Represented and Representing Dimensions in Relational Information Displays

Sung Park and Richard Catrambone
School of Psychology
&
Graphics, Visualization, and Usability Center (GVU)
Georgia Institute of Technology
{gtg116s@mail.gatech.edu, rc7@prism.gatech.edu}

Abstract

It is speculated that the processing of different types of information (e.g., quantitative, ordinal, or nominal data) will be affected by what type of visual display is used to present that information (e.g., line graphs, shapes with varying levels of gray saturation, or shapes of different colors). People are expected to be able to more efficiently and accurately process and answer questions about the visual displays if the type of display (i.e., the representing dimension) provides neither too much nor too little information and matches the type of information (i.e., the represented dimension) being processed. In the present study we found that in general task performance was best when the represented and representing dimensions match. An exception to this is discussed.

1. Introduction

People often have trouble finding the right information when looking at graphs. Sometimes this leads to frustration, causes them to discontinue their searching or results in them acquiring false information due to the poor display. A clear understanding of which display is best suited for various types of information enhances the efficiency and accuracy of the user's task.

A variety of graphic displays have been used such as line graphs, bar charts, pie charts, scatter plots, matrices, tables, networks, maps, and many others to represent different types of information. These displays can be categorized under a common name, relational information displays (RIDs), which are displays that represent relations among types of information (e.g., a person's name and her finish in a race) [1].

In this study, RIDs have been analyzed to answer the question of which display is the most effective in representing different types of information. The study investigated three types of displays: line graph, gray saturation, and color, and three types of information: quantitative, ordinal, and nominal.

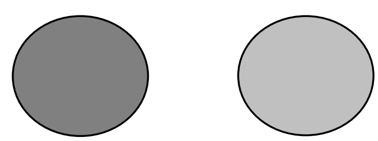
The mapping between type of display and type of information can be studied effectively by the distributed cognition framework because cognitive activity in RID related tasks is distributed into two different representations: an internal representation and an external representation [1].

2. Theory

2.1. Distributed Cognition Framework

Internal representations are information such as propositions, schemas, productions, and mental images that reside in our mind, whereas external representations are physical symbols, graphs, objects as well as the constraints or relations embedded in the physical setting (e.g., the visual and spatial layouts of diagrams) [2].

In an example of a RID task (see Figure 1), the density of the circle codes the value it represents. Viewers are told that the denser the circle, the more value it represents. By a visual inspection of both external representations, the perceptual difference that (a) has more density than (b) can easily be identified. However, the decision of whether (a) represents more value than (b) can be made only by referring to the internal representation (i.e., a rule you have learned) that the denser the circle, the more value it represents. Thus, RID tasks relate two distributed components: an internal representation and an external representation.



(a) 75 (b) 50

Figure 1. An example of a RID task. Each display represents an assigned number and follows the given rule that the denser the circle, the more value it represents.

A key difference between the traditional cognitive view and the distributed cognition view is on how each considers the role of external representations in cognitive processing. The traditional approach to cognition

Proceedings of the Eighth International Conference on Information Visualisation (IV'04)
1093-9547/04 $ 20.00 IEEE

assumes that external objects are limited to being merely a peripheral memory aid [2]. This can be easily accepted when we consider that written digits are used as memory aids for calculation. Consider multiplying 771 by 242 using paper and pencil (see Figure 2). The external representations are the shapes and positions of the symbols and the spatial relations of partial products. Notice the writing of partial products (1542, 3084, and 1542) during the calculation. This frees working memory and allows one to concentrate summing digits [3].

```
       7 7 1
     x 2 4 2
   -----------
       1 5 4 2
     3 0 8 4
   1 5 4 2
 ----------------
   1 8 6 5 8 2
```

Figure 2. An example of an external representation used as a memory aid.

In contrast, the distributed cognition approach explores how cognitive activity is distributed across minds, external artifacts, groups of people, and across space and time [4, 5]. In this view, external representations are not simply memory aids; rather, they are so intrinsic to many cognitive tasks that they can strongly guide and constrain cognitive behavior [6].

One of the most promising methodologies that has been adopted in distributed cognition is the *representational analysis* which investigates representational effect. The *representational effect* refers to the phenomenon that different representations of a common formal structure can cause dramatically different cognitive behaviors. A good example is the representation of numbers [7]. Arabic numerals are more efficient than Roman numerals for multiplication, even when both types of numerals represent the same numbers. For example, 73 x 27 is easier than LXXIII x XXVII because in the latter one, we have to convert each Roman numeral to Arabic prior to multiplication [2]. In this case, we are looking at two different representations, Arabic numerals and Roman numerals, which represent the same information.

