

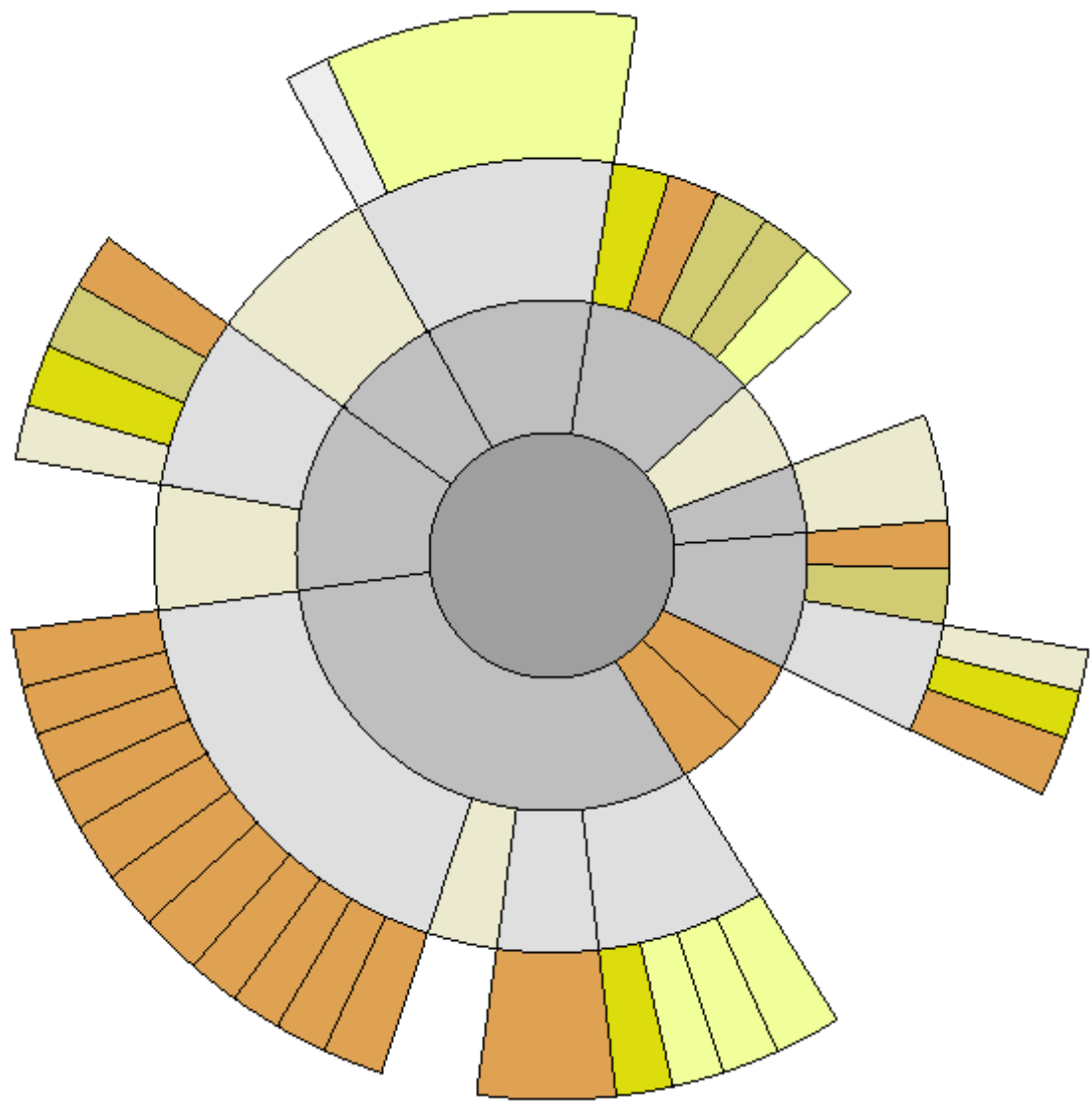
Building effective **scales** for evaluating **subjective experience** in data visualization



