Evaluating Free Rides in Visualizations of Sets

Andrew Blake, Gem Stapleton, Peter Rodgers, and Anestis Touloumis

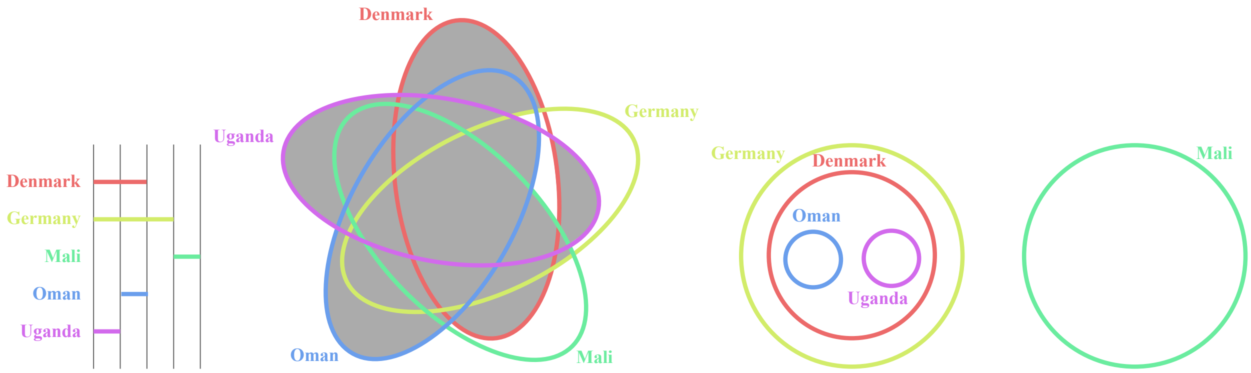

Fig. 1. Diagrams obtained from translating textual statements: (left to right) linear, Venn, and Euler.

Abstract—A consideration for any visualization technique is the added value it brings over other representations of data. A *free ride*, a concept introduced by Shimojima, occurs in a visualization when it reveals some fact that must be inferred from an alternative representation from which the visualization was derived. This paper presents preliminary results from an empirical evaluation of free rides occurring in visualizations of sets as compared to textual representations. By focusing on Euler, linear and Venn diagrams, our results suggest that more research is needed to establish when free rides are cognitively beneficial.

Index Terms—linear diagrams, Venn diagrams, Euler diagrams, free rides, observational advantages

1 Introduction

The purpose of this paper is to present preliminary results from an empirical study that tests the belief that free rides [10], recently generalized to observational advantages [12], are features of visual modes of communication that aid cognition. Free rides occur when one representation of information is translated into another and the resulting representation makes explicit some facts that must be derived (inferred) from the original. Such explicit facts are precisely the free rides. We consider an original representation in textual form which is then translated into diagrammatic form. The specific research question we address can thus be made more precise: does using text *alongside* a semantically equivalent diagram lead to significant performance benefits over *just* using text when identifying information that is conveyed by free rides? We focus on information about sets, visualized by linear diagrams, Venn diagrams, and Euler diagrams as seen in Fig 1.

We now give a simple example. Suppose we have information about people who have visited various countries:

- Everyone who visited Denmark visited Germany
- No one visited both Germany and Mali
- Everyone who visited Oman visited Denmark
- Everyone who visited Uganda visited Denmark
- No one visited Uganda and Oman.

