Exploring Human Factors in Formal Diagram Usage

Andrew Fish[1,*], Babak Khazaei[2], and Chris Roast[2]

[1] University of Brighton, School of Computing,
Mathematical and Information Sciences, UK
Andrew.fish@brighton.ac.uk
[2] Sheffield Hallam University,
Culture, Communication and Computing Research Institute, UK
B.Khazaei@shu.ac.uk, c.r.roast@shu.ac.uk

Abstract. Formal diagrammatic notations have been developed as alternatives to symbolic specification notations. Ostensibly to aid users in performing comprehension and reasoning tasks, restrictions called wellformedness conditions may be imposed. However, imposing too many of these conditions can have adverse effects on the utility of the notation (e.g. reducing the expressiveness). Understanding the human factors involved in the use of a notation, such as how user-preference and comprehension relate to the imposition of wellformedness conditions, will enable the notation designers to make more informed design decisions. Euler diagrams are a simple visualization of set-theoretic relationships which are the basis of more expressive constraint languages. We have performed exploratory studies with Euler diagrams which indicated that novice user preferences strongly conform to the imposition of all wellformedness conditions, but that even a limited exposure diminishes this preference.

1 Introduction

Formal notations have been advocated as important within a variety of software development contexts, since they can offer clarity and precision; the provision of sophisticated tool support can strengthen confidence in the development processes and the quality of the end product. However, the role that such a notation plays is that of a representation that has to be composed, comprehended and updated as part of the development process. Hence, although the formality is a valued facet, there are other significant factors that affect their value. For example, notation appropriateness [1,2] can influence the quality of solutions that a user may entertain. Additional factors can come into play when we consider the role of environments which can affect the particular form in which information is expressed. Users working in different environments may have added difficulties when sharing specifications if their environments enforce the use of different forms of expression for example. The role of a representation is also important, since any specification is likely to be influenced by its primary use (to communicate to others or to record information for instance).

Software developers and design teams who have to work with formal notations are end-users. When designing a formal notation, features that support its intended use

---

and uptake are often provided, but features that may limit its effective use can often be accidentally included. For example, the provision of "macros" or "libraries" for a base notation are features that are useful in supporting the user, whereas a lack of symbol discriminability and limited spatial layout can increase the difficulty in using notations [3]. Difficulties in encouraging user-uptake of symbolic formal specification notations, such as Z, is one of the reasons for the development of diagrammatic specification notations, such as constraint diagrams [4,5]. Also, the Unified Modeling Language is now commonly used in the software development cycle and since the only non-diagrammatic component is the Object Constraint Language (which can be used to express system invariants and pre/post-condition contracts for operations for example), a suitable diagrammatic alternative would fit in with the diagrammatic paradigm. It could also potentially widen the scope of usage of constraints by making them more accessible than their symbolic counterparts, and this may improve readers' understanding of formal specification documentation for instance.

When defining a visual specification notation, the presentation features are important for effective use: they can assist a user in easily identifying syntactic structures and facilitate the interpretation of semantic characteristics. Our perceptual competence at recognizing bounded areas and arcs in diagrams suggests that often less mental effort needs to be devoted to identifying syntactic structures in a diagrammatic representation than in a symbolic one. For a diagrammatic notation to be effective it is important that the relationships in the representation are well-matched with the domain characteristics [6], then the spatial relationships of the representation can lead to free rides [7], which are inferences gained for free due to the well-matching. Despite this potential, nothing limits visual specification notations from being used in a manner that does not exploit these benefits. Hence, it is of interest to examine approaches to ensuring the valuable use of visual notations and the subjective assessments that may also play a role in effective visual notation use.

Restrictions of the presentation of information (such as not allowing three lines to meet at a single point) are called wellformedness conditions. Such conditions are usually imposed with the intention of making the diagrams easier to comprehend and reason with. However, often there has been little or no user testing to determine the actual effects on users of these restrictions. Also, there can be many possible choices of condition and enforcing them can have side-effects such as reducing the expressiveness of the system, or of making it more difficult to present certain statements. Thus a balance needs to be struck between the imposition of some of these conditions and the utility of the system (e.g. being able to visualize as much as possible). Discovering the effects of wellformedness conditions on user preference and performance will help notation designers to determine the correct balance of conditions to be relaxed or enforced for different user groups.

Euler diagrams are a diagrammatic method for representing information about the relationships between sets. They have been used in various forms since Euler [8] first introduced them, and they generalize Venn diagrams [9] which represent all set intersections. Euler diagrams, and notations based on them, are currently being used in many application areas for the presentation of information, including: to represent non-hierarchical directories [10,11]; to visualize complex genetic set relations [12]; to represent ontologies in semantic web applications [13]; to enable the visualization of statistical data [14]. However, the main focus of usage of such diagrams is to model

object oriented software systems [5, 15,16] and to develop formal, diagrammatic logical reasoning systems and automatic theorem provers [4, 17,18-22]. Recently, the consequences of enforcing certain wellformedness conditions on Euler diagrams in the areas of drawability, semantic interpretation and reasoning have been investigated [23], and we wish to acquire further related information about user-preference.