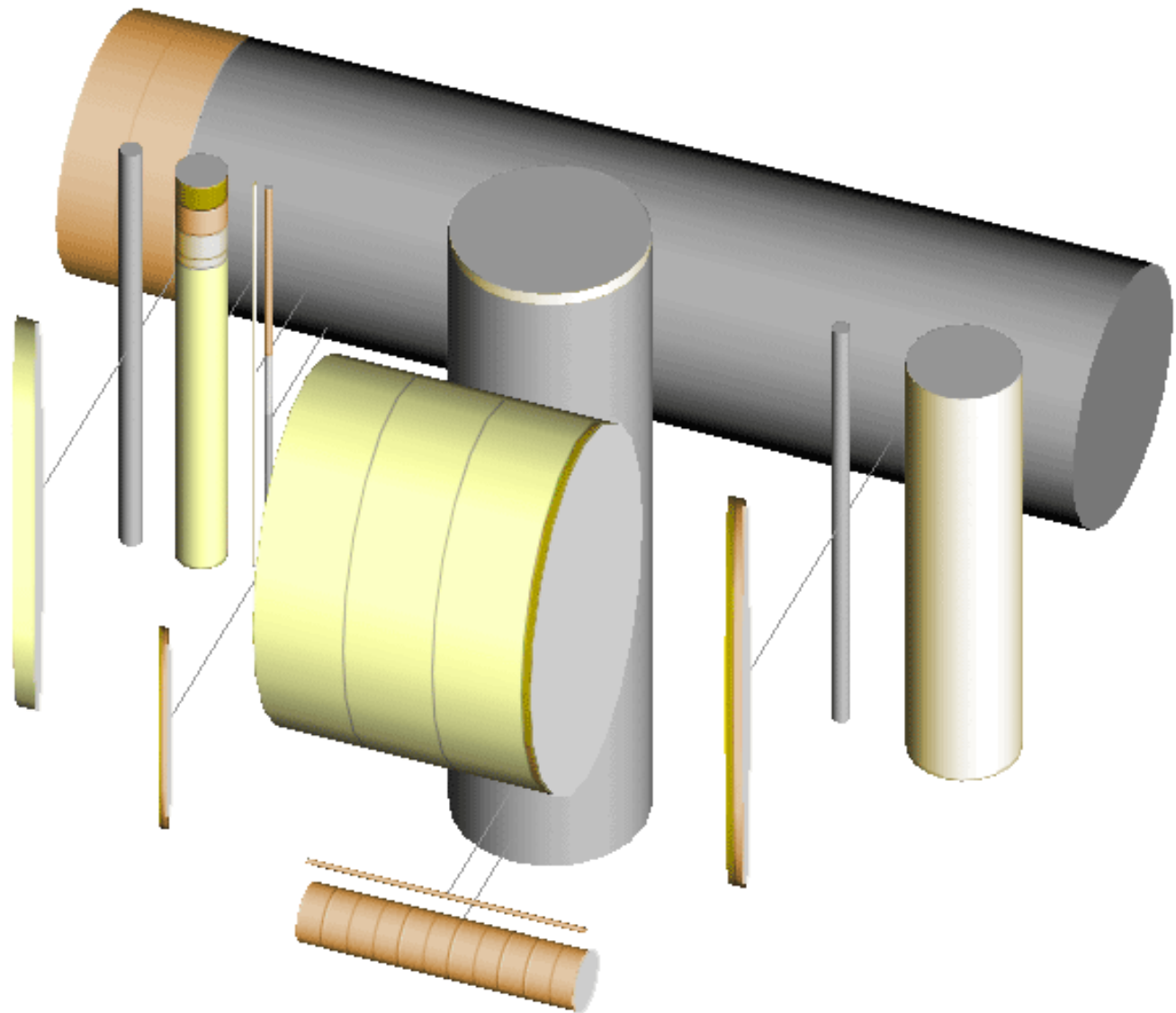
Petra Isenberg and Anne-Flore Cabouat

Université Paris-Saclay, CNRS, Inria, LISN, Télécom Paris



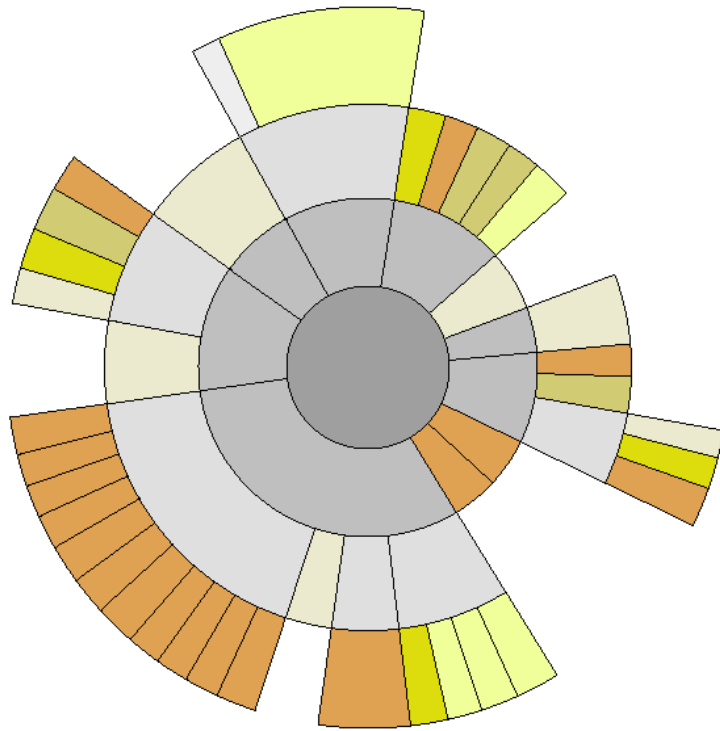