From these statements, which correspond to subset (*Everyone...*) and disjointness (*No one...*) relations between sets, various facts can be derived. These derivations include *Everyone who visited Uganda visited Germany* and *No one visited both Mali and Uganda.* By translating the originally given facts into the Euler diagram in Fig. 1, we make these two derived facts explicit and, so, they are examples of free rides from the diagram. This is because, in the first case, the translation necessarily places the Uganda circle inside the Germany circle. In the second case, the Mali and Uganda circles necessarily do not overlap. Indeed, this Euler diagram also makes *additional* derivable facts explicit, such as *No one visited both Mali and Oman*, and therefore has many free rides. Focusing our evaluation of free rides on visualizations of sets is of particular significance because there are enormous amounts of set-based data available in a wide variety of application areas [2]. Reflecting this abundance of data, the research community is actively devising methods for visualizing it. Set visualization techniques often exploit closed curves (or variations thereof) [4, 5, 7, 8, 11] or lines for representing sets [1, 3, 6, 9]. This paper therefore focuses on such methods by evaluating Venn diagrams and Euler diagrams both of which use closed curves [13–15], see the middle and right of Fig. 1, as well as linear diagrams (which use line segments) [9], see the left of Fig. 1.

2 Summary of the Study Design

We generated 20 tasks for people to perform where they had to choose the correct answer from one of five multiple choices. Ten tasks required them to identify a subset free ride (Everyone ...) and the other ten focused on a disjointness free ride (No one ...). Of the five options, one was always ‘none of the above’, two were subset statement and two were disjontness statements; see Fig. 2. The correct answers were distributed across the option positions and the other three subset and disjointness statements were randomly positioned. The study included four groups, thus adopting a between group design: text only (T), linear diagrams with text (L& T), Venn diagrams with text (V& T) and Euler diagrams with text (E& T). We captured accuracy and time performance data for each task. The study was conducted online using Prolific Academic and participants were randomly assigned to one of the groups. The study started with a short training phase, exposing the participants to four tasks of increasing difficulty. The participants then proceeded to answer the 20 questions from which performance data were gathered. This performance phase also included a few questions

- *Blake, Stapleton and Touloumis are with the University of Brighton.*
- *Rodgers is with the University of Kent.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

Table 1. Summary of the main study data.

Group	No. of Participants	Accuracy	Mean Time
Overall	404	70.21%	37.63
T	99	51.11%	47.63
L&T	99	87.58%	29.47
V&T	103	51.07%	43.33
E&T	103	91.02%	30.17

Table 2. Overall Comparison of Treatments by Accuracy.

Treatments	Odds	CI	p-value	Sig.	Most Accurate
L&T versus T	6.74	$(4.53, 10.04)$	< 0.0001	✓	L&T
L&T versus V&T	6.75	$(4.37, 10.44)$	< 0.0001	✓	L&T
L&T versus E&T	0.70	$(0.42, 1.15)$	0.1537		
T versus V&T	1.00	$(0.75, 1.33)$	0.9906		
T versus E&T	0.10	$(0.07, 0.15)$	< 0.0001	✓	E&T
V&T versus E&T	0.10	$(0.07, 0.16)$	< 0.0001	✓	E&T

designed to identify inattentive participants.

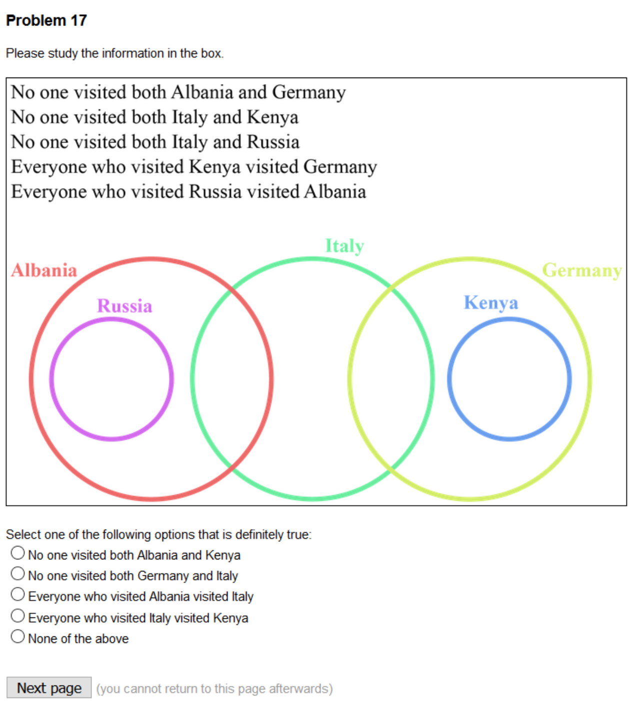

Fig. 2. Euler diagram question: the correct answer is option 1.