Representational analyses have been conducted in various areas, mostly in problem solving and reasoning domains such as Tower of Hanoi and Tic-Tac-Toe [6]. These studies commonly found that the form of a representation determines what information can be perceived, what processes can be activated, and what structures can be discovered in a problem solving situation.

Representational analysis can effectively investigate RIDs as well. For each type of information (quantitative, ordinal, nominal) different features (length, density, color) conveying isomorphic information can be compared to see how they lead to differences in cognitive processing. This ties back to the original question of which display is optimal in conveying different types of information.

2.2. Scale of Measurements in RIDs

Zhang [1] developed a theoretical framework that accounts for the representational effects of all RIDs. The framework suggests that although there are no optimal displays that are efficient for all types of tasks, there can be a correct or incorrect mapping between the representation of a display and the structure of a task. One aspect of his framework involved the psychological scale of the information that the display conveys.

Stevens [8] identified four types of psychological scales: ratio, interval, ordinal, and nominal. Each type can be characterized as having or lacking the following properties: category, magnitude, equal interval, and absolute zero (see Table 1).

Table 1. The Properties of Scales

Properties	Scale Types			
	ratio	interval	ordinal	nominal
category	yes	yes	yes	yes
magnitude	yes	yes	yes	no
equal interval	yes	yes	no	no
absolute zero	yes	no	no	no

The "category" property means that the instances on a scale can be distinguished from one another. The "magnitude" property means that one instance on a scale can be recognized as greater than, less than, or equal to another instance on the same scale. The "equal interval" property means that the magnitude of an instance represented by a unit on the scale is the same regardless of where on the scale the unit is. An "absolute zero" is a value which indicates that nothing at all of the property being represented exists.

Every dimension, whether it is a physical dimension, such as the length of a line, or a more abstract dimension such as an amount of money, is on a certain type of scale [1]. In RIDs, there are two dimensions: (1) the information and, (2) the display that conveys that information. We can identify the scale property for each dimension and form a two dimensional matrix.

2.3. Represented and Representing Dimensions

The two dimensions are formally titled as: represented and representing dimensions. The *represented dimensions* of a RID are the dimensions of an original domain that are to be represented by various RIDs. The *representing dimensions* of a RID are the physical dimensions of the RID that are used to represent the dimensions of the original domain (See Figure 3) [1].

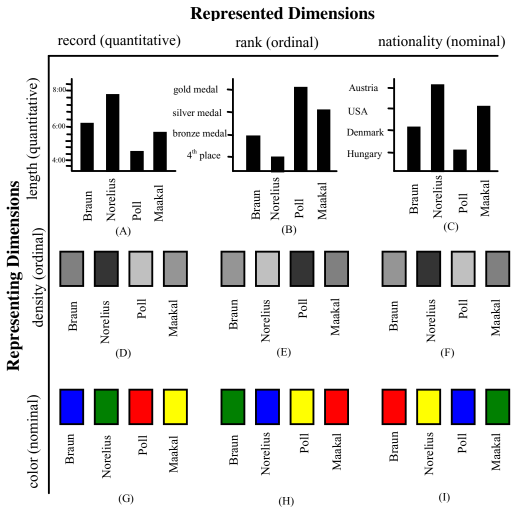


Figure 3. The mapping between the represented and representing dimensions. The scale type of each entity is inside the parentheses. Records are from a fictional women's 400 meter freestyle swimming competition.

In Figure 3, there are three ways of presenting information: length, density and color. The three types of displays were chosen to represent three types of scale properties (quantitative, ordinal, nominal) each conveys.

In order for a representation to be efficient and accurate, the represented and representing dimensions should match in scale. In the three displays on the diagonal in Figure 3 (A, E, I), the scale of the represented dimensions matches the scale of the representing dimensions. For example, in Figure 3A, swimming records are represented by the length of the bars, which has the same quantitative property. In Figure 3E, the ranking of swimmers is represented by density which has the same ordinal property. In this condition, the user is told to assume that the denser the saturation, the higher the rank. In Figure 3I, each individual's nationality is represented by color which has the same nominal property.

