Affective Engagement for Communicative Visualization: Quick and Easy Evaluation using Survey Instruments

Ya-Hsin Hung*
Purdue University

Paul Parsons†
Purdue University

ABSTRACT

As visualization for communication becomes more prevalent, it is important to have ways to evaluate the "success" of communicative visualizations beyond traditional analysis- and performance-oriented approaches. There are many metrics on which the success of communicative visualizations could be viewed, including those related broadly to the user's subjective experience. One construct that has received attention in recent years is user engagement. In this paper, we examine the role of *affective engagement* (AE) in evaluating communicative visualizations. We explore options for assessing AE, and report a literature review on potentially relevant survey instruments. We provide suggestions on how to evaluate AE, discussing steps and analytical methods to develop a self-report assessment based on our ongoing work on AE in information visualization.

Index Terms: Human-centered computing—Visualization—Visualization design and evaluation methods

1 INTRODUCTION

As visual science communication faces a shift from the traditional exploratory vs. explanatory paradigm to their confluence [1, 67], more broad and diverse visualizations have emerged to fill a new communicative role for visualization. This direction is consistent with a growing movement in InfoVis research and practice to expand beyond performance-related concerns to investigate issues related to the broader user experience, where affect, enjoyment, and other non-utilitarian concerns take precedence over performance and usability (e.g., [9, 20, 28, 50]).

One concern within the emerging area of communicative visualization is how to determine when a visualization communicates successfully, and how that "success" can be measured. There are many facets of success for any type of visualization, depending on the goals of the designer and the context of use. For instance, success may be evaluated based on traditional performance metrics and/or non-traditional issues including memorability [4], attitude change [37], persuasion [46], and engagement [20].

While numerous issues are certainly relevant for communicative visualization, our focus here is on the topic of user engagement—specifically, **affective engagement (AE)**. By affective engagement we mean the user's emotional involvement or investment while interacting with a visualization. In recent years, user engagement has received increasing attention within the HCI community [52, 57, 62, 63]. Engagement has been studied from different perspectives by researchers in various disciplines, each having different concerns and viewing engagement through different lenses. From any perspective, user engagement, as with other aspects of subjective experience, is a complex construct that is difficult to define [41].

Recently, as InfoVis scholarship has reached beyond usability driven objectives, investigating aspects of subjective experience is increasingly important [20, 36, 51]. It is reasonable to assume that InfoVis researchers and practitioners are interested in how "engaging" a particular visualization is, wanting to measure levels of engagement to predict or determine success [50]. Behavior-based metrics (e.g., time spent, see [5]) have previously been employed to quantify "engagement" levels in InfoVis.

Evaluation methods for InfoVis should match the goals of the design situation and the context of use [27]. Because many existing evaluation strategies have been aimed at analysis rather than communication, they are not often suitable for evaluating issues relevant for communicative visualization. Thus, there is a need to examine evaluation methods for InfoVis with communication, rather than performance, as the main goal. Here we present AE as one relevant construct for evaluating communicative visualizations, focusing on "quick and easy" ways to evaluate AE.

Our target audience is people looking for quick and easy ways to evaluate AE for communicative visualization. While this target clearly fits visualization practitioners, academics and other researchers often have need for quick and easy evaluation methods. Practitioners often face constraints, such as time, money, and equipment limitations, that make lab-based user testing not feasible. Evaluation methods involving specialized equipment (e.g., eye trackers, EEGs) or considerable money and space to run user studies—while certainly valuable—are outside the scope of our concern here.

Practitioners can benefit from quick and easy evaluation methods that can still provide actionable information regarding AE. When we say we want evaluation to be "quick and easy", we want all stages (i.e., conducting, analyzing, interpreting) of the process to be both quick and easy. By quick, we mean a minimal time spent to conduct the testing, to analyze the collected data, and to make sense of the results for making further decisions. By easy, we mean there is no need for specialized domain knowledge to conduct the testing, no need for specialized equipment to collect the data, and the collected data is easy to process and easy to interpret. With the above criteria, we believe that *a concise self-report survey instrument that can quantify AE* can be an appropriate tool for visualization designers wanting to evaluate AE for communicative purposes.

In what follows, we review existing survey instruments, report the progress of our own work on a survey instrument for AE, explore how practitioners can make use of the instrument in a scenario, and elaborate on the scientific as well as empirical significance of this study. Finally, plans for distributing the final survey instrument and future study will be described.

