rate by query and source material modes.

	Original	Paraphrase	Meaning change	Distractor	Total
Prose	0.094	0.141	0.143	0.172	0.137
Tables	0.141	0.141	0.266	0.234	0.195
Diagram	0.125	0.156	0.156	0.063	0.125
All sources	0.120	0.146	0.190	0.156	0.153

Table 2. Response times (sec.) by query and source material modes.

	Original	Paraphrase	Meaning change	Distractor	Total
Prose	08.87	08.30	08.86	08.81	08.71
Tables	13.67	13.92	13.43	12.68	13.44
Diagram	11.53	11.78	11.43	09.25	10.96
All sources	11.31	11.33	11.23	10.14	11.00

4.2 Effect of Source Material Mode

There was a significant main effect of the mode of SITREP source material on error rate F(2,62) = 3.207, p = 0.049, $\eta^2_G = 0.044$. For these data (see Table 1), error on queries about the diagram was significantly lower than on queries about the tables, t(31) = 2.561, p = 0.016. Error on queries about the SITREP's prose material was lower than on queries about its tables by a marginally significant amount, t(31) = 1.797, p = 0.082.

There was also a significant main effect of source material on response time for correct responses F(2,62) = 25.501, $p < 0.001$, $\eta^2_G = 0.169$. All pairwise differences for these data (see Table 2) passed a t-test. Participants were significantly faster on prose queries than on diagram queries, t(31) = 3.333, p = 0.002; they were in turn significantly faster on diagram queries than on table queries, t(31) = 3.843, p = 0.001. (Participants were significantly faster on prose queries than on table queries, t(31) = 7.067, $p < 0.001$.) Figure 2 plots group means for error and response time by source material mode.

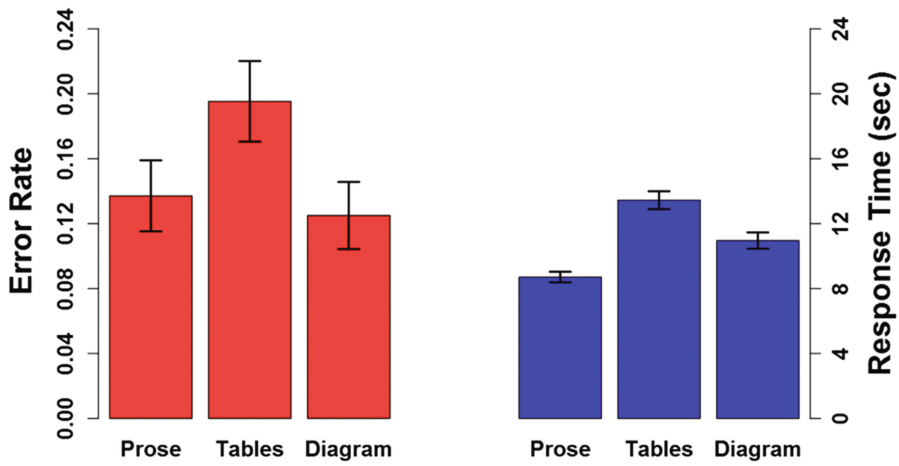


Fig. 2. Group means for error rate (left, red) and response time (right, blue) by source. There were significant main effects on error rate and response time. In pairwise comparisons of error, only the difference between diagram and table queries was significant. All pairwise comparisons of response time were significant. Error bars show the standard error of the mean.

Characteristic differences between these data and patterns seen in earlier studies with the SVT are worth noting. Royer [6] observed a mean error rate of 0.25 (average of all query types), but error rates in the present study are lower by as much as half for all three source material modes (Table 1). In earlier work, Royer [5] also observed that *original* and *distractor* queries were answered correctly at higher rates than *paraphrase* and *meaning change* queries. A plausible explanation for this pattern is that a secondary stage of thought may be required after recognizing that *paraphrase* and *meaning change* queries are only somewhat like quotes of the source materials.