In this paper, we describe exploratory studies which primarily set out to explore the relationship between the use of wellformedness conditions for Euler diagrams and user-preference. In addition, we examine how user comprehension is linked to these wellformedness conditions. The examination of preference is aimed at helping us develop an understanding of whether wellformedness is seen as desirable by users of diagram as well as by those designers who espouse the wellformedness conditions. Related long term goals include understanding how a user's preference may change with experience, how the context of use affects preference and how preference is linked to comprehension. The current exploration provides evidence about the importance of the choice of wellformedness conditions for potential users.
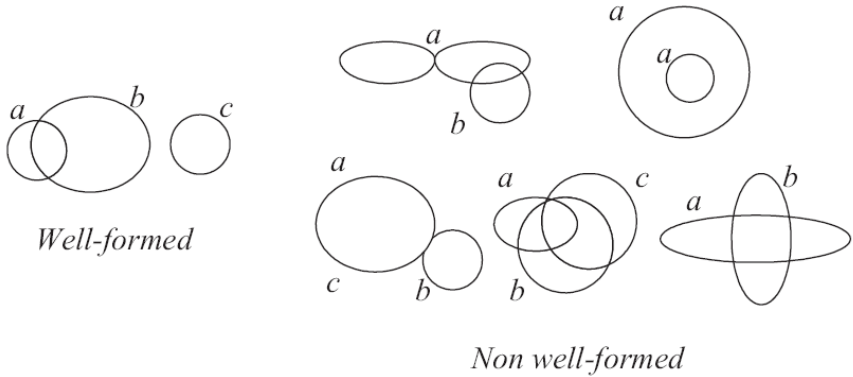
2 Euler Diagrams

We give an informal definition of Euler diagrams and their wellformedness conditions (see [24] for more details). An *Euler diagram* is a finite set of labelled closed curves (called *contours*) in the plane. A *zone* (or *minimal region*) is a connected component of the complement of the contour set, and a *region* is a union of zones. An Euler diagram is *well-formed* if it satisfies a given set of wellformedness conditions. A typical set of *wellformedness conditions* are:

1. **Simple contours:** The contours are simple closed curves.
2. **Unique contour labels:** Each contour has a unique label.
3. **No concurrency:** Contours are not concurrent (that is, they meet at a finite, discrete set of points).
4. **No tangential intersections:** Contours do not touch, but can cross each other transversely wherever they meet.
5. **No multiple points:** No more than two contours meet at any single point.
6. **Unique zone labels:** Each zone can be uniquely identified by the set of contours containing it and the set of contours excluding it.

For example, a well-formed Euler diagram is shown on the left of Figure 1. It has three labels a,b,c, three corresponding contours, and five zones determined by the label sets $\{\}$, $\{a\}$, $\{a,b\}$, $\{b\}$ and $\{c\}$ – corresponding to the regions outside all contours, inside just a, inside both a and b, inside just b, and inside just c, respectively.

Figure 1 also shows five non well-formed diagrams. The top left one shows a "figure of eight" curve (which is a non-simple curve) for the contour labelled by a and so this fails condition 1. The top right diagram depicts two contours both of which are labelled by a and so this fails condition 2. The bottom left diagram fails condition 3 because it has two contours (a and c) which are completely concurrent, that is, one is overlayed on top of the other; it also has two contours (a and b) which meet tangentially and so it fails condition 4. The bottom middle diagram has a triple point (which is a multiple point) – all three contours pass through a single point – and so it fails

Fig. 1. Wellformedness of Euler diagrams

condition 5. The bottom right diagram has zones which are not uniquely identifiable using the contour labels – the region which is inside *a* but outside *b* is disconnected – and so it fails condition 6.

Conditions 1, 2 and 4 are almost always enforced, with simplicity only previously being relaxed to make reasoning easier [25]. Conditions 3 and 5 are sometimes enforced [20] and sometimes not [26]. Condition 6 is usually enforced, and has often been referred to as "no split /disconnected zones" [27].

2.1 Semantics

The semantics of Euler diagrams that we adopt are:

1. the interior of each contour represents the set denoted by its label, and each region of the diagram represents the corresponding set intersection determined by the labels.
2. a shaded or missing region of the diagrams represents an empty set, whilst a non-shaded region that is present in the diagram represents a non-empty set.

Figure 2 shows two semantically equivalent diagrams, each depicting three non-empty sets (*A, B* and *C*) such that $A \cap C = \emptyset$ (that is, *A* and *C* are disjoint), $B \cap C = \emptyset$ and $A \cap B \neq \emptyset$. The second utilizes shaded zones to indicate emptiness, whereas the first used the absence of zones for this. The introduction of shading into the system enables more varied forms of expression, and whilst it has the benefit of explicitly depicting emptiness (which can only be depicted by omission without the use of shading) it also brings with it greater diagrammatic complexity arising from more overlapping of the contours, as well as the need to understand more syntax.