2 ENGAGEMENT IN INFORMATION VISUALIZATION

In the HCI literature, several theoretical frameworks regarding engagement have been proposed [42]. User engagement has been discussed in relation to constructs such as flow, curiosity, surprise, and joy [58]. It has also been defined as the emotional, cognitive, and behavioral connection that exists between a user and a resource in time or possibly over time [3]. User engagement is also construed as the positive interaction quality of user experience, and has been associated with being captivated and motivated to use a website [30]. Sometimes it is treated as a user's general level of involvement with a product [48]. Related concepts of flow [11, 16], immersion [8], and playfulness [2] have been investigated in related research areas,

*e-mail: hung17@purdue.edu
†e-mail: parsonsp@purdue.edu

	Name	Year	Main Construct	# of items	Communicative effectiveness	Visual aspects	Performance metrics	Engagement metrics	Affect metrics	Commercial version	Specific Platform/Scenario	Intended platform
R1	System usability Scale (SUS) [7]	1986	Usability	10			●		▲			System/Technology
R2	NASA Task Load Index (TLX) [18]	1986	Subjective workload	6			▲		▲			Interface/System
R3	Questionnaire for User Interface Satisfaction (QUIS) [10]	1988	Satisfaction	27	▲	▲	●	▲	▲	✓	✓	Interface
R4	Perceived Usefulness and Ease of Use (PUEU) [12]	1989	Usefulness and Ease of Use	12			●				✓	System/Technology
R5	The After-Scenario Questionnaire (ASQ) [32]	1990	Satisfaction	4			●					System/Technology
R6	The Post-Study System Usability Questionnaire (PSSUQ) [31]	1992	Usability	16	▲	▲	▲	▲	▲			System/Technology
R7	Nielsen's Heuristic Evaluation [40]	1994	Usability	10	▲	▲	●	▲	▲			System/Technology
R8	Computer System Usability Questionnaire (CSUQ) [33]	1995	Usability	19	▲		●		▲			System/Technology
R9	Software Usability Measurement Inventory (SUMI) [24]	1995	User experience	50	▲		●	▲	▲	✓	✓	Software
R10	Presence Questionnaire Item Stems [66]	1998	Presence	32			▲	●	▲		✓	Virtual environment
R11	Website Analysis and Measurement Inventory (WAMMI) Questionnaire [23]	1998	User experience	20	▲		▲	▲	▲	✓	✓	Website
R12	USE Questionnaire [35]	2001	Usability	30			●	▲	▲			System
R13	Unified Theory on Acceptance and Use of Technology (UTAUT) [64]	2003	Technology Acceptance	31			▲	▲	●			Technology
R14	Fun questionnaire [15]	2005	User experience (fun)	14			▲	●	●			Educational system
R15	Cognitive absorption and TAM [49]	2005	TAM+cognitive absorption	22			▲	▲	▲		✓	System/Technology
R16	The Single Ease Question (SEQ) [59]	2006	Ease of use	1			●				✓	System/Technology
R17	Immersive Experience Questionnaire (IEQ)-a [22]	2008	Experience of immersion	31	▲	▲		●	●		✓	Video game
R18	Immersive Experience Questionnaire (IEQ)-b [22]	2008	Experience of immersion	33			▲	●	●		✓	Video game
R19	User Experience Questionnaire (UEQ) [29]	2008	User experience	26	▲	▲	▲	▲	▲	✓		System/Technology
R20	Gaming Engagement Questionnaire (GEQ) [6]	2009	Deep engagement	19				●	●		✓	Video game
R21	The Subjective Mental Effort Question (SMEQ) [54]	2009	Ease of use	1			●				✓	System/Technology
R22	User Engagement Scale (UES) [43]	2010	User engagement	31		▲	▲	●	●			System/Technology
R23	Measurement Model of User Engagement [44]	2015	User engagement	12			▲	●	●		✓	Website
R24	Standardized User Experience Percentile Rank Questionnaire (SUPR-Q) [53]	2015	User experience	8			▲	▲	▲	✓		System/Technology

Figure 1: 24 survey instruments meeting all inclusion criteria, ordered chronologically. Rows are collected survey instruments, columns are characteristics of the survey instruments (● = "mostly satisfied", ▲ = "partially satisfied", and ✓ = "Yes"). Digital version can be found at `https://yahsin.github.io/CommViz/`

some of which are close to the concept of user engagement.

As for measuring user engagement, there are a number of physiological indicators that might be related to engagement; metrics such as blood pressure, heart rate, and nervous system activity [25] all can be directly measured. However, costs and effort to conduct such measurements are considerably high, limiting their practicality. Some behavioral indicators (e.g., mouse clicks, page visits, time spent) have been used as indicators of users' subjective experiences as well [19,48]. Still, considerable interpretation is required to make causal connections to subjective phenomena [65].