3 Results: Comparison of Representations

Here we report on the overall comparison between the four treatment groups. The accuracy rates and mean times (in seconds) are summarised in table 1. Whilst the accuracy rates and mean times are in *indicator* of relative performance, it is important to note that the statistical methods employed do not compare these data: methods that compare means (e.g. ANOVA) do not account for correlated responses from participants and make other assumptions that our data violate.

Using a GEE based statistical model for the accuracy data, we estimated a 95% confidence interval (CI) for the odds of providing a correct answer with one treatment compared to another. The estimated odds of correctly answering questions with L&T was 6.74 (to 2d.p.) times higher than that of T with a 95% CI of $(4.53, 10.04)$ and p-value of < 0.0001 (to 4d.p.). Therefore, L&T supported significantly more accurate task performance than T only. Results for the other pairwise comparisons are given in table 2, from which we can derive an overall accuracy ranking for the treatments: *L&T* $=$ *E&T* $>$ *V&T*$=$ *T*.

Table 3. Overall Comparison of Treatments by Time.

Treatments	Ratio	CI	p-value	Sig.	Fastest
L&T versus T	0.62	$(0.55, 0.69)$	< 0.0001	✓	L&T
L&T versus V&T	0.63	$(0.56, 0.70)$	< 0.0001	✓	L&T
L&T versus E&T	0.95	$(0.85, 1.06)$	0.3492		
T versus V&T	1.01	$(0.89, 1.15)$	0.8373		
T versus E&T	1.54	$(1.36, 1.74)$	< 0.0001	✓	E&T
V&T versus E&T	1.52	$(1.34, 1.71)$	< 0.0001	✓	E&T

Using a GEE based statistical model for the time data, we estimated a 95% CI for the ratio of the time (measured in seconds) needed to answer a question correctly with one treatment compared to another. The CI and its corresponding p-value allowed us to determine whether two treatments were significantly different. The model estimated that the time needed to answer a question correctly with L&T was 0.62 times (2d.p.) that with T with a 95% CI of $(0.55, 0.69)$ and p-value of < 0.0001. Therefore, linear diagrams with text supported significantly faster task performance than text only. Results for the other pairwise comparisons are given in table 3, from which we can derive an overall time ranking for the treatments: *L&T* $=$ *E&T* $>$ *V&T*$=$ *T*.

Therefore, our accuracy and time analysis support the superiority of linear diagrams and Euler diagrams, when used as a support for a textual representation, as compared to Venn diagrams alongside text or just text alone. Taking into account both the accuracy and time analysis, we have consistent rankings, from which we derive an overall ranking of the four treatments: ***L&T*** $=$ ***E&T*** $>$ ***V&T***$=$ ***T***.

The odds and ratios computed for the accuracy and, respectively, time data give insight into the effect size. For instance, the odds of producing a correct answer using linear diagrams alongside text compared to text alone are 6.74 (odds of approximately 1 would indicate no significant difference[1]). From the perspective of time, we would expect correctly answering a question using linear diagrams alongside text to be 0.62 of the time taken (i.e. to take 62% of the time) to provide a correct answer using text alone. The other effect sizes, where we saw significant differences, are similar and are evident from the odds and ratios given in the tables.

4 Conclusion

This study has revealed that the role of free rides in explaining the cognitive benefits of diagrams is not clear-cut. Whilst Euler and linear diagrams were both effective supports for the textual representations, Venn diagrams did not lead to significantly better task performance. We can suggest that just because a diagram expresses other, consequential information, it does not *necessarily* mean that the diagram facilitates inference and saves deductive cost.