In the three displays above the diagonal (B, C, F), the representing dimensions have more information than the represented dimensions. The extra information of the representing information may cause misperceptions on the represented dimensions. For example, in Figure 3B, a gold medalist may be perceived as twice as superior as a bronze medalist. This is a misperception, because gold, silver, and bronze medal only indicate the relative (ordinal) standings among the players. In other words, the display does not intend to provide quantitative information. This type of potential misleading is termed the *language problem* in Mackinlay's psychological analysis [9].

In the three displays below the diagonal (D, G, H), the representing dimensions have less information than the represented dimensions. In these displays, the extra information of the represented dimensions either has to be represented internally or is not represented at all. For example, in Figure 3H, color possesses no ordinal information about the rankings of the players. If a rule is given, for example, that blue is higher than red in relative standings, the user is forced to refer to their internally represented rule prior to making a decision about which color represents bigger value [1].

The goal of the present research is to examine implications of the represented and representing dimensions framework. The key hypothesis is that when the represented and representing dimensions match, a person's performance will be best. This leads to three predictions:

1) Performance in retrieving quantitative information is optimal when the display has the same quantitative scale property.
2) Performance in retrieving ordinal information is optimal when the display has the same ordinal scale property.
3) Performance in retrieving nominal information is optimal when the display has the same nominal scale property.

3. Experiment

3.1. Method

3.1.1. Participants Participants were 48 undergraduate students that were enrolled in introductory psychology courses at the Georgia Institute of Technology who volunteered for the experiment to earn course credit.

3.1.2. Materials Participants viewed a set of RIDs consisting of a display and question on the same screen (see Appendix 1-9). Flash and Flash Actionscript were used to implement the condition sets on IBM PCs.

3.1.3. Procedure Each person completed a pre-session in order to learn how to respond to the question set when viewing a display. An example question that is comparable to the actual question set from all three levels (quantitative, ordinal, nominal) was provided. The experimenter read the question out loud and walked through each step of understanding the question, viewing the display, acquiring appropriate information, and selecting an answer. Participants were asked at each step whether they understood. This orientation procedure was used so that everyone would begin the experiment already accustomed to the environment.

After this orientation, participants performed nine information retrieval tasks that corresponded to nine types of trials shown in Figure 3. There were four instances of each trial type resulting in 36 trials per participant. All contents were records from a fictional 400 meter swimming competition. The task for each condition depended on its represented dimension (column). For example, in 3A, 3D, and 3G (see Figure 3), a participant was asked to answer a question that could be solved only by considering the quantitative property of the display such as asking who was three times faster than Emily (see Appendices 1, 4, 7). Participants were asked to click on the answer at the bottom.

In 3B, 3E, and 3H, a participant was asked to answer a question that could be solved only by considering the ordinal property of the display such as asking participants to order the rankings of the swimmers (see Appendices 2, 5, 8).

In 3C, 3F, and 3I, a participant was asked to answer a question that could be solved only by considering the nominal property of the display such as asking for swimmer's home state (see Appendices 3, 6, 9).

One methodological comment about the density display needs to be made. We wanted to be sure that each density level was discriminable. For example, in Appendix 5, the density level that represents Emily is distinguishable against the others. When levels are laid from black to white, there is a limited number of separate levels that can be acquired maintaining those discriminable. Each RGB level for gray saturation increases together from 0 to 255 (e.g. R = 2, G = 2, B = 2 and R = 4, G = 4, B = 4) where 0 is black and 255 is white. Informal pilot testing suggested that when the interval between levels becomes less than about 36, the visual distinction among levels starts to break down. Therefore, the experiment used eight density levels (i.e., a set of eight choices) that had an interval greater than 36.

3.1.4. Design The experiment was a mixed design with two within-subjects variables (represented and representing dimensions) and one between-subjects variable (four possible orders of the 36 trials). The dependent measures were whether or not a task was done correctly and the time taken to do the task.

To avoid possible sequential effects, counterbalancing methods were used. The questions and the choice sets were ordered so as to prevent participants from answering with the information from previous trials. In addition, two constraints were implemented when randomizing the order of trials. First, the same type of question (quantitative, ordinal, nominal) could not be asked consecutively. Second, the same type of display (length, density, color) could not be shown on consecutive trials. These constraints resulted in four sets of computer generated random orders (pseudo orders) with each order having at least 10 participants.