There are many slight variations in diagram semantics in the literature (as well as extensions). The most notable variation here is that often dots or crosses (called spiders, constant sequences or x-sequences for example) are required to be placed in regions to represent non-emptiness [20 – 22], but we wished to burden the subjects with
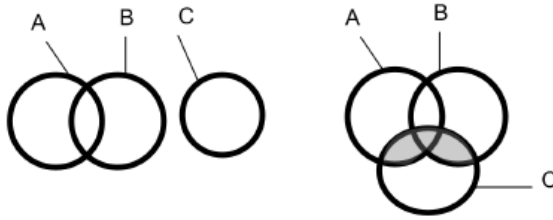
Fig. 2. Semantically equivalent diagrams

as little additional syntax as possible. An avenue that warrants further study is the testing of different choices of semantics for Euler diagrams – how these affect user preference and understanding, especially in the presence of various wellformedness conditions.

2.2   Roles of Euler Diagrams

Euler diagrams are thought to be an effective representation since the set-theoretic re-lationships that they represent are well-matched by the spatial relationships that they use. For example, the proper subset relationship is well-matched to the proper inclu-sion of curves in the plane (both are transitive, but not reflexive or symmetric). We can obtain the inference "$A \subset C$" from "$A \subset B$" and "$A \subset C$"for free from the corre-sponding Euler diagram with three concentric circles shown in Figure 3, for instance.
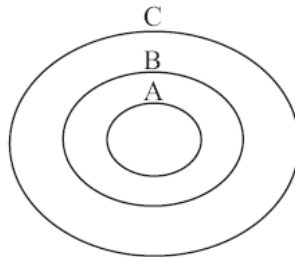

Fig. 3. Well-matching and free rides

   An example demonstrating the importance of the role of notation and environment for Euler diagrams occurs in interactive Euler diagram theorem proving environ-ments. Information is stored and reasoned with at an abstract level, which is useful for computations, but is not so appropriate for presentation to a user (to enhance user faith in an automated proof, or as an aid to understanding proof techniques). Imposing too many wellformedness conditions can prevent certain set theoretic statements from being represented diagrammatically, and a change of representation here is likely to be undesirable for a user. Furthermore, one may wish to tailor the presentation of a proof to individual user preference. For example, an environment could offer a choice between a short proof containing non wellformed diagrams and a longer proof using wellformed diagrams. In general there is a balance to be found between the number of diagrams in a proof to be displayed and both the size of the deduction steps used,

and the complexity of the diagrams used. An experienced user might prefer to view a non wellformed diagram which is compact but contains a lot of information over a collection of wellformed diagrams which are individually easier to comprehend but one also has the added cognitive load of having information spread across more diagrams.

It is also important to remember that diagrams may be authored, read and edited by different people (possibly in different groups, in different countries, using different environments). Therefore, being too rigid in the enforcement of wellformedness conditions may have detrimental effects on communication.

3   Preliminary Study

Before conducting a fully fledged study examining users perceptions of, and competence with, Euler diagrams a preliminary study was conducted to get some feedback on the likely outcomes of the main study. The pilot study was conducted with five subjects who were all half-way through a second year degree option on human-computer interaction.

Introducing the concept of Euler diagrams to subjects who may be unfamiliar with discrete maths concepts presented the problem that they may easily view diagram comprehension tasks as being assessments of their ability. Because of this we wished to focus upon non-abstract examples of Euler diagrams (for example, not using alphabetical labelling of the contours), while also ensuring that any examples would not encourage subjects to guess at answers based upon personal knowledge or expectation. To this end we developed a contextual setting in which Euler diagrams were proposed as a form of graphical output to an internet search facility - termed "Oigle"! Within this setting, subjects could be easily encouraged to focus upon judging the value and utility of the diagrams presented as output. The labels used for the concrete examples were motivated by lists of popular internet search terms. As far as possible, labels which were considered to have strongly related meanings were not used together in the same diagram. We also adopted a slightly different labelling convention than usual (compare Figures 1 and 2), with the aim of reducing potential ambiguity caused by the placement of a label.

The preliminary study took around 40 minutes, with time equally divided between: familiarization and training, and comprehension questions. The familiarization and training involved a short introduction to the "Oigle" concept and some basic examples of its output. Subjects were given four well-formed diagrams involving no more than four contours, asked to briefly describe them, and given the chance to compare their answers with model descriptions.

After the familiarization phase, subjects answered twenty "yes/no" comprehension questions. The questions were related to a sheet of nine diagrams, four of which were not well-formed; each of these diagrams involved no more than four contours. The results showed an average score of 16/20 indicating a good level of comprehension. The greatest variability in the answers was found for those questions concerning non well-formed diagrams, indicating that wellformedness within the diagrams used was influential. Additional feedback from the subjects indicated a preference for "avoiding unnecessary area divisions", "providing a neater layout of diagrams with more symmetry" and "clearer separation where regions were separate". Within the confines of

the study, no clarifications of the descriptions were offered. Subsequent follow-up studies will involve more focused interviews with the subjects who volunteered these explanations.

