DECSYS – Discrete and Ellipse-based response Capture SYStem

Zack Ellerby‡, Josie McCulloch‡ (Member, IEEE), John Young†, Christian Wagner‡ (Senior Member, IEEE)
‡LUCID Research Group, University of Nottingham, UK
†Sociotechnical Security Group, National Cyber Security Centre, UK
zack.ellerby; josie.mcculloch; john.y@ncsc.gov.uk; christian.wagner@nottingham.ac.uk

Abstract—Data-driven techniques that capture uncertainty through intervals or fuzzy sets can substantially improve systematic reasoning about uncertain information. Recent years have seen renewed interest in the capture of intervals from a variety of sources – including experts and general survey participants. This approach avoids the more cumbersome batteries of questions that are otherwise required to capture individual uncertainty, and which may not obtain the same degree of fidelity. It also enables respondents to effectively communicate any range (e.g., vagueness) inherent in their response, allowing generation of models that represent this additional information. However, manual methods of obtaining and processing interval-valued data – such as through paper-based questionnaires, are labour and time intensive. This has provided a practical barrier to adoption of interval-valued response-formats in the wider community, from research to industry (e.g., marketing). We argue that establishing an effective and accessible method for interval-valued data-capture will greatly encourage research in and application of uncertainty-aware models of data. Thus, we present DECSYS, a newly developed open-source software tool, which enables the creation and administration of digital surveys that elicit both conventional and interval-valued responses. DECSYS incorporates a range of features, and is designed to maximise versatility for experimenters and usability for participants. Surveys can be conducted either locally or online, and results easily exported. We welcome community feedback, including on how to best tailor the tool in the future to maximise value and support multidisciplinary adoption of uncertainty-aware data collection.

I. INTRODUCTION AND BACKGROUND
A. Why are intervals valuable for survey responses?

The quantitative elicitation of responses is fundamental in a variety of domains, including perceptual and categorical judgement, attitudinal measurement, and public opinion. Up until now, a broad body of discussion and empirical research has amassed examining the most effective methods of obtaining these responses, along with determining best practice, in terms of both implementation of these methods and interpretation of the resulting data in as accurate and unbiased a manner as is reasonably possible, cf. [1]–[9]. However, while this literature is comprehensive in many respects, we believe that methods of capturing interval-valued data have, up to this point, remained relatively overlooked.

Interval-valued responses provide a powerful and natural means for respondents to communicate any inherent range that may be present in the appropriate response. They also enable the capture and better quantification of variable degrees of uncertainty at an individual-response level, going beyond more conventional between-subject or between-item measures. Here, the position of the interval captures the strength of the response, e.g., level of agreement, whilst the width of the interval captures the uncertainty in the same response (see Fig. 1). This uncertainty may reflect a limitation of the participant’s knowledge, or a lack of specificity in the question, and is often present when a response is made, even if it goes unreported by traditional measures.

Although it is possible to collect information about uncertainty without using intervals, for example by using a traditional ordinal scale paired with further associated questions, adopting the direct capture of intervals avoids the need for additional questions. We argue that an interval-valued response-format can therefore substantially streamline the process of capturing individual response-uncertainty, by reducing the burden on respondents in terms of both length and complexity of surveys. It may also improve the fidelity of vague or uncertain responses, through offering respondents...
a more natural and cohesive method of representing these aspects of a response.

Once an effective paradigm for capturing interval-valued response-data becomes established, this could provide complementary benefits of both a quantitative and qualitative nature. Using interval-valued data can improve the accuracy of model outputs. This is achieved by virtue of retaining potentially valuable information about individual response uncertainty, range, or vagueness, which traditional response-formats would otherwise discard. For example, interval-valued responses can clearly distinguish between ambivalent and uncertain responses, which cannot be similarly discriminated using conventional measures \[12\]–\[16\]. This is particularly useful in application of fuzzy data analysis, \[12\]–\[16\] and Computing with Words \[17\], \[18\], where interval-values are aggregated and modelled through fuzzy sets (see Fig. 2, for example). Additionally, intervals can not only improve precision of model outputs, but also better inform their interpretation, through better quantification of the highly variable degrees of uncertainty that surround them.