3.2. Results and Discussion

Statistical analyses were performed by taking the following steps: (1) three-way ANOVA (represented dimensions vs representing dimensions vs pseudo orders), (2) simple effect analysis based on planned hypotheses, and (3) a step-wise regression analysis. Pseudo order has no effect and all results and analyses will be collapsed over this variable.

ANOVA Analysis among the three dimensions (See Table 3) indicates that there is a main effect of display type (length, density, color), $F(2,88) = 125.51$, $p < .001$, and task (quantitative, ordinal, nominal), $F(2,88) = 392.07$, $p < .001$. A significant two-way interaction between display and task was found, $F(4, 176) = 71.61$, $p < .001$.

This analysis was based on reaction time that included both correct and incorrect trials. For completeness, another analysis was performed excluding incorrect trials.

Excluding incorrect trials, there continued to be a main effect of display type (length, density, color), $F(2,88) = 8.61$, $p < .001$, and task (quantitative, ordinal, nominal), $F(2,88) = 33.51$, $p < .001$. A significant two-way interaction between display and task was also found, $F(4, 176) = 6.46$, $p < .001$.

Simple Effects Analysis The interaction between display and task suggests analyzing simple effects at each level. Simple main effects and simple contrasts were analyzed to test the three hypotheses (see Table 2). Results indicate that all three levels (quantitative, ordinal, nominal) of tasks have a significant simple effect on displays.

Table 2. p Values for Testing Hypotheses

Comparisons	Including Incorrect Trials	Excluding Incorrect Trials
Quantitative Tasks		
Simple Main Effect	< .001	< .05
Length vs Density	< .001	< .05
Length vs Color	< .001	< .001
Ordinal Tasks		
Simple Main Effect	< .001	< .001
Density vs Length	> .94	> .96
Density vs Color	< .001	< .001
Nominal Tasks		
Simple Main Effect	< .001	< .001
Color vs Length	< .001	< .001
Color vs Density	< .001	< .001

To test the first hypothesis, that performance in retrieving quantitative information is optimal when the display has the same quantitative scale property, simple contrasts between length (in line graphs) and density and between length and color were analyzed. Results indicate that the length display conveying quantitative information yielded significantly faster reaction time than the other two (see mean differences in Table 3). Therefore, the first hypothesis was supported.

To test the second hypothesis, that performance in retrieving ordinal information is optimal when the display has the same ordinal scale property, simple contrasts between density and length and between density and color were analyzed. Results indicate that while the density display conveying ordinal information yielded significantly faster reaction time than color, it did not differ from length. Potential reasons for this are addressed in the general discussion.

To test the third hypothesis, that performance in retrieving nominal information is optimal when the display has the same nominal scale property, simple contrasts between color and length and between color and density were analyzed. Results indicate that the color display conveying nominal information yielded significantly faster reaction time than the other two. Therefore, the third hypothesis was supported.

Table 3. Average reaction time in seconds including incorrect responses for each condition. Percent correct is inside the parenthesis.

	Task		
Display	Quantitative	Ordinal	Nominal
Length	9.4 (95%)	28.4 (85%)	6.7 (100%)
Density	27.7 (42%)	28.5 (95%)	15.3 (71%)
Color	21.8 (91%)	43.3 (98%)	5.3 (100%)

Regression A stepwise multiple-regression analysis was carried out on reaction time. There were 1728 data points, 36 (one per task) for each of the 48 participants.

To code categorical predictors such as display and task, dummy coding was used. For example, display-type was represented by two variables (diplaydummy1 and displaydummy2) that could take on values of 0 or 1 (0, 0 would mean quantitative, 0, 1 would mean ordinal, and 1, 0 would mean nominal). Suppose displaydummy1 were to enter into the regression equation. This would mean that the nominal display was significantly different from the other two displays. Similarly, if displaydumm2 entered the equation, it would mean that the ordinal display was significantly different from the other two displays. In order to find out whether the quantitative display was different though, the dummy variables would have to be recoded so that either 0, 1 or 1, 0 was assigned to represent a quantitative display. Thus, two regression analyses were run in order to cover all possible levels.

Table 3 summarizes a regression equation that includes length and color levels in display factor and quantitative and nominal levels in task factor while Table 4 summarizes a regression equation that includes density and color levels in display factor and ordinal and nominal levels in task factor.