This initial study provided validation for the experimental setting, timing, and the potential to explore the influence of wellformedness conditions on comprehending the Euler diagram notation, especially when working with novice subjects. From this study we concluded that the questions could be more difficult, and we could include more complex diagrams. Therefore, a slightly more complex Euler diagram convention was chosen for the main study: we chose to employ the concept of a shaded region to indicate an empty set, as illustrated in Figure 2. Introducing shaded regions into the notation allows the same information to be represented by a greater variety of diagrams, and this provides a useful way of adding to the complexity of the experimental materials.

The preliminary study highlighted the variety of concrete layouts that exist for Euler diagrams, whether well-formed or not. For example, two overlapping contours can be drawn varying the relative sizes of the two contours and their relative position, and of course, their individual shapes. For our study we wish to limit the unwarranted impact that this variety may have, and focus specifically on the wellformedness conditions. For this reason we selected some "scoping" heuristics designed to ensure a level of conformance in the style and layout of diagrams, thereby enabling wellformedness conditions to be assessed more accurately:

1. Keep regions of a similar consistent size (except in purposefully ambiguous cases). Hence, non-trivial overlaps should be shown clearly as such, but if a tangential intersection is to be displayed then the presence (or absence) of an overlap need not be clearly shown.
2. Do not stretch contours unnecessarily. Hence contours do not become distorted unnecessarily.
3. The bounding rectangle of a diagram (where this means a rectangle containing all of the contours) should be close to square.
4. All labels should be outside the bounding rectangle of the diagram, whilst being closest to the contours that they label; they should appear alphabetically in a clockwise order.

These heuristics were proposed as "good practice" that should be followed where possible.

4 Main Study

The main study was directed towards establishing an understanding of how users react or respond to the visual language, and how this relates to issues such as comprehension, wellformedness and less precise, though still significant, concepts such as visual appeal. Although user preferences for specific diagrams may be highly subjective, it is valuable to know how closely their preference follows the notion of wellformedness and also how influential it might be upon comprehension and thus, utility. It is quite possible, especially with novice users, that preferences can influence comprehension, both in terms of accuracy and also willingness to engage with diagram related problems.

4.1   Experimental Design

Based on previous studies requiring subjective responses [1,28], there is evidence that experience can be an influential factor, and so we chose to gather user preference data both before and after using Euler diagrams. A set of comprehension questions about Euler diagrams similar to those of the preliminary study served as a (limited) Euler diagram experience. We employed a form of subjective preferences reporting that allows preferences to be easily identified. It is not uncommon in some experimental settings for subjects to proffer responses that they believe to be those desired by the study. This effect can be limited by providing subjects with a comparative judgment task. In this case we asked subjects to indicate their most preferred and least preferred diagrams within given sets.

Although motivated by specific concerns about wellformedness, the study was primarily exploratory, focused upon revealing factors that may be relevant for further studies. If we were to posit hypotheses driving the study these would be:

1. Wellformedness conditions concur with user comprehension and user preferences.
2. Experience with Euler diagrams influences user preferences.

The study consisted of three phases, an *a priori* preferences assessment, a comprehension phase and a *post priori* preference assessment:

1. Subjects were presented with four questions showing groups of three similar diagrams and were required to indicate which they preferred the most and which they preferred the least. Each question had at least one well-formed diagram, and prior to the study an expert assessment of the quality of the diagrams was also recorded. Figure 4 shows an example of one of these questions.
2. Two relatively complex Euler diagrams were provided and ten "yes/no" comprehension questions given. The questions were balanced both between positive and negative answers, and between the two diagrams. The responses to two of the questions were contingent upon whether tangential intersections created a nontrivial region, so that the set intersection was non-empty (a *liberal* reading), or they did not create a non-trivial region, so that the set intersection was empty (a *conservative* reading). Figure 5 shows the two diagrams together with some of the questions used.
3. Subjects were presented with four further preference questions in the style of the first phase. Figure 6 shows an example of one of these questions.

In order to gather information on subjects' preferences, in both phases 1 and 3 subjects were given no indication of criteria by which to judge preference other than their own — they were simply asked to indicate which diagram they thought was "best" and which they thought was "worst". Subjects were given the opportunity to report back on any reasoning or rationale that they used for each phase. The diagrams employed for each question in phases 1 and 3 were similar in complexity, although not all of the diagrams within each question were semantically equivalent since this could imply a received interpretation of the inherently ambiguous non-wellformed cases. One of the primary purposes of Phase 2 was to provide the end users with some experience of comprehension, and so it was felt that natural language questions would suffice (as opposed to more formal questions posed in symbolic logic for instance).

The experimental preparation involved an initial phase of familiarization and training taking 25 minutes. The students were then asked to participate in a half an hour experiment involving the above phases. It was explained that this was a comprehension exercise indirectly relevant to their course and that the purpose was not to assess them but to help us to understand both their preferences and their difficulties in comprehending some diagrams. Twenty five second year B.Sc. Software Engineering students took part in the study. The students were covering elements of system analysis and design based on the Object Constraint Language (OCL) and the Unified Modelling language (UML). The students had some mathematical background - the minimum being the pass level in GCSE Maths. The nature of the course and the focus of the experiment had some similarities as the students were covering the range of diagrammatic notations within the UML. All of the students had successfully completed one and a half academic years of programming and software engineering study.