Given that our measures (errors and time to respond correctly) are not Royer's (proportion correct), this trend is nevertheless only weakly evident in our error rates for diagram queries and is not observed in our prose and table query data. Moreover, no evidence of our secondary stage argument—a lag needed for additional thought—is seen in our response time data. Although additional work is needed, key factors in these disparities may be sample size, statistical power, and underlying performance differences in the populations studied by Royer et al. (primary school students) and the largely college-educated population sampled here, with many holding graduate degrees.

4.3 Effect of Contextual Display Format for Tables and Diagram

There was no main effect of the contextual display formats (*with context*, *minus context*, *context free*, and *written*, described above in Subsect. 3.2) on the error rate F (3,93) = 1.261, p = 0.293, $\eta^2_G = 0.024$. There was a marginally significant interaction between source (prose, table, diagram) and context for error rate F(8,248) = 1.817, p = 0.093, $\eta^2_G = 0.043$.

There was a main effect of the contextual display format on the response time for correct responses F(3,93) = 6.384, p = 0.001, $\eta^2_G = 0.056$. Performance with the *written* contextual display format was significantly faster than *with context*, *minus context*, and *context free*, t(31) > 2.452, p < 0.020. Also, *with context* was marginally faster than *context free*, t(31) = 1.712, p = 0.097.

There was also a significant interaction between source and context for response time F(8,208) = 4.765, p = 0.001, $\eta^2_G = 0.082$. In paired comparisons, queries about prose-based information were answered correctly significantly faster than table queries in all contextual display formats, t(26) > 4.395, p < 0.001[1]. Prose queries were also answered significantly faster than diagram queries in the *context free* and *minus context* formats, t(26) > 3.335, p < 0.003. No significant differences were found among paired comparisons of response times for diagram queries in each contextual format. Mean response times for table queries were slower than all others in the study. Among these, queries in the *with context* format were answered significantly faster than queries in the *written* format, t(26) = 2.620, p = 0.014, and marginally faster than queries in the *minus context* and *context free* formats, t(26) = 2.001, p = 0.055 and t(26) = 1.738, p = 0.094, respectively (Table 3).

Table 3. Error rates/response times (sec.) by source material modes and contextual formatting.

Prose	0.137/8.71			
	Written	With context	Minus context	Context free
Tables	0.266/13.38	0.172/11.86	0.203/13.87	0.141/14.63
Diagram	0.141/10.20	0.109/11.07	0.172/11.17	0.078/11.35

[1] As was noted earlier, our response time data only corresponds to correct responses. Consequently, five participants who were unable to answer *any* of the queries in one or more of the paired comparisons in this set of t-tests correctly were removed from this part of the analysis.

5 Discussion and Conclusion

Underlying the work in this paper is an applied need for better ways to assess individuals' comprehension of information presented in formats that are different from, but often ancillary to, prose. Most techniques for eliciting a measure of this understanding tend to be time-consuming to compose and also rely heavily on the proficiency of subject matter experts and skilled test writers. To a certain extent, and in spite of its focus on text, the SVT is a proven and successful response to this problem. As a rubric for assessing comprehension, its paradigm of queries that fall into "old" versus "new" information categories is straightforward to implement, and only requires a test writer to know the underlying information domain well enough to derive a plausible set of alternative sentences that do not readily betray their status to a cursory reader. The attractive simplicity of altering source material to query an individual's understanding of it also makes the idea of extending the SVT to other kinds of information seem viable. Doing this, however, presumes that certain functional equivalencies can be made across information formats, e.g., that something like an informational "unit" corresponding to a "sentence" can be defined for, or identified in, information that, on its surface, is not sentence-based.