SunBurst

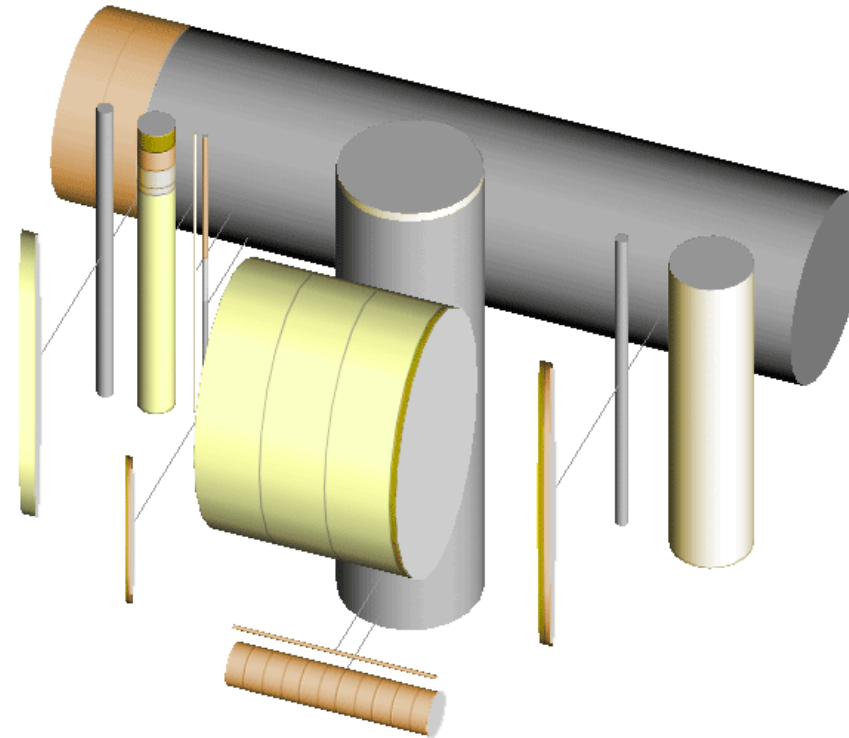


BeamTree

- ▶ Which visualization is more **beautiful**?
- ▶ Do you all agree?
- ▶ By how much?