**C. What have we done to solve this problem?**

In this paper, we present newly developed, open-source, platform-independent software that enables capture of interval-valued responses. This permits both creation and administration of digital surveys, and can be run both over local networks and online. Distinctively, this will provide researchers with the capability to elicit not only conventional responses, but also interval-valued data, using an ellipse-based response-format. We call this tool the Discrete and Ellipse-based response Capture SYStem (DECSYS). This software is available under a multi-license model, offering a free, open-source license (details online) for academic use, while other licensing options are available on request. It is available to download at: http://www.lucidresearch.org/software.html.

We are currently working to demonstrate and explain the added-value provided by interval-valued survey responses of the type obtained through our ellipse response-format; as well as examining the usability of this response-format, across a variety of domains, through soliciting subjective user feedback.

**D. Why ellipses?**

The ellipse-based response-format leverages the recent proliferation of touchscreen and stylus-compatible devices to obtain interval-valued responses in a quick and intuitive manner, with minimal divergence from how such a response would be made on paper.

Related work has been conducted into developing a ‘Fuzzy Graphic Rating Scale’ (FRS) \[32\] to capture ‘fuzzified’ questionnaire responses \[33\]–\[35\]. By comparison with this, and other possible mechanisms of obtaining interval-valued data, e.g. through soliciting estimates of each endpoint independently, the ellipse-based response-format allows each response

---

\[1\] Here the term interval refers to the technical definition designated by Stevens \[27\]. This is semantically different from how this term is more commonly used throughout the rest of this paper, and indeed most of the sciences. Where referred to elsewhere in this paper, ‘interval’ will refer to the mathematical notion of a closed set of values on a continuous real line, which is defined by two endpoints and possesses the property of width, which is zero only in the case that both endpoints are identical.
to be made in a single, natural and cohesive action. This represents a substantial practical advantage, reducing questionnaire size, and therefore the workload of both the survey respondent and administrator. Further, this method should be easier for respondents to intuitively understand, reducing initial training requirements and engendering higher response fidelity. In addition, the ellipse approach is assumption free, in that it bypasses a fundamental requirement of the FRS to assume the underlying distribution of the membership function representing each response. By contrast, the ellipse method treats all points on the specified interval equally. Such a response can then be combined with others, either within or between subjects, to construct a fuzzy set from multiple responses, the distribution of which will be derived entirely from the input data (cf. Fig. 2).

E. Paper overview.

The remainder of this paper will document the key features of the DECSYS tool. Section II-A will specify technical aspects of the software. Section II-B will present the range of options available when creating and adapting surveys. Section II-C will discuss features relevant to administering surveys, which can be done either locally or online. Section II-D will then address mechanisms by which surveys are accessed by the respondent. Section III will provide a brief summary of the key aspects of the paper, including justification, features and aims of DECSYS.

II. FEATURES AND APPLICATION

The DECSYS survey platform provides experimenters with the capacity to both create and administer surveys according to their particular specifications, in terms of content, general appearance and response-format. The core and unique component of DECSYS, which sets it apart from already available survey tools, is the capacity to effectively and efficiently capture interval-valued responses. However, DECSYS incorporates a range of features, designed to make it a well-rounded and versatile standalone tool for the purposes of both survey design and data-collection. This section will document these features.

A. Technical specification

DECSYS is a web application written in C#. It uses the ASP.NET Core Web Framework and targets .NET Core, and can therefore run cross-platform. In addition, as it is a web application, much of it is written with HTML, CSS and JavaScript. Notably, the question components, enabling Likert-type, ellipse-based and free-text responses (discussed in Section II-B) are written entirely in JavaScript, using the React framework.

B. Creating surveys

An administrator homepage (see Fig. 3) enables new surveys to be created. In addition, this page shows an overview of all previously built surveys, with the option to preview, export, duplicate or delete any existing survey, as well as to edit any survey that has not yet been launched.
The level of flexibility to tailor survey content and formats. This then have the option to either confirm or re-draw the response. The response will be recorded before it is finalised. Respondents because it allows participants to clearly view how any given range markers appear (shown in Fig. 5 in black & Fig. 6(c), be customised. When a participant draws an ellipse response, and colours of both the scale and the drawn response may also be included, to visually divide the scale. The position, width and colour of the text in which the labels are shown. Many aspects of an ellipse component are also configurable. Labels may be added underneath the response-scale, to denote minimum, midpoint and maximum values, with specified text content, font, size and colour. Vertical markers can optionally be included, to visually divide the scale. The position, width and colours of both the scale and the drawn response may also be customised. When a participant draws an ellipse response, range markers appear (shown in Fig. 5 in black & Fig. 6(c), in blue), to indicate the left and right-most bounds of each response, according to where the extremities of the ellipse cross the horizontal response bar. This feature is valuable because it allows participants to clearly view how any given response will be recorded before it is finalised. Respondents then have the option to either confirm or re-draw the response. The broad range of customisable options provide a high-level of flexibility to tailor survey content and formats. This facilitates accurate replication of previous questionnaire studies. Another convenient feature of DECSYS is that existing questions can be duplicated and then adapted. This substantially streamlines the process of maintaining a cohesive format and style throughout the questionnaire, as settings don’t have to be re-specified from defaults for each individual question.