Our goal in the study presented here was to think through and test some of the ramifications of extending the SVT beyond prose. Our decision to work with tabular and diagrammatic information as it might appear in a SITREP arose from an interest in informational factors that affect decision-making, but we could have as easily chosen comparable, adjunct materials from newspapers, textbooks, or technical reporting. The chief concerns we faced were: determining (a) if (and what) sentence-like units of information could be identified in each of these formats and (b) whether information gleaned from these sources would be sufficiently recalled by a given query without some degree of its original visual context. Addressing the "sentence equivalence" problem was more than a matter of simply depicting individual table cells or single diagrammatic entities. Instead, it was necessary to inventory coordinated relationships that could be depicted and verbalized as assertions. The basic relational pattern identified most frequently in each format was then used as the basis for selecting visual quotes. For tables, this meant showing a table entry with its corresponding row and column headers, and for the diagram, it meant showing node entities linked by an edge together with the diagram's legend. (Other patterns we identified, e.g., multiple informational items in a table cell and grouped nodes, were taken to be variants of sentence equivalence.)

The "visual context" problem, in contrast, centered on the idea of asking readers to evaluate queries displayed as "visual quotes." To be clear, how readers construe context in information formats other than prose is open to question. Along with the knowledge and inferential skills target readers are presumed to have, "context" can arguably be thought of as associative and structural information writers use to coordinate what they want to convey so it can be effectively grasped. Many sentences carry associative references to their prose context with them, and this is generally enough to invoke or recall a sufficient sense of the source's informational cohesion to discern the nature of an SVT query. Most readers, for instance, will likely recognize "the SVT is a

proven and successful response to this fiction" as a *meaning change* of part of a sentence from the beginning of this section because of its tacit reference to the first paragraph's topic, viz., how to better assess comprehension of information in formats other than prose. Structural information in prose is implicit in the ordering of sentences and in larger divisions of text, but its contextual importance for comprehension is arguably not as key as referential cohesion is. In tabular and diagrammatic information formats, however, the situation is somewhat inverted. Structure—how and where content is shown—is a conspicuous and potentially meaningful presentation factor, and references to a table or diagram's informational topic are not necessarily carried by individual sentence-like assertions without the inclusion of additional coordinating details. This thinking, then, and intuitions about locational imagery that is often present in focused recalls of depicted information (cf., [12]) led to a straightforward set of contextual formatting conditions, per Fig. 1 above, that were used to assess the extent to which the source information's layout is or is not needed to most effectively pose visual SVT queries.

Our findings clearly show that the SVT can, in fact, be functionally extended beyond prose and, moreover, that depicting queries based on "visual quotes" is a viable way to evaluate a reader's understanding of both tabular and diagrammatic information. Visual queries took up to six seconds longer to assess than prose, but error rates overall were lower and somewhat more uniform than expected. To explore these patterns, post-study, we scored each information mode in the SITREP (excluding the map) for readability. Using [14] for the prose portion and [10] for the tables and the diagram, with the latter recast as an adjacency matrix, all three categories were found to be approximately at, but no higher than, a first-year undergraduate level of difficulty and thus well below the general educational level of our readers. This comparative lack of complexity across all three sources may have contributed to the observed pattern of error rates, both broadly and across query types. It cannot, however, account for the clear differences in response times observed in the study. Two factors that may have hindered the table and diagram query assessments are information content and cognitive effort. In our inventory of assertions, we respectively identified 20, 47, and 31 propositions in the prose, table, and diagram portions of the SITREP. The correspondence between these numbers and the pattern of response times in Fig. 2 suggests that readers may have simply needed more time to consider a query when a larger degree of underlying information was involved. In addition to any effort arising from these information loads, the respectively different skills readers need to attend to and encode different media [4] may have also had a role in the pattern of response times. Plausible evidence of this can be glimpsed in the source/context interaction in Sect. 4.3 wherein showing table queries *with context* notably improved response times relative to the other contextual formats. This was one of the outcomes we anticipated, based in part on cognitive performance principles relevant to our experimental design, e.g., Gestalt laws of grouping [15] and encoding specificity [12], but an analogous result was not found for diagram queries. Still, this may only reinforce the fact that there are substantial procedural differences between the practiced skill of reading prose and less frequently exercised skills involved in reading tables and diagrams that continue to warrant further study. Use of a talk-aloud protocol per [16] in future work may help to provide a clearer account of these differences.