4.2 Results

All of the participants completed the study without any expressed difficulty. About half of the subjects provided feedback on the reason for their choices, and a few subjects did not identify both "best" and "worst" preferences in some questions. The results for each phase are explained below and are summarized in Tables 1, 2 and 3. There is a strong conformance between subjects, suggesting that despite the subjective nature of phases 1 and 3, there is a considerable amount of agreement between subjects. For phases 1 and 3, Table 1 shows the total number of responses (over all subjects and all questions in the phase) of "best" (and "worst") actually being a well-formed diagram. For example, in phase 1 there were 79 user choices of "best" which were well-formed, but only 21 user choices of "best" which were not well-formed. Both the a priori and post priori responses show a strong correlation with the proposition that novice subjects' preferences match well-formed Euler diagrams. Comparing results for phases 1 and 3 in Table 1, we also see that experience with Euler diagrams appears to lessen the conformance. The significance of this relationship was assessed by comparing the average score for each subject (where the score for a subject is the number of choices of "best" that are wellformed) with the probable score for no effect (see Table 2). The probable score for each question was given by the proportion of well-formed diagrams available in the question. The results for both the a priori case and the post priori case are highly significant when compared with their predicted averages ($p < 0.001$ with Wilcoxon matched-pairs signed-ranks test, $N = 25$).
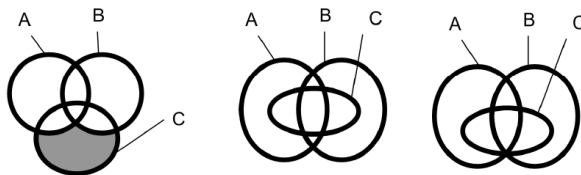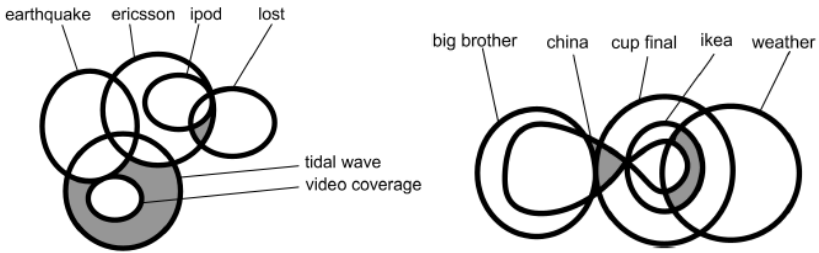

Fig. 4. Phase 1: choose the best and worst

There are some sites with results about all three of "ipod", "ericsson" and "lost".	[YES/NO]
There are some sites matching "earth-quake" and "tidal wave".	[YES/NO]
Every site about "ipod" is also about either "ericsson" or "tidal wave".	[YES/NO]
Pages about "cup final" and "weather" include some "ikea" pages.	[YES/NO]
There are no pages about "big brother" and "weather".	[YES/NO]

Fig. 5. Phase 2: comprehension
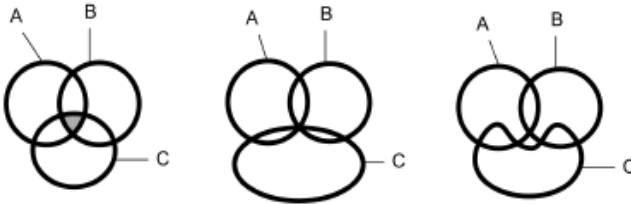
Fig. 6. Phase 3: choose the best and worst

Table 1. Results: matching preference to wellformdness in phases 1 and 3

well-formed	Phase 1		Phase 3	
	"best"	"worst"	"best"	"worst"
yes	79	12	68	35
no	21	81	31	57

Table 2. Results: comparing average score with expected average in phases 1 and 3

	Phase 1	Phase 3
average score ($N = 25$)	3.20	2.75
probable score	1.62	2.00

Subjects were also asked to provide feedback on the reasons for their preferences. From these the overarching theme was a preference for "clarity" and "readability". Some subjects explicitly stated that their choice was based on visual appeal, occasionally with an explicit criterion such as having "symmetry". Others alluded to more semantic concepts such as "set theory" and "equivalence", with one subject explicitly referring to the problematic nature of tangential intersections. Some interesting cases focused more on comparative judgments, such as the degree of unnecessary complexity and the number of shaded regions; one subject even suggested that the textual equivalent of a diagram would be simpler to understand. Generally, we found a rich mixture of factors being employed ranging from visual and aesthetic concerns through to semantic clarity. This mix of informal comments combined with the high correlation of results suggests that, for the cases examined, wellformedness promotes visual representations that users can interpret as being both visually clear and also easy to interpret logically.