At each step, an "*F*-to-enter" statistic was calculated for each predictor to determine which one would enter the equation next. The predictor with the largest *F*-to-enter value was entered into the equation at step *n*. The criteria set for *F*-to-enter were at .05 and *F*-to-remove at .1.

Table 4. Summary of Stepwise Regression (first run) Analysis for Variables Predicting Reaction Time in Tasks

Variable	Final Coefficient	Standard Coefficient	R^2 Total	Corre-lation
y-intercept	49.39			
Nominal	-24.48	-.69	.242	$-.49^{**}$
Quantitative	-15.41	-.43	.355	$-.04^{*}$
Line	-6.98	-.19	.417	$-.24^{**}$
Correctness	-8.19	-.16	.440	$-.15^{**}$
English first	-4.82	-.11	.456	$-.12^{**}$
Trial	-0.19	-.11	.469	$-.11^{**}$
Color	2.00	.06	.472	$.18^{**}$
Gender	1.49	.04	.474	$.06^{**}$

Table 5. Summary of Stepwise Regression (second run) Analysis for Variables Predicting Reaction Time in Tasks

Variable	Final Coefficient	Standard Coefficient	R^2 Total	Corre-lation
y-intercept	26.99			
Ordinal	15.41	.43	.242	$.54^{**}$
Nominal	-9.07	-.25	.355	$-.49^{**}$
Correctness	-8.19	-.16	.391	$-.15^{**}$
Color	8.98	.25	.416	$.12^{**}$
Density	6.98	.20	.442	$.13^{**}$
English first	-4.82	-.11	.458	$-.12^{**}$
Trial	-0.19	-.11	.472	$-.11^{**}$
Gender	1.49	.04	.474	$.06^{**}$

$^{*}p < .05.$ $^{**}p < .01.$

The predictors that ended up in the final two equations are as follows:

1. *task (quantitative, ordinal, nominal)*: Collectively, three levels in task explained the variance of reaction time the most. While the ordinal task (see Appendices 2, 5, 8) required participants to analyze and select eight choices (marginal mean of 33.41 sec), the nominal task (see Appendices 3, 6, 9) required a single response (marginal mean of 9.11 sec).
2. *correctness:* This represents whether each response was correct or not. It indicates that participants were slower on the trials on which they were incorrect.
3. *display (length, density, color)*: Collectively, the three levels in display explained the variance of reaction time significantly. Participants were the slowest in the color task (marginal mean of 38.69) and the fastest in the length task (marginal mean of 14.85).
4. *English first*: This represents whether English was the participant's first language. Because the task requires understanding the question in English it makes sense that domestic students were faster. However, it is interesting to see that SAT verbal scores did not enter the regression equation.

5. *trial*: This codes for effect of serial order. It shows that participants became faster over tasks.

6. *gender*: This codes participant's gender. This shows males were slower than females.

Potential predictor variables that did not end up in the final equations were: age, GPA, SAT verbal scores, and SAT math scores. Ultimately, about 47% of the variance in participant's reaction time on each task was accounted for by the final equation.

4. General Discussion

The key hypothesis in this study was generally supported: optimal performance will occur when there is a match between the represented and representing dimensions. However, an exception was observed in length versus density for conveying ordinal information.

In the original represented and representing dimensions framework, length has an additional scale of measurement (quantitative) over density (ordinal). Zhang [1] suggested that due to this redundant scale of measurement, viewers may be misled to perceive unintended quantitative information in the line graph. For example, in Appendix 2, the difference between Paul and Alexis is that one was 60th and the other 30th in rankings. While this is only a relative standing, due to the quantitative nature of length the viewer might conclude that Alexis was twice as fast as Paul.

We assumed that such mistakes would lead to longer reaction times. One reason we did not find a cost to using line length for ordinal information is because our ordinal task clearly directs users to focus on ordinal information. This might reduce the chance of users being misled. Future research might use less-directive instructions (better mirroring the real world) to see if the cost of the mismatch emerges.

Another possible factor to consider is Mackinlay's psychological analysis on ranking of perceptual tasks over ordinal information. He suggested that density is better than length in conveying ordinal information primarily because while length judgment suffers from distance effect density does not.

For example, in Figure 4, while the comparison of size (ordinal information) in lines (length) is easily perceived in (A), it is difficult to do so in (B). On the other hand, density judgment does not suffer from the distance effect.