Last, this work illuminates an important challenge for comprehension studies with different information formats in an integrated document. Our experimental design required repeated checks to ensure that *meaning changes* and *distractors* did not make assertions that could be unwittingly conflated with the reader's encoding of the SITREP and foil other queries. The prose query that had to be removed from our data (see Sect. 4 preamble) underscores the extent of this risk, and references to use of the SITREP's map in post-study questionnaires show that merging adjunct informational materials into a larger narrative is a common reading and encoding strategy.

Acknowledgements. Wende Frost contributed to early portions of this work. This research was supported by the Naval Research Laboratory Base Program.

References

1. United States Marine Corps: Submission Requirements for the Commanders' Operational Situation Report (SITREP) (Marine Corps Order 3000.2J). Washington, DC (2011)
2. United Nations Children's Fund (UNICEF): The Situation Analysis of Children and Women in Belize 2011: An Ecological Review, Belize City (2011)
3. World Health Organization (WHO): Ebola Situation Reports: Archive (2016). http://www.who.int/csr/disease/ebola/situation-reports/archive/en/
4. Rayner, K.: Eye movements in reading and information processing: 20 years of research. Psychol. Bull. **124**(3), 372–422 (1998)
5. Royer, J.M., Hastings, C.N., Hook, C.: A sentence verification technique for measuring reading comprehension. J. Read. Behav. **11**(4), 355–363 (1979)
6. Royer, J.M.: Developing reading and listening comprehension tests based on the sentence verification technique (SVT). J. Adolesc. Adult Lit. **45**(1), 30–41 (2001)
7. Brock, D., McClimens, B., Trafton, J.G., McCurry, M., Perzanowski, D.: Evaluating listeners' attention to and comprehension of spatialized concurrent and serial talkers at normal and a synthetically faster rate of speech. In: 14th International Conference on Auditory Display. IRCAM, Paris (2008)
8. Brock, D., Wasylyshyn, C., McClimens, B., Perzanowski, D.: Facilitating the watchstander's voice communications task in future navy operations. In: Wilson, D. (ed.) IEEE Military Communications Conference, pp. 2222–2226. IEEE, New York (2011). https://doi.org/10.1109/milcom.2011.6127692
9. Royer, J.M., Cunningham, D.J.: On the theory and measurement of reading comprehension. Technical report No. 91, Center for the Study of Reading, University of Illinois at Urbana-Champaign (1978)
10. Mosenthal, P.B., Kirsch, I.S.: A new measure for assessing document complexity: the PMOSE/IKIRSCH document readability formula. J. Adolesc. Adult Lit. **41**(8), 638–657 (1998)
11. Kosslyn, S.M.: Graph Design for the Eye and Mind. Oxford University Press, New York (2006)
12. Tulving, E., Thomson, D.M.: Encoding specificity and retrieval processes in episodic memory. Psychol. Rev. **80**(5), 352–373 (1973)

13. Livingston, M.A., Brock, D., Maney, T., Perzanowski, D.: Report on an Extension of the Sentence Verification Technique to Multimedia Documents with Tables and Node-link Diagrams. NRL Information Management and Decision Architectures Report IMDA-11 (2018)
14. Kincaid, J.P., Fishburne Jr., R.P., Rogers, R.L., Chissom, B.S.: Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy Enlisted Personnel. Research Branch Report 8–75. Chief of Naval Technical Training: Naval Air Station Memphis (1975)
15. Pinker, S.: A theory of graph comprehension. In: Artificial Intelligence and the Future of Testing, pp. 73–126. Lawrence Erlbaum Associates, Mahwah (1990)
16. Cromley, J.G., Snyder-Hogan, L.E., Luciw-Dubas, U.A.: Cognitive activities in complex science text and diagrams. Contemp. Educ. Psychol. **35**(1), 59–74 (2010). https://doi.org/10.1016/j.cedpsych.2009.10.002