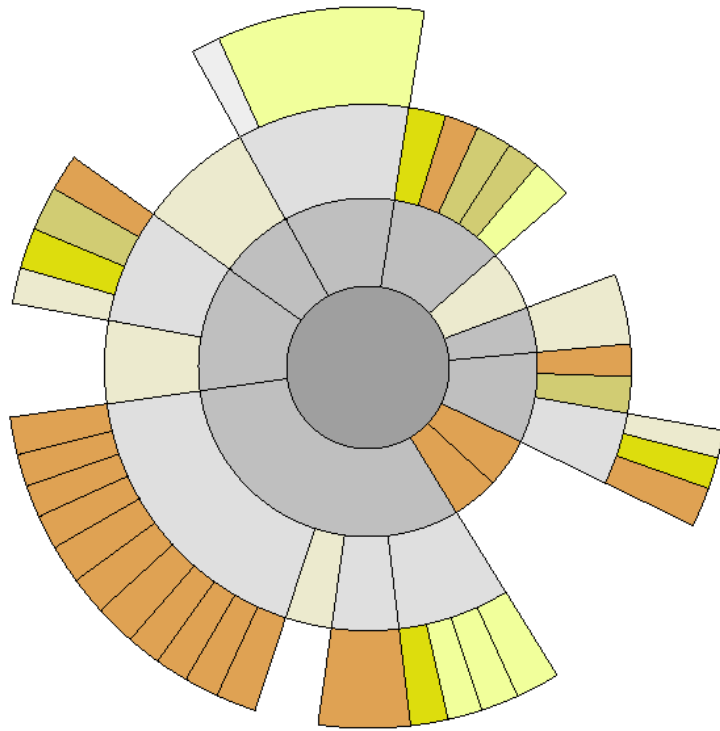


SunBurst

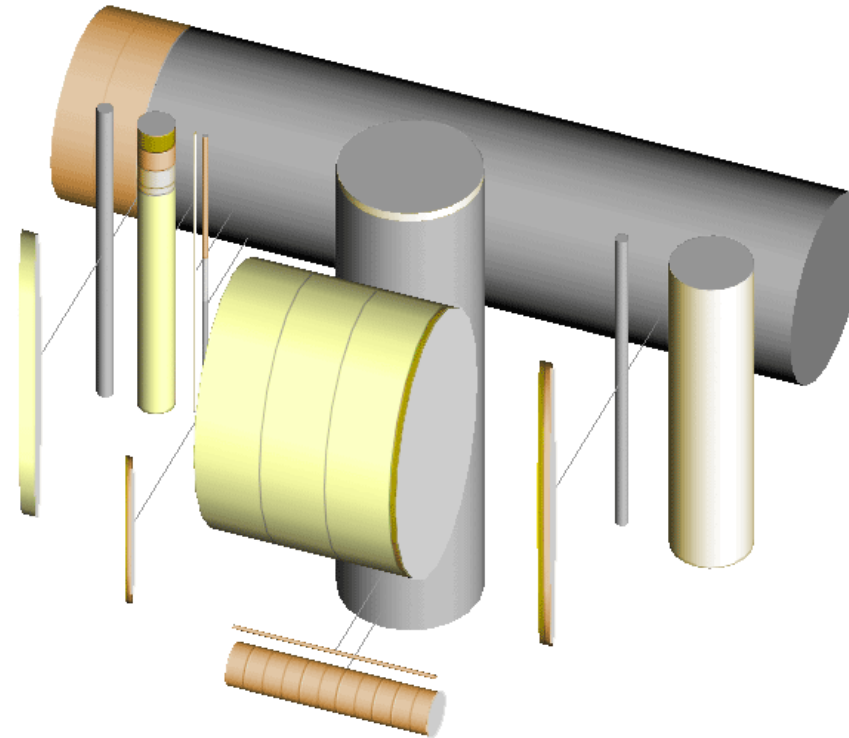


BeamTree

- ▶ Which visualization do you find more **readable**?
- ▶ Do you all agree?
- ▶ By how much?



SunBurst



BeamTree

Rating Scales

To what extent do you agree or disagree with the following statements: _____

	Strongly disagree							Strongly agree
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

- ▶ Which visualization is more **beautiful**?
- ▶ Do you all agree?
- ▶ By how much?

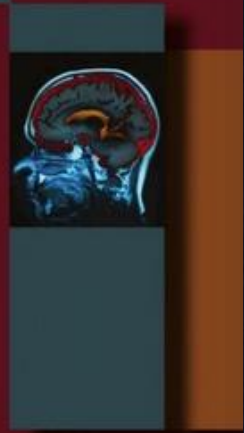


<https://tinyurl.com/beauvis>

Rating
and T
for M
and E
Achie
Mus

HAND
Clin
Scale
in Ps
and

Har
Child
Neu
Rat



Second E

Includes
access to
download
PDFs
rating s

Edited by
Lee Baer,

Rober



Quantitative
Applications
in the
Social Sciences
193

Exploring Rating Scale Functioning for Survey Research

Stefanie A. Wind

HOW ABOUT VIS?