Further notable features include the capability to show an image alongside a question, allowing DECSYS to be used for stimulus-based judgements. In addition, the experimenter can specify whether questions will be presented in a fixed or random order. This feature is designed such that individual questions can be toggled to a fixed or random position (see Figs. 4 & 5). This provides the flexibility to either begin or end a survey with a block of specific questions (e.g. demographic or feedback related), while the remainder of the questionnaire comprises randomly ordered items. Third, whilst editing a component, a small-scale visual preview is shown of how it will appear to the respondent (see Fig. 5). Similarly, it is possible for the experimenter to preview a whole survey before launch. This option allows progression through the entire survey, without recording responses and with the option to exit at any point.

C. Conducting surveys

DECSYS has the functionality to run surveys over local networks, in ‘Workshop Mode’. This allows a high-level of intrinsic data security, through the use of a closed network. It is also designed for use over the internet, in ‘Online Mode’ (presently still in development). This will provide greater flexibility in obtaining participants and broaden the potential user base. The appearance, functions and user interface of the online version are designed to be generally equivalent to the local version, but with the addition of further security measures to control who can access the survey, as either an administrator or participant.

When running in ‘Workshop Mode’, the survey is managed on a host computer and respondents connect to the survey from client computers using a web-browser. The administrator accesses the survey homepage on the host computer, also through a web-browser. All surveys that have been created on the host computer are visible and accessible from the administrator homepage (see Fig. 3), and each survey is indicated here as either currently active or inactive. From this page, any inactive survey can be activated, and an active survey can be closed. In workshop mode, only a single survey can be activated at any one time, as a result of how they are hosted and accessed over local network (see also Section II-D).

The majority of options for survey management (preview, export, duplication or deletion), remain available from the admin page irrespective of its current activity status. However, although any survey can be edited prior to its initial launch, this is not permitted afterwards, even if it is subsequently made inactive again. This feature ensures that all results stored with a given survey pertain to the same questions and format (i.e. data and survey remain consistent), even if these were obtained over multiple sessions. If required, the ‘Duplicate’ function can
be used to copy an entire existing survey following its initial launch. This can then be adapted and saved as an updated version of the same survey.

Once a survey has been made active, the administrator can view the real-time progress of each participant (see Fig. 7). This informs the experimenter about the number of participants that have signed up to take part, allows estimation of time until completion, and also indicates whether any respondents may be having trouble completing the questionnaire.

After launching a survey it is also possible to view results, even as responses come in (see Fig. 8). These are timestamped and linked to a unique identifier for each respondent, and show core response content, e.g. text input, discrete response value, or left and right ellipse endpoints, alongside question number and the order in which each question was presented (important in the case of randomisation). If the survey has been run over multiple sessions, results from each session will be accessible here under different tabs. Results from an individual session can be exported for external processing and analysis (in .json format). Alternatively, results from all sessions can be exported in a single file, together with survey parameters.

D. Accessing and completing surveys

Surveys can be distributed in multiple ways. Responses can be entirely anonymous, in the sense that anyone can sign up to the survey. This offers the broadest potential uptake, but provides minimal guarantees concerning the quality of the data. For instance, the same respondent could potentially complete the survey multiple times, leading to bias and entailing that data is non-independent. Alternatively, responses could be intrinsically linked to a specific individual or account. This approach requires more safeguards regarding considerations of privacy and data security, but offers the option of comparing and cross-validating responses across multiple independent surveys. There is also a middle-ground, by which responses can be made on a semi-anonymous basis, in that only a select group of respondents are invited to take part, and can only do so a single time, but each set of responses remains separate from the identity of the individual that provided them. DECSYS is designed to provide the survey administrator with the capability to collect data according to each of these approaches, depending upon what is most appropriate for their specific use case. Potential respondents will therefore have multiple ways of accessing surveys, dependent upon survey specifications.