Our results show that, whilst free rides in effective diagrams can lead to performance improvements, it is also important to use well-designed diagrams suitable for the task at hand. One cannot just rely on the presence of free rides as an argument for the efficacy of a representation of information. In some cases, such as Venn diagrams, there are clearly other factors that are important to understand in order to determine the potential benefits or negative consequences of alternative representations. Our results point to the need to better understand the role of free rides and other features of representations in solving inference problems. We suggest that a more comprehensive and integrated theory of features of diagrams that make them effective for cognition, including a deeper understanding of free rides, should be a major goal of the information visualization community. Future work should also consider identifying the role of text, in the treatments that included diagrams, when participants performed the tasks.

Acknowledgments

This research was partially funded by a Leverhulme Trust Research Project Grant (RPG-2016-082) for the project entitled Accessible Reasoning with Diagrams.

[1]Strictly speaking, a confidence interval containing 1 would indicate no significant difference.

References

[1] B. Alper, N. Henry Riche, G. Ramos, and M. Czerwinski. Design study of LineSets, a novel set visualization technique. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2259–2267, 2011.

[2] B. Alsallakh, L. Micallef, W. Aigner, H. Hauser, S. Miksch, and P. Rodgers. Visualizing sets and set-typed data: State-of-the-art and future challenges. In *Eurographics Conference on Visualization (EuroVis)*, pp. 124–138. 2014.

[3] P. Cheng. Probably good diagrams for learning: Representational epistemic recodification of probability theory. *Topics in Cognitive Science*, 3:475–496, 2011.

[4] S. Chow and F. Ruskey. Towards a general solution to drawing area-proportional Euler diagrams. In *Euler Diagrams 2004*, vol. 134 of *ENTCS*, pp. 3–18. ENTCS, 2005.

[5] C. Collins, G. Penn, and M. S. T. Carpendale. Bubble sets: Revealing set relations with isocontours over existing visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1009–1016, 2009.

[6] B. Gottfried. A comparative study of linear and region based diagrams. *Journal of Spatial Information Science*, 2015(10):3–20, 2015.

[7] W. Meulemans, N. Henry Riche, B. Speckmann, B. Alper, and T. Dwyer. Kelpfusion: A hybrid set visualization technique. *IEEE Transactions on Visualization and Computer Graphics*, 19(11):1846–1858, 2013.

[8] N. Riche and T. Dwyer. Untangling Euler diagrams. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1090–1099, 2010.

[9] P. Rodgers, G. Stapleton, and P. Chapman. Visualizing sets with linear diagrams. *ACM Transactions on Computer-Human Interaction*, 22(6):article 27 (28 pages), 2015.

[10] A. Shimojima. *Semantic Properties of Diagrams and Their Cognitive Potentials*. CSLI Publications, Stanford, CA, USA, 2015.

[11] P. Simonetto, D. Auber, and D. Archambault. Fully automatic visualisation of overlapping sets. *Computer Graphics Forum*, 28(3):967–974, 2009.

[12] G. Stapleton, M. Jamnik, and A. Shimojima. What makes an effective representation of information: A formal account of observational advantages. *Journal of Logic, Language and Information*, 26(2):143–177, 2017.

[13] G. Stapleton, L. Zhang, J. Howse, and P. Rodgers. Drawing Euler diagrams with circles: The theory of piercings. *IEEE Transactions on Visualization and Computer Graphics*, 17(7):1020–1032, 2011.

[14] J. Venn. On the diagrammatic and mechanical representation of propositions and reasonings. *Phil.Mag*, 1880.

[15] L. Wilkinson. Exact and approximate area-proportional circular Venn and Euler diagrams. *IEEE Transactions on Visualization and Computer Graphics*, 18(2):321–331, 2012.