Self-report refers to methods that rely on what users say or recall from their experience. Interviews, one of the most common and powerful methods for understanding people and collecting qualitative data, have been used for identifying attributes of user engagement in technology [26]. In general, interviews are recorded (voice or video) and transcribed into text format, then analyzed qualitatively [14]. Another way to collect qualitative data from users is via verbal protocols, which ask people to verbalize their thoughts and feelings, with an aim to get insight into the participant's cognitive processes [17,45].

Advantages of self-reporting methods include interpretability, information richness, and practicality [47]. Additionally, for affective traits that cannot be directly observed (e.g., AE), self-report methods provide more interpretable and richer information than performance metrics such as heart rate or pupillary response [60]. A handful of surveys or questionnaires for assessing engagement in various disciplines have been proposed. For example, in the gaming community, the Immersive Experience Questionnaire (IEQ) [22] and the Gaming Engagement Questionnaire (GEQ) [6] have been proposed to evaluate immersion and engagement respectively. Models and frameworks relating to engagement have been proposed for online services and web interfaces [3]. Questionnaires and surveys in information science and technology have also been proposed (e.g., [43]). Related constructs, such as flow [39] and playfulness [34] have received considerable efforts at measurement and assessment.

3 Existing Survey Instruments

To provide context before presenting our own work, we first investigate established self-report instruments that researchers and practitioners can use to evaluate their visualizations. To do so, we conducted a brief survey of relevant evaluation instruments. There are two intentions here: (1) a collection of these instruments can be a valuable resource on its own, especially for visualization practitioners who might not have easy access to surveys of this kind; and (2) the survey helps to highlight where gaps might be—for communicative issues in general, and AE in particular, in the context of visualization for communication.

To conduct the survey, the authors searched for and collected relevant self-report instruments. The initial search was very broad; besides some general resources from HCI and UX handbooks [55,56,61], we also searched online using the following keywords: "visualization", "user experience", "engagement", "communication", "persuasion", "emotion", "survey", "questionnaire", "scale" and their various combinations. Although we found many instruments related to communication, affect, satisfaction, and various psychological constructs, we excluded all that were not concerned with human-technology relationships. Thus we excluded instruments dealing with constructs such as human-human communication, anxiety, customer satisfaction, and so on.

In the end, we settled on 3 inclusion criteria—each instrument should: (1) be concerned with human-technology relationships; (2) be associated with a publication, (3) not require specialized equipment. The authors met multiple times to identify characteristics that are relevant for communicative visualization and could be used to

code the instruments. The characteristics include whether an instrument is concerned with (a) communicative effectiveness of the technology; (b) visual aspects of the technology; (c) performance (e.g., time, error); (d) user engagement; (e) affect; (f) a particular platform or scenario; and (f) whether the instrument has a commercial version that needs to be purchased. Examples of the characteristics include: communicative effectiveness: "*Prompts for input is confusing/clear.*" (R3 [10]); visual aspects: "*The screen layout of this website is visually pleasing.*" (R22 [43]); performance metrics: "*I can recover from mistakes easily and quickly.*" (R12 [35]); engagement metrics: "*I really get into the game.*" (R20 [6]); affect metrics: "*The system is somewhat intimidating to me.*" (R13 [64]).

At this point our review is not exhaustive, yet due to our systematic approach, we believe it is reasonably representative of a more complete sample. By following two strategies—investigating popular books and conducting our own search—we believe that we have covered at least the popular and well-established instruments. In follow-up work we plan to do a more exhaustive review.

3.1 Findings

Our search resulted in 24 instruments that met the inclusion criteria. A summary of these is shown in Fig. 1. For each instrument, we list the name, publication year, the construct being evaluated by the instrument, the total number of items (questions) or heuristics included, and the instrument characteristics described previously.

In general, most of the collected instruments deal with constructs like usability and user experience, and more than half are developed based on general system/technology/artifact platforms. Usability oriented instruments usually have substantial numbers of performance-related items and few items related to affect or engagement (e.g., R3 [10], R8 [33]). On the other hand, engagement-related instruments have more focus on affect and engagement, with fewer performance-related items. (e.g., R20 [6], R22 [43])

The contexts in which communicative visualizations are used are different from many of the instruments in Fig. 1 (e.g., to influence or to persuade viewers). Thus, while the compiled instruments may be a useful resource for evaluating communicative visualizations, we believe that the existing instruments are not entirely suitable for the following reasons:

- **Scope:** Most surveys that include affect- or engagement-related items aim to cover a much broader construct. Thus the relevant information gained about AE may not be very substantial (e.g., may be related to only 1 or 2 items). Also, there are problems with using only portions of an instrument without using it in its full and originally intended context.
- **Specific media or environment:** Some instruments are measuring a construct that is tied to a specific medium or context of use that is not very relevant for communicative visualization (e.g., video game [6], presence in virtual environment [66]).
- **Context of measurement scale:** Although some instruments share similar key factors, the context of their measurement is not always appropriate for communicative visualization. For example, "captivation" may be a sub-component for both "immersion" and "engagement". However, an item that asks "I felt detached from the outside world" (see [22]) is likely not appropriate for communicative visualization, yet makes sense for assessing engagement in virtual reality.
- **Length:** Some instruments contain a high number of items, and may take substantial time to answer and administer. Although more items may be desirable for precision, long instruments may not be "quick and easy", quickly becoming a barrier for practitioner use.