- ▶ We need new & dedicated measurement instruments
- ▶ For subjective experiences

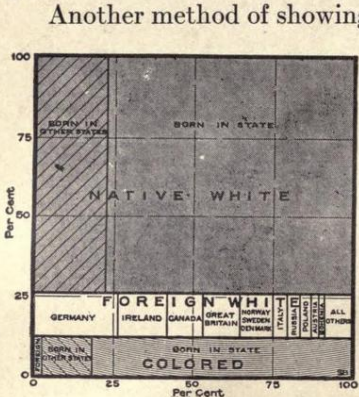
BeauVis

A Validated Scale for Measuring
the Aesthetic Pleasure of Visual
Representations



Tingying He, Petra Isenberg, Raimund Dachzelt, and Tobias Isenberg
Université Paris-Saclay, CNRS, Inria, LISN & Technische Universität Dresden

fractions of a division at both beginning and end of the component to be measured. The best thing in charts of this kind is to use unruled paper, and specify the scale. The reader, if he wants the exact data, can take his measurements with a ruler. The engineer's scale has its subdivisions in decimals and hence is the most convenient scale for chart work. An engineer's scale should be part of the equipment of every person who has charts to make.



United States Statistical Atlas for the Census of 1900
 Fig. 5. Elements of the Population of the United States in 1900

Here the components of the total population are shown in their relative sizes on the vertical scale. Each component is also divided into different subdivisions whose percentage size may be read from the horizontal scale. This is an admirable method of presentation if components must be subdivided

out detracting from the utility of the chart itself. The lettering was done by hand and shows the possibilities for neatness resulting from hand work when a skilled draftsman is employed. In many ways the method of Fig. 5 is preferable, but it is probably true that Fig. 6 would be more readily understood by the average untrained reader.

Where there are a large number of items to be compared and the components of each item are given, the method of Fig. 7 is a very convenient one. The Census Atlas for the 1900 Census contained many pages of charts of this type for its comparisons of different States. By placing bars for all the States on one page, the total for the country is shown as 100 per cent in the vertical direction. No vertical scale

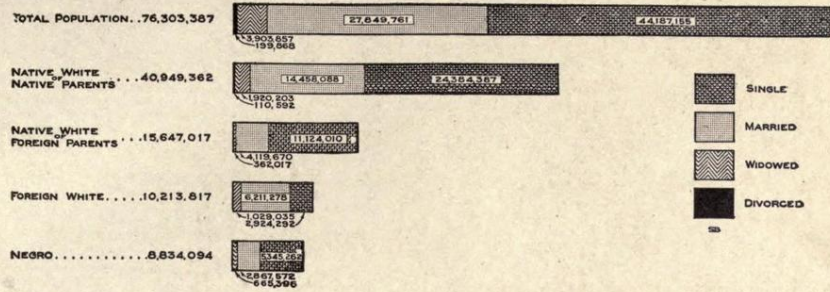


Fig. 6. Conjugal Condition of the Population of the United States in 1900
 The four lower bars show components of the total population represented by the upper bar. The combined length of the four lower bars equals the length of the upper bar

is used, all the bars being made of equal width. In this particular case (Fig. 7) we have a total split into its components and again subdivided so as to show the prevalence of a second factor which is included in the first. Thus, we see the proportion of illiterates in each of the main groups of population for each State. All of the States are shown on the same basis, since all are depicted by bars of the same length representing 100 per cent. It is not easy to make a clear black-and-white drawing if one kind of cross-hatching must be placed on top of another kind. Fig. 7 shows that it is possible, however, to superimpose two kinds of cross-hatching and get a drawing that is fairly clear. The facts in this chart would have been brought out better if colors had been used for the main divisions of population. Ruled cross-hatching in black to represent the percentages of illit-

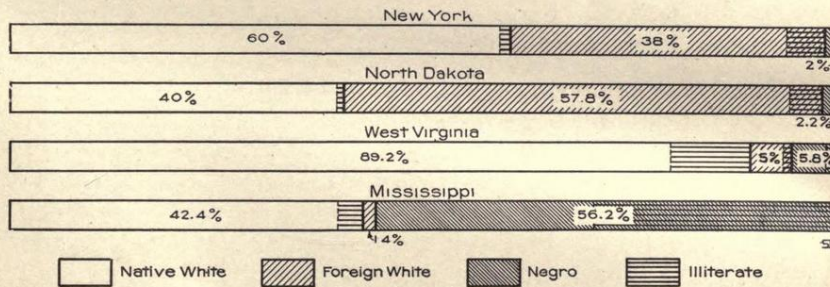
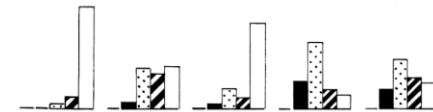
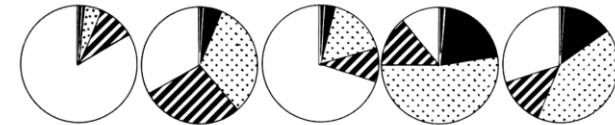