However, while it is true that length suffers from the distance effect, density also has its own limitation in that there are not many levels between black and white that are discriminable from one another. As mentioned earlier, while eight density levels seemed to be the perceptual limit with our computer monitors, length can have an unlimited number of representations as long as the screen real estate permits it.

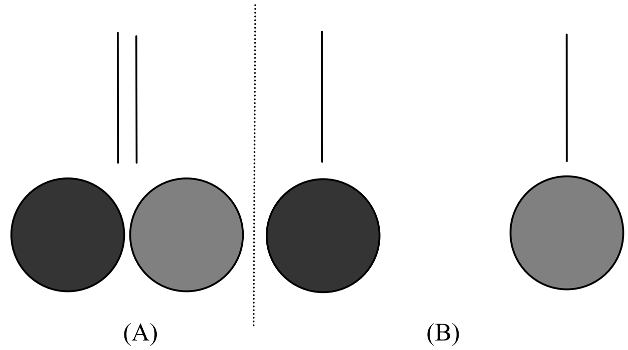


Figure 4. An example of distance effect

5. Conclusion

This study contributes to the foundational theory of Information Visualization. The main focus is to provide a fundamental advancement in our understanding of how visualization should build on human limitations and capabilities.

While Bertin provided a classical theory on graphical semiotics [10], Cleveland and McGill [11] developed a ranking of quantitative tasks using psychophysics law and their own empirical study. Mackinlay [9] complements this theory by providing psychological analysis over ordinal and nominal tasks. Zhang [1] combines all this into the represented and representing dimensions in the light of distributed cognition framework.

This study has shown that with the exception of the ordinal task, the scale of measure between the task (i.e., represented dimension) and the display (i.e., representing dimension) should match in order for a representation to be efficient and accurate.

References

[1] Zhang, J. A representational analysis of relational information displays. In *International Journal of Human-Computer Studies*, 45, 59-74. 1996.

[2] Zhang, J., & Norman, D. A. Representations in distributed cognitive tasks. In *Cognitive Science*, 18, 87-122. 1994.

[3] Zhang, J. External Representations in Complex Information Processing Tasks. In *Encyclopedia of Library and Information Science*. New York, Marcel Dekker. 2000.

[4] Hutchins, E. The technology of team navigation. In J. Galegher, R. Kraut, & C. Egido (Ed.), *Intellectual teamwork: Social and technical bases of collaborative work*. Hillsdale, NJ, Erlbaum. 1990.

[5] Hutchins, E., & Norman, D. A. Distributed cognition in aviation: *A concept paper for NASA* (Grant No. NCC 2-591). San Diego: University of California, Department of Cognitive Science. 1988.

[6] Zhang, J. The nature of external representations in problem solving. In *Cognitive Science*, 21(2), 179-217. 1997.

[7] Nickerson, R. Counting, computing, and the representation of numbers. In *Human Factors*, 30, 181-199. 1988.

[8] Stevens, S. S. On the theory of scales of measurement. In *Science*, 103(2684), 677-680. 1946.

[9] Mackinlay, J. D. Automating the design of graphical presentations of relational information. In *ACM Transactions on Graphics*, 5(2), 110-141. 1986.

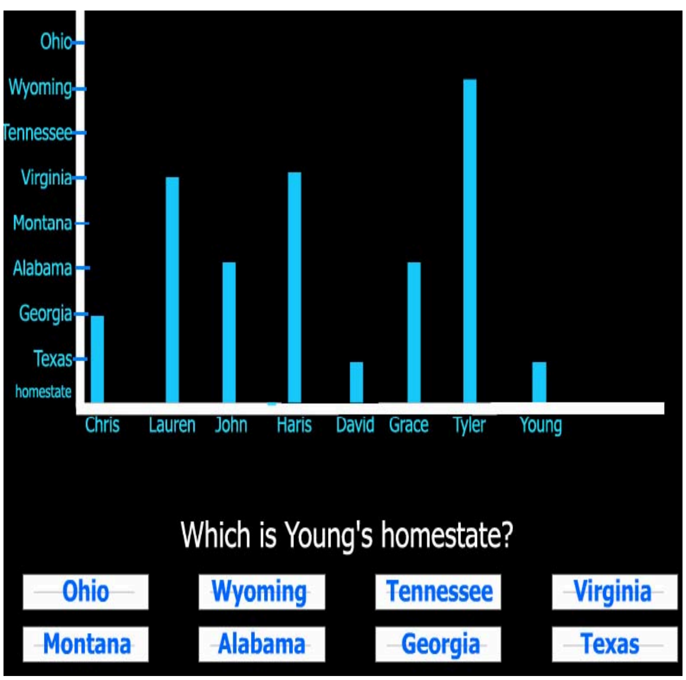
[10] Bertin, J. Semiology of Graphics: Diagrams, Networks, Maps. Madison, WI. University of Wisconsin Press. 1967.