The comprehension task (phase 2) illustrates that the subjects, on the whole, have a reasonable grasp of Euler diagram semantics, especially as the diagrams used were considerably more complex than those of the preliminary study (see Figure 5). Table 3 shows the results of this task, with the "average" column indicating the average number of correct answers. An interesting feature is that diagram 2 (the bottom diagram in Figure 5) has a significantly lower level of accurate comprehension than diagram 1 (the top diagram in Figure 5). Diagram 2 involves more instances of non wellformedness than diagram 1. The difference between the diagrams was also apparent in the subjects' feedback, with several subjects referring to the complexity of the second diagram. Phase 2 also showed that in a minority of cases (2/7) tangential intersections were interpreted "liberally" (i.e. as being a non-empty intersection). Perhaps the most interesting observation here is that five of subjects (20%) were not consistent in their interpretation of the two diagrams. Hence, these subjects altered their interpretation based on the diagram encountered - an effect that could be attributed to the complexity of diagram 2 or the informal interpretation of the labels used in the diagram.

Table 3. Results from phase 2 (averages and interpretation of tangential intersections)

	Average		*liberal* reading
Diagram 1	3.92	(78%)	2
Diagram 2	2.64	(53%)	5
Total	6.56	(66%)	7

Our study was also interested in the influence of experience with Euler diagrams in phase 2 on subjects' preferences. Although phase 1 and phase 3 were not formally balanced in the experimental design, the conformance of subjects preferences is reduced for the post priori case (phase 3). The difference between the a prior and post priori can be confirmed statistically: the proportionate score in the two cases were compared using Wilcoxon and showed a significance of $p < 0.05$. The potential influences behind this effect are the subjects' experience with phase 2, and possible differences in complexity between phase 1 and phase 3. In order to exclude the second of

these factors, the variances of individual questions within phase 1 and phase 3, as well as the likely complexity of the questions was examined. In addition, the question complexity (or, more precisely, the complexity of the diagrams used in the questions) was assessed using a variety of metrics including: the range of wellformedness conditions contravened; the number of labels; the number of regions; the number of shaded regions; the clutter metric of [29]. From this inspection, no obvious candidate factors for differentiating between the question complexity in phase 1 and phase 3 were found.

To clarify whether the limited experience introduced by phase 2 was influential on user-preference, a supplementary study was conducted in which the questions from phases 1 and 3 were combined. The study was with 15 Masters level students, the high educational level of these subjects appeared to be reflected in the subjects' comprehension performance which was on average 73% (which is 7% higher than the average for the main study). Responses to the preference questions of phases 1 and 3 were less conformant than those in the main study, with an overall average score of 4.8 (see table 4 for a direct comparison of the results). Responses to questions from phase 3 were more conformant than those from phase 1 (though not significantly so). Hence, the effect observed in the main study was reversed in the supplementary when phase 2 was excluded. These additional results confirm that the differing question complexity of the phases 1 and 3 is not significant, and thus is unlikely to be influential in the main study. This strengthens the observation that the conformance is weaker by virtue of the experience of comprehension questions in phase 2.

Table 4. Conformance results for preference questions in the studies

well-formed	Main Study		Supplementary Study	
	"best"	"worst"	"best"	"worst"
yes	74%	25%	66%	36%
no	26%	75%	34%	64%

Given the conformance between subject preferences and well-formed diagrams, the impact of the different wellformedness conditions was also examined. The influence of individual wellformedness conditions on subject preferences were found to be ordered by relevance as follows: *multiple points*, *tangential intersection*, *non-simple contours*, *unique zone labels*, and *concurrency condition*. Hence, the presence of multiple points is the most influential condition upon user responses, and concurrency the least. However, this influence is not strong, and the order varies dramatically when examined for phase 1 and phase 3 separately. In addition, it is hard to draw firm links between this order and the feedback from users, as the users were not introduced to the specific conditions being examined. One area for further study is to encourage subjects to articulate their preference rationales in terms that can be easily related to the proposed wellformedness conditions. Also, the systematic testing of individual sets of conditions is a next step, using the feedback from this study to prioritize which sets of conditions are worth investigating first (given that there are many possible sets of conditions that can be imposed).

5   Discussion and Conclusions

In our exploratory studies with Euler diagrams, we tried to keep the diagrams as simple as possible, whilst still allowing enough variety to present collections of semantically similar diagrams for the subjects to choose from. One reason for this was to ensure that novices to the subject could easily engage with the tasks; the results of the subjects in the comprehension task indicate that the study was set at a level which was sufficiently cognitively demanding. Despite having only just been introduced to the concept of Euler diagrams, we found that subject preferences strongly conformed with the imposition of the wellformedness conditions (a significant result confirming our motivating hypothesis), but that experience with Euler diagrams influences user preference. The link between the wellformedness conditions and user comprehension is only very weakly evidenced by the results from phase 2 of the study (which shows a slight a drop in comprehension for questions relating to the diagram with more instances of non-wellformedness). The authors are currently preparing to examine the link between user-preference and comprehension in the presence of the wellformedness conditions in more detail. We believe that comprehension needs to be examined in both of the settings of interpretation and construction of diagrams.