Fig. 7. Males of Voting Age in certain States, in 1900, by Color and Nativity and by Illiteracy

A method for two successive divisions into components is shown here. The proportion of illiterates in each group is brought out by the horizontal ruling

The beauty of b/w textures



[Bertin, 1967]

[Brinton, 1914]

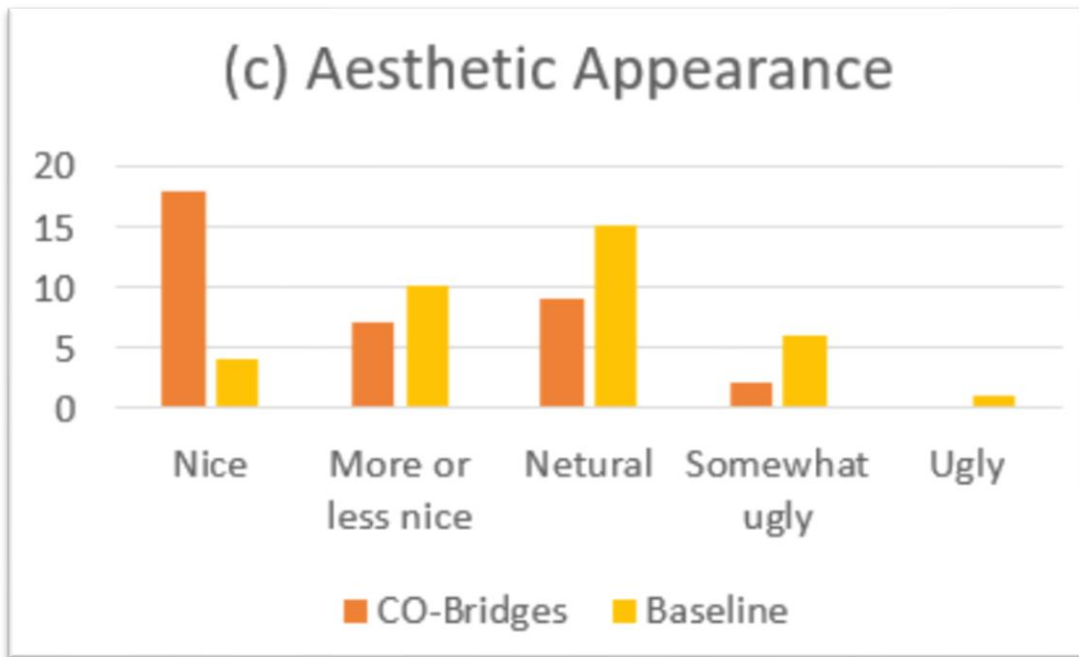
Beauty

An important aspect
in visualization

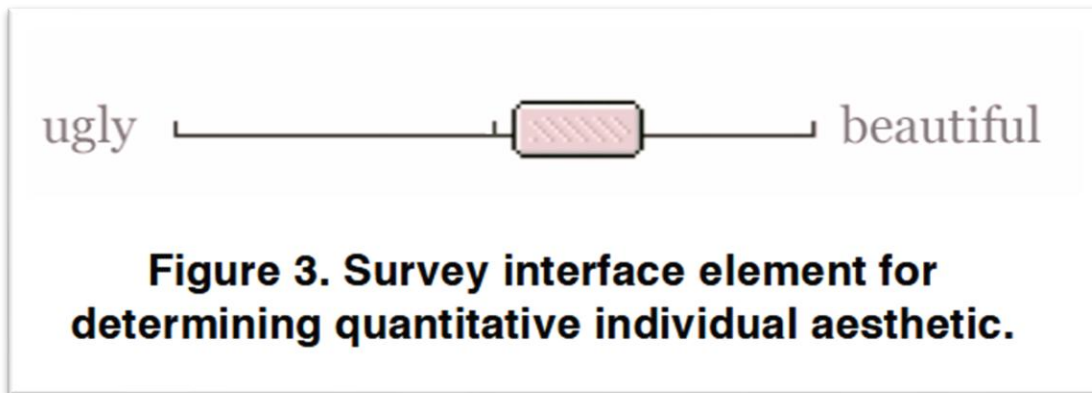
- ▶ Affects usability and effectiveness
[Cawthon & Vande Moere, 2007; Healey & Enns, 2022]
- ▶ Communicates
[Brath et al., 2005]
- ▶ Engages viewers
[Bach et al., 2013; Tateosian et al., 2007]