[11] Cleveland, W. S. & McGill, R. Graphical perception and graphical methods for analyzing scientific data, In *Science*, 828-833. 1985.

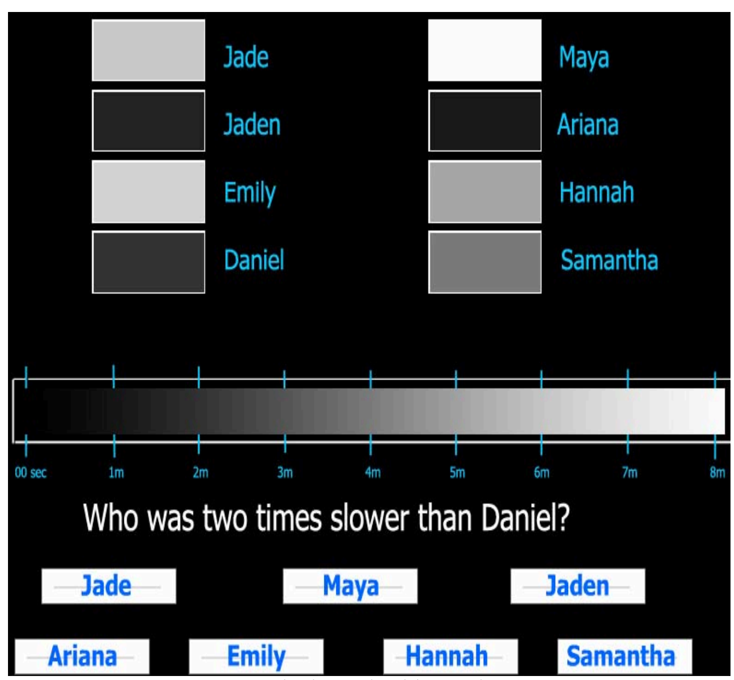
Appendix

1. Quantitative task with Line Graph (length)

2. Ordinal task with Line Graph (length)

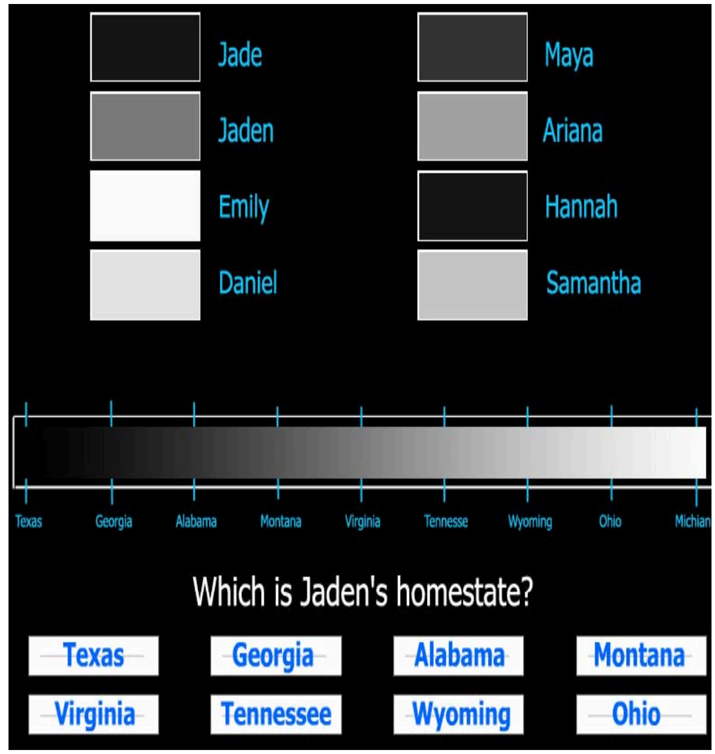