In **Workshop Mode**, respondents join a survey via a pre-set local network address (the IP address and port number of the host computer). This is input into the web browser of a client computer that is connected to same local network. Doing so immediately takes the participant to the survey welcome page, from which they can then progress through whichever survey is currently active.

In **Online Mode**, different methods will be used to control access to the survey platform. For survey administrators, this will take the form of an account-based log-in. However, as in other facets of the software, DECSYS is designed to provide a high degree of flexibility with respect to authentication of survey participants. Respondents will be able to access a survey either using a single or limited use authentication token, sent to them via email, through account-based log-in, or, where appropriate, entirely anonymously, by simply entering an openly available URL into a web-browser. This will depend upon the prior specification of the survey administrator.

III. CONCLUSIONS AND FUTURE WORK

We propose that collecting interval-valued survey responses is a useful method of obtaining information about vagueness, uncertainty or fuzziness in participants’ answers, where the interval captures a range of possibly correct responses. Parallel research studies are ongoing, which aim to empirically establish the added-value offered by interval-valued responses, alongside the reported usability and acceptance of our novel ellipse-based response-format by survey respondents. These span a variety of domains, including Cyber-Security, user experience, food rating, personality inventory, and judgement under uncertainty. The present paper aims to address the lack of currently available software that is purposely designed for the efficient and intuitive capture of interval-valued data within surveys. Without such a digital tool, there is a substantial disincentive for the wider adoption of interval-valued survey response-formats, as paper-based versions require increased workload compared to more traditional response-formats.
In this paper we present DECSYS, an open-source and platform-independent software tool that enables the collection of interval-valued responses, as well as traditional free-text and Likert response-formats. DECSYS incorporates a broad range of features, which provide a high level of versatility for the survey administrator, in terms of both function and design, and which also maximise ease-of-use for the respondent. The innovative ellipse-based method used to capture interval-valued data is designed to be both quick and intuitive. This allows interval-valued responses to be provided in a single, cohesive and familiar action, which diverges minimally from how an equivalent paper-based response would be made.

While an initial version of DECSYS has been completed, development of the software is ongoing; look, feel and functionality are continuously being refined. Both provision of additional features and improvements to existing features are planned, including the expansion of security and authentication protocols to provide flexible and secure access control in an online setting. We also aim to facilitate future third-party development, and thereby encourage the wider community to take full advantage of the potential for adaptation of the tool for a broad range of applications. DECSYS is an open-source software project, available at: http://www.lucidresearch.org/software.html.