How can we assess
beauty in
visualization



[Chen et al., 2021]



[Cawthon & Van de Moere, 2007]

1. The graph has good **aesthetic** appearance. 1 2 3 4 5 6 7

2. The graph has good **readability**. 1 2 3 4 5 6 7

[Bu et al., 2020]

Table 6. Overall ratings of DataShot on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree).

Assessment	Measurement	Mean	SD
Content	Insightful	3.90	.57
	Comprehensive	3.90	.74
	Effective	4.60	.52
Visual	Easy to understand	4.70	.48
	Aesthetic	3.90	.74
	Expressive	3.70	.67
System	Useful	4.00	.67
	Usability	4.30	.67

[Wang et al., 2020]

Problem 1

There is **no standard** way to evaluate the “aesthetics” of a visualization

Problem 2

What do we **mean** by
“beauty” or “aesthetics”?

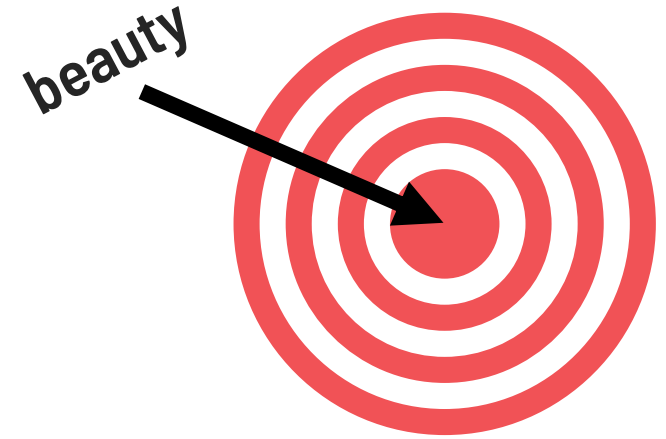
Reliability vs. Validity

- ▶ Do people rate a phenomenon consistently?