4 Current Work

In our previous work [20], we reviewed literature related to user engagement and technology across disciplines such as website analysis, game design, education, psychology, and HCI, and compiled a list of potentially relevant characteristics—57 in total. We merged the overlapping ones and removed ones that were too broad or vague (e.g., entertaining, usability) as well as ones only relevant in specific contexts (e.g., brutality, eroticism). In the end, we compiled 11 preliminary characteristics that had the highest frequency in the literature and were most relevant to engagement in InfoVis: *aesthetics, captivation, challenge, control, discovery, exploration, creativity, attention, interest, novelty, and autotelism.*

However, as we tried to further develop a formal survey instrument to assess AE, several concerns emerged:

1. Most literature was from outside of InfoVis, thus many visualization-related aspects (e.g., visual elements or underlying data) were not well-covered. This can be seen in Fig. 1.
2. No common understanding of engagement exists; thus, the elicited characteristics could be from different constructs—even though they may all be labeled as "engagement".
3. Although our process was systematic, our findings were not grounded in empirical observation.
4. Our identified "characteristics" were not equal to the sub-components of the construct, which are necessary to identify for development of a reliable instrument [60].

Due to the above concerns, we decided to conducted a new mixed-methods study that involved 25 participants interacting with visualizations for both explanatory and exploratory contexts. Data from think-aloud, eye-tracking, questionnaires, and semi-structured interviews were collected and analyzed using a grounded theory approach for qualitative data and triangulation with qualitative data [21]. While the analysis is still work-in-progress, with InfoVis context in mind, we identified behaviors or activities that can serve as potential indicators of proposed key factors. Some examples of behavior indicators and their corresponding key factors are shown in Fig. 2.

Key factor	Behavior indicators (examples)
Discovery	▪ Learn something that not known before (e.g., a new fact, concept, or piece of information). ▪ Figure out how to use the InfoVis along the way.
Creativity	▪ Imagine things not directly related to what can be seen in the InfoVis. ▪ Generate new and original thoughts or ideas.
Interest	▪ Feel the content (e.g., topic, message) of the InfoVis is interesting. ▪ Feel the features (e.g., interactions, animations) provided in the InfoVis is interesting.

Figure 2: Some examples of behavior indicators and corresponding key factors for affective engagement in information visualization. Items will be written to assess those behaviours and quantify into scores. Please note that this list is only an excerpt of the full list.

4.1 Development of Survey Instrument

In this section, we briefly explain our ongoing work developing a quick and easy survey instrument for assessing AE in InfoVis. Since AE is a complex and unobservable (latent) construct, to measure or assess it there are several steps to go through [38]. As shown in Fig. 3, first, the (latent) target construct needs to be decomposed into several sub-components/key factors based on its conceptual space. Second, the observable behavior indicators need to be established for each sub-component/key factor. Finally, based on those behavior indicators, items (questions) can be written that can locate respondents' level on the indicator scale.

There are some general steps to develop an instrument to assess an affective construct [38, 60]. Three primary processes of measurement development need to be employed: (1) identify intended use

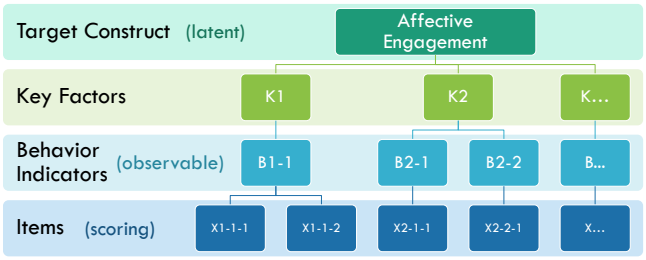

Figure 3: Structural overview of the survey instrument for affective engagement in InfoVis, from top to bottom: Target construct (affective engagement), Key factors (denoted as K), Behavior indicators (denoted as B), and Items (denoted as X). As it is possible to have multiple key factors for a construct, it is also possible to have multiple behavior indicators and items from their previous layer.

of test scores; (2) specify score interpretation related to AE and test-taker population(s); and (3) develop a conceptual definition (space) of the target construct and establish a list of behaviors that are taken to indicate a person's location on the trait continuum (see Fig. 3). Subsequently, expert reviews and a pilot tryout can be conducted. Finally, the instrument will be field tested with respondents.