The "good practice" heuristics that were teased out in our preliminary study (section 3) enabled a greater degree of consistency and a better style of diagrams (which is something that often comes from user-experience). Thus, such scoping heuristics could improve the effectiveness, and help ease the uptake, of the notation. Future work will involve user testing to refine and justify the choice of such heuristics.

One of the long term aims of examining user preferences of wellformedness is to identify which conditions are of greatest use or value to users, and in what circumstances. Any resulting prioritization of wellformedness conditions can be useful in identifying which conditions can be relaxed with minimal disruption for users. Furthermore, it could inform the prioritization of theoretical work on the automatic generation of diagrams (to display output in a diagrammatic theorem proving environment for example). Although a prioritization of the conditions was identified for the study reported, it was not significant or stable and so the examination of specific conditions and their relevance for users requires further study. One feature of considering prioritized conditions is that they may vary between users, in which case a more fundamental question might be whether or not the wellformedness conditions represent coherent presentation constraints for users. For instance, in phase 2 of the study some subjects' interpretations of tangential intersections were not consistent, in that they varied between diagrams. The examination of the effect of working on the comprehension task indicates that this weakened the initial conformance of preference and wellformedness. This suggests that the experience of working with the diagrams enables subjects to tease out and discriminate between aesthetic factors and cognitively demanding factors when interpreting diagrams. Expert users may also have different preferences to novices as they may wish to represent or see information presented more compactly or laid out in a certain manner according to their task. For example, for the question shown in Figure 7, a diagram expert who completed the study preferred the non wellformed middle diagram (which has split zones), and not the wellformed first diagram (which had shaded regions). Therefore, we believe that wellformedness conditions for diagrammatic notations should be treated as presentation

aids, enforced and dismissed as the user wishes in order to aid understanding of the diagrams. Adopting the use of wellformedness conditions is likely to be especially useful for communication with novice users, but their a-priori imposition can restrict the notational utility and have an adverse effect on user perception of the notation.
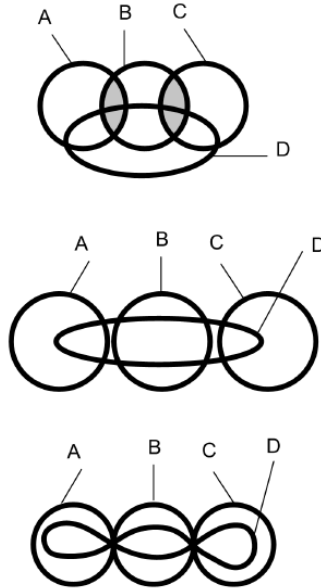
Fig. 7. Phase 3: choose the best and worst

Subjects' informal feedback showed a broad range of concerns ranging from aesthetic, through diagrammatic clarity to more cognitively demanding issues such as diagram semantics. As well as demonstrating the variety of possibly conflicting concerns, this indicates other possible conditions which may influence preference and comprehension (such as symmetry and area proportionality). After suitable testing, notation designers could use such information either to introduce new, or refine existing, wellformedness conditions, or as possible improvements to the scoping heuristics. Ideally this would lead to an improvement in user communication when using these diagrams.

Constraint diagrams [17,4,5] are an extension of Euler diagrams, with added facilities to explicitly express quantification and navigation expressions. They were designed for use as a formal specification and reasoning system in an object oriented setting. We imagine that as systems become more complex (and expressive) by adding extra syntax, the difficulties in understanding by a user will increase and that this may affect user preference for wellformedness. Future studies will be performed to test these more complex notations, but an incremental testing approach was deemed necessary in order to build up a realistic understanding of user's preferences (otherwise there are too many complex interactions to be able to easily isolate any properties to test).

One of the many benefits of this exploratory research is the number of areas and questions that have been identified as needing further testing. In the long term we intend to provide a general framework for testing user preference and comprehension which will enable developers and users to gain insights into the human factors of their favourite notations. Our empirical study points to the need for the thorough investigation of any conditions imposed on a specification notation as a potential source for usability problems, and to consider possible improvements of the choice of conditions imposed based on user preferences. We advocate that a good design of an environment for authoring, viewing and editing diagrammatic specifications would in fact allow a user-based choice of which wellformedness conditions are imposed (which could be different for different users).