3. Nominal task with Line Graph (length)

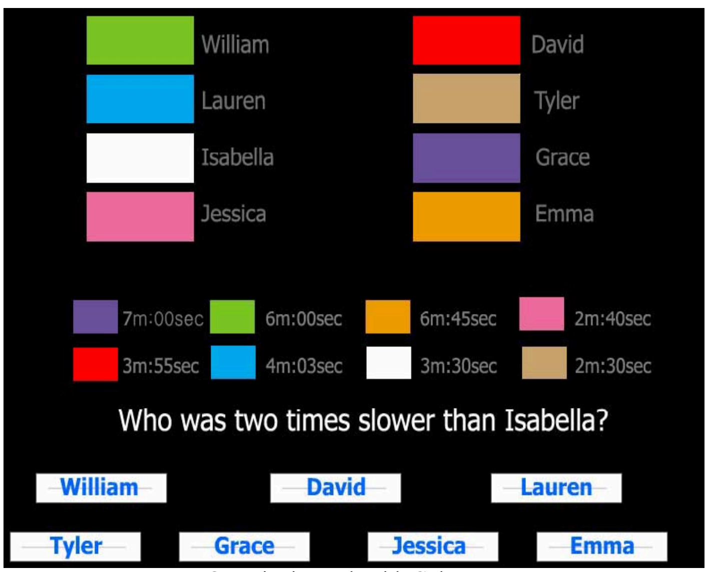


4. Quantitative task with Density

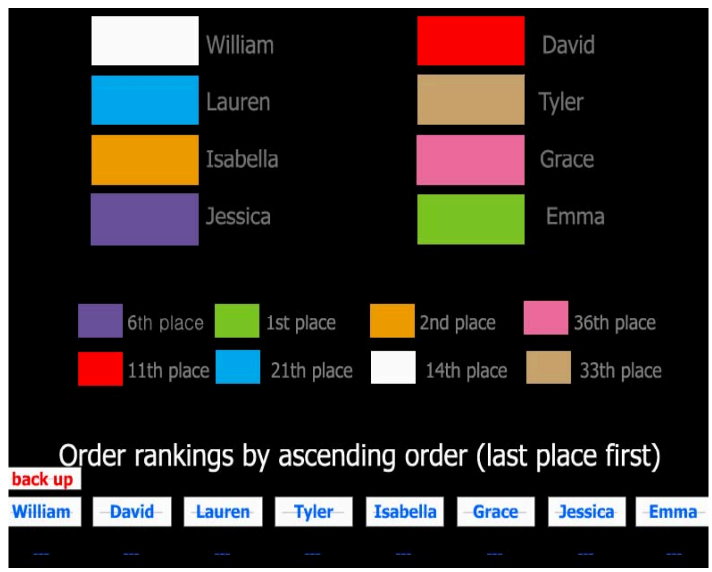
5. Ordinal task with Density



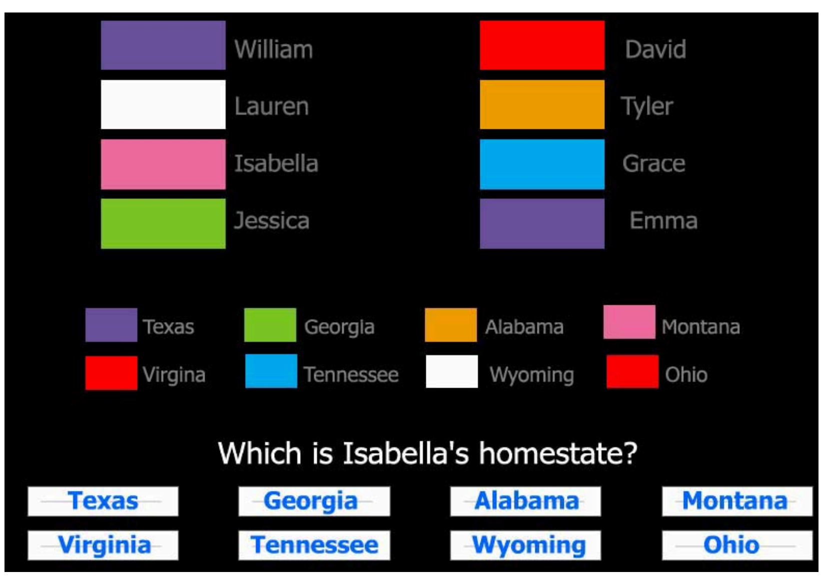
6. Nominal task with Density



7. Quantitative task with Color



8. Ordinal task with Color



9. Nominal task with Color