Analytical Methods With the collected data from a field test, several statistical techniques can be employed for preliminary diagnosis: classical item difficulty (p-value) as each binary item's mean response value, and biserial correlation between each item and the sum score with item removed. Furthermore, Reliability Coefficient (Omega) can also be helpful, where reliability is a property of observed test scores from a particular instrument in a specified examinee population.

- **Factor Analysis (FA):** Results from FA could be helpful for testing internal structure and external structure of the instrument, or making decisions on whether to remove or revise certain instrument items [38].
- **Item Response Theory (IRT):** An IRT model (e.g., Rasch model, Nominal response model, Graded response model) is able to show the relationship between the ability or trait (target construct) measured by the instrument and an item response [13]. Researchers can retrieve more information about items' characteristics (e.g., discrimination ability, test information function) compared to traditional CTT (Classical test theory) method.

5 Using our Survey Instrument

Consider a scenario where a visualization practitioner wants to evaluate their communicative visualization (e.g., an interactive visualization incorporated with an online magazine article) according to levels of AE within a target group of users. The practitioner can recruit a group of respondents (more is generally better, but size can be adjusted depending on resources and other factors) from their target population (e.g., readers of that online magazine). By asking respondents to answer the items after interacting with the visualization, *the practitioner can calculate the level of AE of those particular participants. By averaging scores, user's AE levels can be estimated for that visualization*. Fig. 4 (a) provides a visual depiction of an evaluation scenario where a survey instrument is being used to assess AE.

Visualization practitioners can make use of survey instruments at various stages to their design process. For instance, when a working (functional) prototype is ready, a visualization designer can conduct an evaluation session, either as a tryout or a more structured user testing described as follows:

- **Pilot tryout**: A small number of participants will be invited to use the working prototype (e.g, free exploration, tryout specific features) and then provide their comments and opinions. Generally, this type of quick tryout works best when the designer requires instant feedback (i.e. the work is at an early stage).
- **User Testing**: A group of participants will be recruited and asked to conduct specific tasks (e.g., solve problems, interpret visualizations, identify insights). Performance data (e.g., task accuracy and efficiency) and/or subjective responses (e.g., interview and survey) are collected via several metrics. The assigned tasks are often more structured than pilot tryout and tend to be conducted when a more thorough investigation is necessary (i.e. competitor comparison).

Fig. 4 (b) shows how a designer can utilize a survey for a pilot tryout and for user testing. Note that for both cases, the evaluation can be conducted on-site (e.g., lab study) or remotely (e.g., online crowd-sourcing). A short self-report survey instrument (with roughly 10 items) will not take too much time, which makes a larger scale user testing more feasible (e.g., online crowd sourcing).

Even for a simple pilot tryout session, one potential benefit to employing survey instruments is that the listed items or key factors in it can stimulate rich feedback from participants. Furthermore, this scenario (see Fig. 4 (b)) also demonstrates how other performance measurements (e.g., error rate) and behavior observation methods (e.g., eye-tracking) can be integrated if more data is required.

5.1 Limitations

We have briefly discussed why and how a self-report survey instrument that assesses AE for InfoVis can be beneficial for communicative visualization, and can be a useful option for visualization evaluation. Still, there are some limitations of this approach that should be noted:

- **Interpretation:** The survey instrument is not meant to measure a visualization's AE, or any property of an artifact. Instead, it is measuring respondents' latent construct that consists of their emotional involvement, which is "labeled" as AE. Thus, the survey result cannot be interpreted as indicating the visualization's quality; instead, we have to weigh the value of the visualization against other considerations (e.g., tasks, goals).
- **Administration:** With different populations of respondents, the scores on the same visualization are expected to be different. Since *scores here only represent AE levels of respondents that have been chosen*, the recruitment of appropriate respondents (i.e. sampling proper participants from target population), is considerably important in influencing the results of the survey.
- **Usage:** This survey instrument is intended to measure AE for communicative visualizations, not visualizations for analysis. User intention, context, or motivation can influence AE—e.g., a high-stakes task with safety implications as part of the user's job will impact AE in ways different from low-stakes news stories for casual users. Thus, the use of the instrument is for communicative situations/scenario where stakes are low and visualization use is non-utilitarian is important.