REFERENCES

[1] J. M. Converse and S. Presser, Survey questions: Handcrafting the standardized questionnaire, vol. 63. Sage, 1986.
[2] F. J. Fowler, Improving survey questions: Design and evaluation, vol. 38. Sage, 1995.
[3] R. M. Groves, F. J. Fowler Jr, M. P. Couper, J. M. Lepkowski, E. Singer, and R. Tourangeau, Survey methodology, vol. 561. John Wiley & Sons, 2011.
[4] J. A. Krosnick and L. R. Fabrigar, “Designing rating scales for effective measurement in surveys,” Survey measurement and process quality, pp. 141–164, 1997.
[5] R. Likert, “A technique for the measurement of attitudes,” Archives of Psychology, 1932.
[6] A. W. Meade and S. B. Craig, “Identifying careless responses in survey data,” Psychological methods, vol. 17, no. 3, p. 437, 2012.
[7] L. Payne Stanley, “The art of asking questions,” 1951.
[8] W. E. Saris and I. N. Gallhofer, Design, evaluation, and analysis of questionnaires for survey research. John Wiley & Sons. 2014.
[9] L. L. Thurstone and E. J. Chave, The measurement of attitude. Univ. of Chicago Press, 1929.
[10] C. H. Coombs and L. C. Coombs, “Don’t know” item ambiguity or respondent uncertainty?, Public Opinion Quarterly, vol. 40, no. 4, pp. 497–514, 1976.
[11] F. J. Klopferr and T. M. Madden, “The middlemost choice on attitude items: Ambivalence, neutrality, or uncertainty?,” Personality and Social Psychology Bulletin, vol. 6, no. 1, pp. 97–101, 1980.
[12] J. Navarro, C. Wagner, U. Ackielin, L. Green, and R. Ashford, “Exploring differences in interpretation of words essential in medical expert-patient communication,” in Fuzzy Systems (FUZZ-IEEE), 2016 IEEE International Conference on, pp. 2157–2164, IEEE, 2016.
[13] A. Pourabrodollah, C. Wagner, S. Miller, K. Wallace, et al., “Towards data-driven environmental planning and policy design-leveraging fuzzy logic to operationalize a planning framework,” in Fuzzy Systems (FUZZ-IEEE), 2014 IEEE International Conference on, pp. 2230–2237, IEEE, 2014.
[14] C. Wagner, M. Smith, K. Wallace, and A. Pourabrodollah, “Generating Uncertain Fuzzy Logic Rules from Surveys: Capturing Subjective Relationships between Variables from Human Experts,” in Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on, pp. 2033–2038, IEEE, 2015.
[15] M. Smith, K. Wallace, L. Lewis, and C. Wagner, “A structured elicitation method to identify key direct risk factors for the management of natural resources,” Helson, vol. 1, no. 3, p. e00043, 2015.
[16] C. Wagner, S. Miller, J. Garibaldi, D. Anderson, and T. Havens, “From Interval-Valued Data to General Type-2 Fuzzy Sets,” Fuzzy Systems, IEEE Transactions on, vol. 23, pp. 248–269, Apr., 2015.
[17] F. Liu and J. M. Mendel, “Encoding words into interval-type 2 fuzzy sets using an interval approach,” IEEE Transactions on Fuzzy Systems, vol. 16, no. 6, pp. 1503–1521, 2008.
[18] D. Wu, J. M. Mendel, and S. Coupland, “Enhanced interval approach for encoding words into interval type-2 fuzzy sets and its convergence analysis,” IEEE Transactions on Fuzzy Systems, vol. 20, no. 3, pp. 499–513, 2012.
[19] S. Miller, C. Wagner, J. M. Garibaldi, and S. Appleby, “Constructing general type-2 fuzzy sets from interval-valued data,” in Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on, pp. 1–8, IEEE, 2012.
[20] C. Wagner, S. Miller, and J. M. Garibaldi, “Similarity based applications for data-driven concept and word models based on type-1 and type-2 fuzzy sets,” in Fuzzy Systems (FUZZ), 2013 IEEE International Conference on, pp. 1–9, IEEE, 2013.
[21] G. González-Rodríguez, A. Colubi, and M. Á. Gil, “Fuzzy data treated as functional data: A one-way anova test approach,” Computational Statistics & Data Analysis, vol. 56, no. 4, pp. 943–955, 2012.
[22] T. Nakama, A. Colubi, and M. A. Lubiano, “Two-way analysis of variance for interval-valued data,” in Combining Soft Computing and Statistical Methods in Data Analysis, pp. 475–482, Springer, 2010.
[23] A. Blanco-Fernández, M. R. Casals, A. Colubi, R. C oppi, N. Corral, S. D. I. R. de Sáa, P. D’Urso, M. B. Ferraro, M. García-Bárzana, M. Á. Gil, et al., “Arithmetic and distance-based approach to the statistical analysis of imprecisely valued data,” in Towards Advanced Data Analysis by Combining Soft Computing and Statistics, pp. 107–118, Springer, 2013.
[24] S. S. Stevens, On the theory of scales of measurement. Bobbs-Merrill, College Division, 1946.
[25] E. P. Ahearn, “The use of visual analog scales in mood disorders: A review,” in Towards Advanced Data Analysis by Combining Soft Computing and Statistics, pp. 19–31, Springer, 2013.
[26] M. A. Lubiano and W. Trutschnig, “Anova for fuzzy random variables using the r-package safd,” in Combining Soft Computing and Statistical Methods in Data Analysis, pp. 449–456. Springer, 2010.
[27] W. Trutschnig, M. A. Lubiano, and J. Lastra, “Safd—an r package for statistical analysis of fuzzy data,” in Towards Advanced Data Analysis by Combining Soft Computing and Statistics, pp. 107–118, Springer, 2013.
[28] C. Wagner, “Generating Uncertain Fuzzy Logic Rules from Surveys: Capturing Subjective Relationships between Variables from Human Experts,” in Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on, pp. 2033–2038, IEEE, 2015.