References

1. Khazaei, B., Roast, C.: The Influence of Formal Representation on Solution Specification. Requirements Engineering (8), 69–77 (2003)
2. Roast, C.R., Siddiqi, J.I.: Contrasting Models for Visualisation (Seeing the wood through the trees). In: Duce, D., Puerta, A. (eds.) Design, Specification and Verification of Interactive Systems 1999, EuroGraphics. Springer, Wein (1999)
3. Britton, C., Jones, S.: The Untrained Eye: How languages for software specification support understanding in untrained users. Human-computer Interaction 14, 191–244 (1999)
4. Fish, A., Flower, J., Howse, J.: The Semantics of Augmented Constraint Diagrams. Journal of Visual Languages and Computing 16, 541–573 (2005)
5. Kent, S.: Constraint Diagrams: Visualizing Invariants in Object Oriented Modelling. In: Proceedings of OOPSLA 1997, pp. 327–341. ACM Press, New York (1997)
6. Gurr, C.: Effective Diagrammatic Communication: Syntactic, Semantic and Pragmatic Issues. Visual Languages and Computing 10(4) (1999)
7. Shimojima, A.: Operational constraints in diagrammatic reasoning. In: Allwein, G., Barwise, J. (eds.) Logical Reasoning with Diagrams, pp. 27–48. Oxford University Press, Oxford (1996)
8. Euler, L.: Lettres a une Princesse d'Allemagne sur divers sujets de physique et de philosophie. Letters Berne, Socit. Typographique 2, 102–108 (1775)
9. Venn, J.: On the diagrammatic and mechanical representation of propositions and reasoning. Phil. Mag. (1880)
10. Chiara, R.D., Erra, U., Scarano, V.: Vennfs: A venn diagram file manager. In: Proceedings of Information Visualisation, pp. 120–126. IEEE Computer Society, Los Alamitos (2003)
11. Chiara, R.D., Erra, U., Scarano, V.: A system for virtual directories using euler diagrams. In: Proceedings of Euler Diagrams 2004. Electronic Notes in Theoretical Computer Science, vol. 134, pp. 33–53 (2005)
12. Kestler, H., Muller, A., Gress, T., Buchholz, M.: Generalized venn diagrams: a new method of visualizing complex genetic set relations. Journal of Bioinformatics 21(8), 1592–1595 (2005)
13. Hayes, P., Eskridge, T., Saavedra, R., Reichherzer, T., Bobrovnikoff, D.: Collaborative knowledge capture ontologies. In: Proceedings of K-CAP 2005 (2005)
14. Chow, S., Ruskey, F.: Drawing area-proportional venn and euler diagrams. In: Liotta, G. (ed.) GD 2003. LNCS, vol. 2912, pp. 466–477. Springer, Heidelberg (2004)
15. Harel, D.: On visual formalisms. In: Glasgow, J., Narayan, N.H., Chandrasekaran, B. (eds.) Diagrammatic Reasoning, pp. 235–271. MIT Press, Cambridge (1998)

16. Howse, J., Schuman, S.: Precise visual modelling. Journal of Software and Systems Modeling 4, 310–325 (2005)
17. Fish, A., Flower, J.: Investigating reasoning with constraint diagrams. In: Visual Language and Formal Methods 2004, Rome, Italy. ENTCS, vol. 127, pp. 53–69. Elsevier, Amsterdam (2004)
18. Flower, J., Masthoff, J., Stapleton, G.: Generating readable proofs: A heuristic approach to theorem proving with spider diagrams. In: Proceedings of Diagrams 2004, Cambridge, UK, pp. 166–181. Springer, Heidelberg (2004)
19. Hammer, E.: Logic and Visual Information. CSLI Publications (1995)
20. Howse, J., Stapleton, G., Taylor, J.: Spider diagrams. LMS J. Computation and Mathematics 8, 145–194 (2005)
21. Shin, S.J.: The logical Status of Diagrams. Cambridge University Press, Cambridge (1994)
22. Swoboda, N., Allwein, G.: Using DAG transformations to verify Euler/Venn homogeneous and Euler/Venn heterogeneous rules of inference. Journal of Software and Systems Modeling 3(2), 136–149 (2004)
23. Fish, A., Stapleton, G.: Formal issues in languages based on closed curves. In: Proceedings of VLC 2006, Visual Languages and Computing, Grand Canyon, USA, Knowledge Systems Institute, pp. 161–167 (2006)
24. Flower, J., Howse, J.: Generating Euler diagrams. In: Proceedings of Diagrams 2002, Callaway Gardens Georgia, USA, pp. 61–75. Springer, Heidelberg (2002)
25. Swoboda, N., Allwein, G.: Heterogeneous reasoning with Euler/Venn diagrams containing named constants and FOL. In: Proceedings of Euler Diagrams 2004. ENTCS, vol. 134. Elsevier Science, Amsterdam (2005)
26. Ruskey, F.: A survey of Venn diagrams. Electronic Journal of Combinatorics (1997), http://www.combinatorics.org/Surveys/ds5/VennEJC.html
27. Howse, J., Molina, F., Shin, S.J., Taylor, J.: Type-syntax and token-syntax in diagrammatic systems. In: Proceedings FOIS 2001: 2nd International Conference on Formal Ontology in Information Systems, Maine, USA, pp. 174–185. ACM Press, New York (2001)
28. Roast, C.R., Steele, R.A.: Using interfaces and liking interaction. In: Sharp, H., Le Peuple, J., Chalk, P., Rosbottom, J. (eds.) Proceedings of Human-Computer Interaction 2002. BCS, vol. 2, pp. 46–49 (2002) ISBN:1-902505-48-4
29. John, C., Fish, A., Howse, J., Taylor, J.: Exploring the notion of Clutter in euler diagrams. In: Barker-Plummer, D., Cox, R., Swoboda, N. (eds.) Diagrams 2006. LNCS, vol. 4045, pp. 267–282. Springer, Heidelberg (2006)