6 Future Work and Implications

In the future, we plan to conduct a follow-up lab study that investigates potential correlations between AE and other relevant factors (e.g., memorability and perceived usability). The follow-up experiment can help us: (1) evaluate the AE instrument developed in this project, and (2) learn how AE relates to other user or visualization characteristics. The results will be able to tell us whether the user's

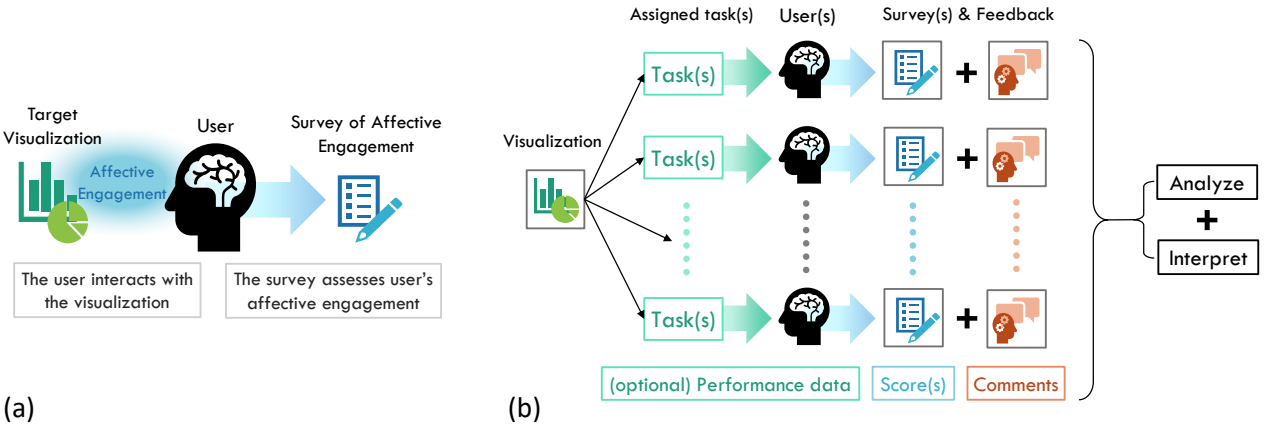

Figure 4: (a) How a survey instrument assesses single user's affective engagement on one visualization. (b) The evaluation scenario of utilizing survey instrument to assess multiple users' AE on a communicative visualization. User study results including survey instruments scores, user's subjective feedback, and (optional) user's performance data could be collected along the way.

AE level can predict other important factors relevant for communicative visualization, or vice versa, and serve as another piece of validity evidence of the AE instrument.

Regarding distribution of the final survey instrument, we plan to upload a freely available digital version (for both online and paper use) to our project website. Tools (e.g., simple spreadsheet) to help users analyze and interpret the result scores will also be provided.

With this instrument, the utilityt and implications of AE can possibly be investigated. Correlations between AE and other factors that are important for communicative visualization may also be investigated. The instrument can hopefully become a handy tool for visualization researchers and designers to evaluate their work quickly and easily. In the future, this research may expand the theoretical basis of engagement in the fields of human-computer interaction and information visualization.

REFERENCES

[1] Alyssa Goodman. The road from explanation to exploration, and back. `https://youtu.be/OfA7XkivjAA`, 2018.

[2] J. Arrasvuori, M. Boberg, J. Holopainen, H. Korhonen, A. Lucero, and M. Montola. Applying the PLEX Framework in Designing for Playfulness. DPPI '11, pages 24:1–24:8. ACM, 2011.

[3] S. Attfield, G. Kazai, M. Lalmas, and B. Piwowarski. Towards a science of user engagement (position paper). In *WSDM workshop on user modelling for Web applications*, pages 9–12, 2011.

[4] M. A. Borkin, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, and H. Pfister. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2306–2315, 2013.

[5] J. Boy, F. Detienne, and J.-D. Fekete. Storytelling in Information Visualizations: Does It Engage Users to Explore Data? In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 1449–1458, New York, NY, USA, 2015. ACM.

[6] J. H. Brockmyer, C. M. Fox, K. A. Curtiss, E. McBroom, K. M. Burkhart, and J. N. Pidruzny. The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, 45(4):624–634, jul 2009.

[7] J. Brooke. SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[8] E. Brown and P. Cairns. A Grounded Investigation of Game Immersion. In *Proceedings of ACM CHI 2004 Conference on Human Factors in Computing Systems*, pages 1297–1300. ACM Press, 2004.

[9] N. Cawthon and A. V. Moere. A conceptual model for evaluating aesthetic effect within the user experience of information visualization. In *Information Visualization, 2006. IV 2006. Tenth International Conference on*, pages 374–382. IEEE, 2006.

[10] J. P. Chin, V. A. Diehl, and L. K. Norman. Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '88*, pages 213–218, New York, USA, 1988. ACM Press.

[11] M. Csikszentmihalyi. The flow experience and its significance for human psychology. In *Optimal experience*. Cambridge University Press, 1988.

[12] F. D. Davis. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3):319–340, 1989.

[13] R. J. de Ayala. *The Theory and Practice of Item Response Theory*. Methodology in the Social Sciences. Guilford Publications, 2013.

[14] B. DiCicco-Bloom and B. F. Crabtree. The qualitative research interview. *Medical Education*, 40(4):314–321, 2006.

[15] A. Donker. *Human factors in educational software for young children*. PhD thesis, 2005.

[16] Y. Douglas and A. Hargadon. The Pleasure Principle: Immersion, Engagement, Flow. HYPERTEXT '00, pages 153–160. ACM, 2000.

[17] K. A. Ericsson and H. A. Simon. Verbal reports as data. *Psychological Review*, 87(3):215–251, 1980.

[18] S. G. Hart and L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock and N. B. T. A. i. P. Meshkati, editors, *Human Mental Workload*, volume 52, pages 139–183. North-Holland, 1988.

[19] J. Huang, R. W. White, and S. Dumais. No clicks, no problem: using cursor movements to understand and improve search. page 1225. ACM Press, 2011.

[20] Y.-H. Hung and P. Parsons. Assessing User Engagement in Information Visualization. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17*, pages 1708–1717. ACM Press, 2017.

[21] Y.-H. Hung and P. Parsons. Evaluating User Engagement in Information Visualization Using Mixed Methods. In *IEEE VIS '17: Proceedings of the 2017 IEEE Conference on Information Visualization*, Phoenix, Arizona, USA, 2017.

[22] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9):641–661, 2008.

[23] J. Kirakowski and B. Cierlik. Measuring the Usability of Web Sites. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 42(4):424–428, oct 1998.

[24] J. Kirakowski and M. Corbett. *SUMI: the Software Usability Measurement Inventory*, volume 24. oct 2006.

[25] S. D. Kreibig. Autonomic nervous system activity in emotion: A review. *Biological Psychology*, 84(3):394–421, jul 2010.

[26] M. Lalmas, H. O'Brien, and E. Yom-Tov. Measuring User Engagement. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 6(4):1–132, nov 2014.

[27] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE transactions on visualization and computer graphics*, 18(9):1520–1536, 2012.

[28] H. Lam and T. Munzner. Increasing the utility of quantitative empirical studies for meta-analysis. In *Proceedings of the 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization*, page 2. ACM, 2008.

[29] B. Laugwitz, T. Held, and M. Schrepp. Construction and Evaluation of a User Experience Questionnaire. pages 63–76. Springer, Berlin, Heidelberg, nov 2008.

[30] J. Lehmann, M. Lalmas, E. Yom-Tov, and G. Dupret. Models of User Engagement. In J. Masthoff, B. Mobasher, M. C. Desmarais, and R. Nkambou, editors, *User Modeling, Adaptation, and Personalization*, Lecture Notes in Computer Science, pages 164–175. Springer Berlin Heidelberg, 2012.

[31] J. R. Lewis. Psychometric Evaluation of the Post-Study System Usability Questionnaire: The PSSUQ. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 36(16):1259–1260, oct 1992.

[32] J. R. Lewis and J. R. Psychometric evaluation of an after-scenario questionnaire for computer usability studies. *ACM SIGCHI Bulletin*, 23(1):78–81, dec 1990.

[33] J. R. Lewis and J. R. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7(1):57–78, jan 1995.

[34] A. Lucero, J. Holopainen, E. Ollila, R. Suomela, and E. Karapanos. The Playful Experiences (PLEX) Framework As a Guide for Expert Evaluation. In *Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces*, DPPI '13, pages 221–230, New York, USA, 2013. ACM.

[35] A. M. Lund. Measuring usability with the use questionnaire. *Usability interface*, 8(2):3–6, 2001.

[36] N. Mahyar, S.-H. Kim, and B. C. Kwon. Towards a taxonomy for evaluating user engagement in information visualization. In *Workshop on Personal Visualization: Exploring Everyday Life*, page 2, 2015.

[37] S. Mason and T. Azzam. In need of an attitude adjustment? the role of data visualization in attitude change and evaluation influence. *American Journal of Evaluation*, 2018.

[38] D. B. McCoach, R. K. Gable, and J. P. Madura. *Instrument Development in the Affective Domain: School and Corporate Applications*. Springer Science & Business Media, may 2013.

[39] G. B. Moneta. On the Measurement and Conceptualization of Flow. In S. Engeser, editor, *Advances in Flow Research*, pages 23–50. Springer New York, 2012.

[40] J. Nielsen. Usability inspection methods. In *Conference companion on Human factors in computing systems*, pages 413–414, 1994.

[41] H. O'Brien. Theoretical perspectives on user engagement. In *Why Engagement Matters*, pages 1–26. Springer International Publishing, 2016.

[42] H. L. O'Brien and E. G. Toms. What is User Engagement? A Conceptual Framework for Defining User Engagement with Technology. *J. Am. Soc. Inf. Sci. Technol.*, 59(6):938–955, 2008.

[43] H. L. O'Brien and E. G. Toms. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, 61(1):50–69, jan 2010.

[44] J. Oh, S. Bellur, and S. S. Sundar. Clicking, Assessing, Immersing, and Sharing: An Empirical Model of User Engagement with Interactive Media. *Communication Research*, 45(5):737–763, sep 2015.

[45] E. L. Olmsted-Hawala, E. D. Murphy, S. Hawala, and K. T. Ashenfelter. Think-aloud protocols: A comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2381–2390, New York, USA, 2010. ACM.

[46] A. V. Pandey, A. Manivannan, O. Nov, M. Satterthwaite, and E. Bertini. The Persuasive Power of Data Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2211–2220, 2014.

[47] D. L. Paulhus, S. Vazire, R. W. Robins, R. C. Fraley, and R. F. Krueger. The self-report method. *Handbook of research methods in personality psychology*, 1:224–239, 2007.

[48] K. Rodden, H. Hutchinson, and X. Fu. Measuring the User Experience on a Large Scale: User-centered Metrics for Web Applications. CHI '10, pages 2395–2398. ACM, 2010.

[49] R. Saadé and B. Bahli. The impact of cognitive absorption on perceived usefulness and perceived ease of use in on-line learning: an extension of the technology acceptance model. *Information & Management*, 42(2):317–327, 2005.

[50] B. Saket, A. Endert, and J. Stasko. Beyond usability and performance: A review of user experience-focused evaluations in visualization. In *Proceedings of the Beyond Time and Errors on Novel Evaluation Methods for Visualization - BELIV '16*, pages 133–142, New York, USA, 2016. ACM Press.

[51] B. Saket, C. Scheidegger, and S. Kobourov. Towards Understanding Enjoyment and Flow in Information Visualization. *arXiv:1503.00582 [cs]*, mar 2015.

[52] K. Salehzadeh Niksirat, S. Sarcar, H. Sun, E. L.-C. Law, T. Clemmensen, J. Bardzell, A. Oulasvirta, C. Silpasuwanchai, A. Light, and X. Ren. Approaching engagement towards human-engaged computing. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, pages 1–4, New York, USA, 2018. ACM.

[53] J. Sauro. SUPR-Q: A Comprehensive Measure of the Quality of the Website User Experience. *J. Usability Studies*, 10(2):68–86, feb 2015.

[54] J. Sauro and J. S. Dumas. Comparison of three one-question, post-task usability questionnaires. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*, page 1599, New York, USA, 2009. ACM Press.

[55] J. Sauro and J. R. Lewis. *Quantifying the user experience : practical statistics for user research*. Amsterdam, Netherlands : Morgan Kaufmann, second edi edition, 2016.

[56] B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmqvist, and N. Diakopoulos. *Designing the user interface: strategies for effective human-computer interaction*. Pearson, 2016.

[57] S. S. Sundar, S. Bellur, J. Oh, H. Jia, and H.-S. Kim. Theoretical Importance of Contingency in Human-Computer Interaction: Effects of Message Interactivity on User Engagement. *Communication Research*, 43(5):595–625, may 2014.

[58] A. Sutcliffe. Designing for User Engagement: Aesthetic and Attractive User Interfaces. *Synthesis Lectures on Human-Centered Informatics*, 2(1):1–55, 2009.

[59] D. Tedesco and T. Tullis. A comparison of methods for eliciting post-task subjective ratings in usability testing. *Usability Professionals Association (UPA)*, 2006:1–9, 2006.

[60] R. M. Thorndike and T. M. Thorndike-Christ. *Measurement and Evaluation in Psychology and Education*. Pearson, Boston, 8 edition edition, may 2009.

[61] T. Tullis and W. Albert. *Measuring the User Experience, Second Edition: Collecting, Analyzing, and Presenting Usability Metrics*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2nd edition, 2013.

[62] P. Turner. The anatomy of engagement. In *Proceedings of the 28th Annual European Conference on Cognitive Ergonomics - ECCE '10*, page 59, New York, USA, 2010. ACM Press.

[63] P. Turner. The figure and ground of engagement. *AI & SOCIETY*, 29(1):33–43, 2014.

[64] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis. User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3):425, 2003.

[65] J. S. Wiggins. *Personality and prediction: principles of personality assessment*. Addison-Wesley Pub. Co., 1973.

[66] B. G. Witmer and M. J. Singer. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators & Virtual Environments*, 7(3):225–240, 1998.

[67] A. Ynnerman, J. Löwgren, and L. Tibell. Exploranation: A New Science Communication Paradigm. *IEEE Computer Graphics and Applications*, 38(3):13–20, 2018.