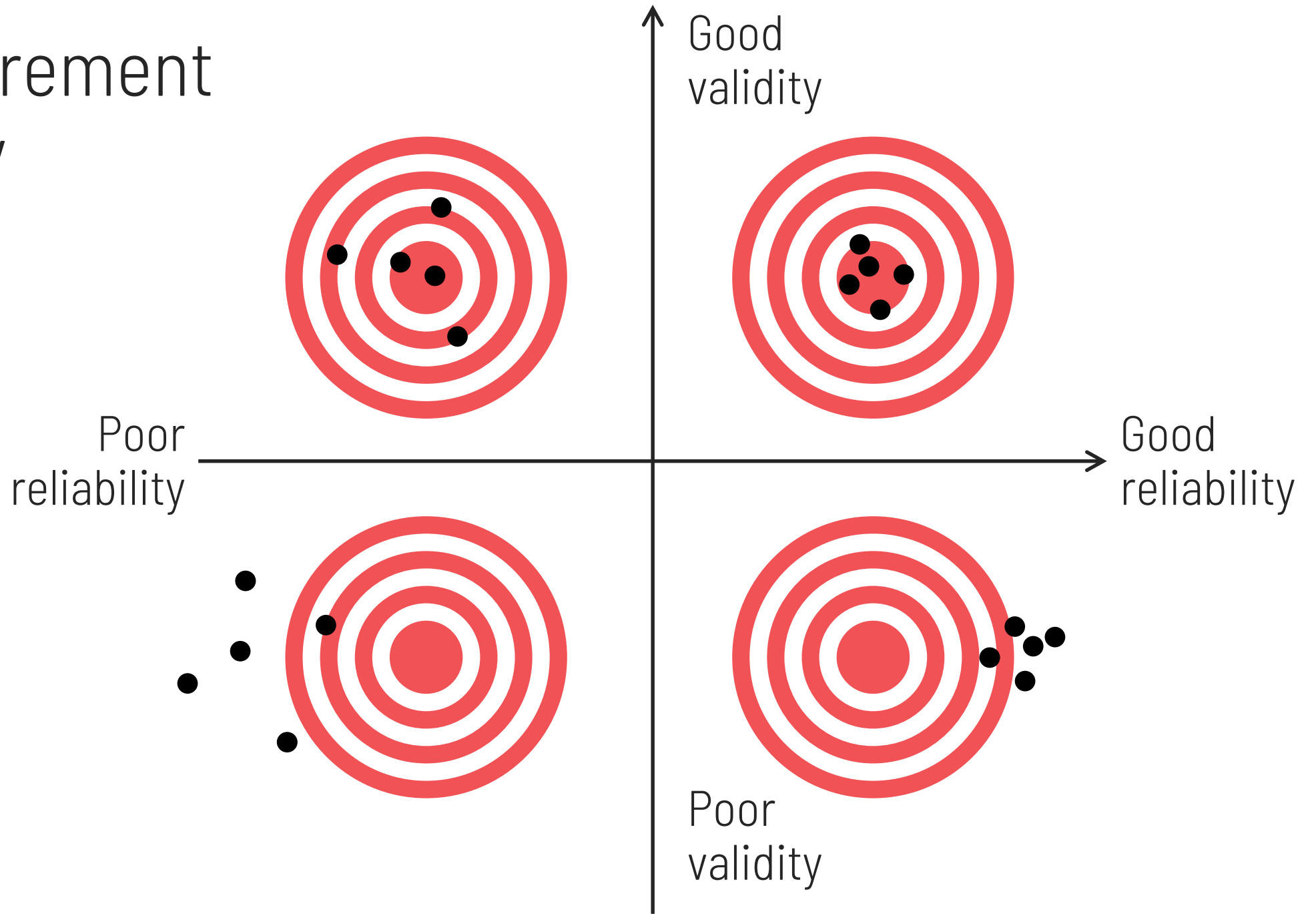
RELIABILITY

- ▶ Do people really rate the phenomenon and not something else?

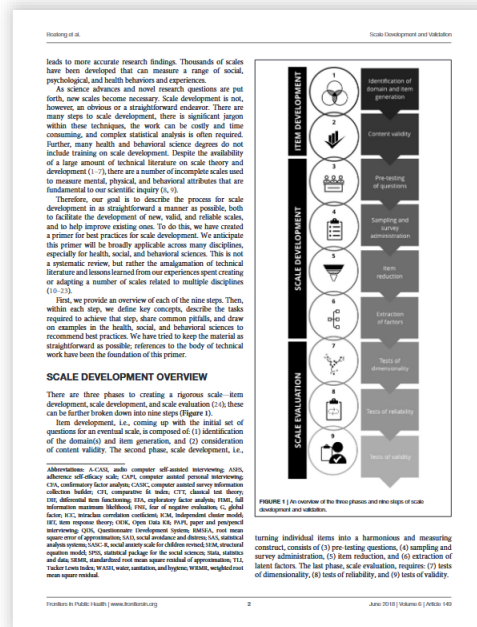
VALIDITY



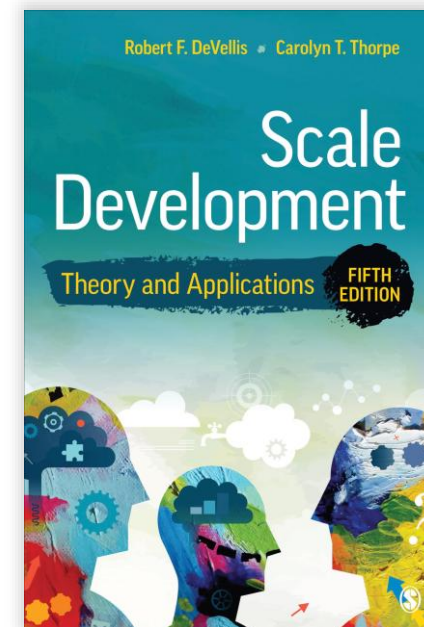
Measurement quality



Researchers in psychology, health, and social sciences have decades of experience to share



Boateng et al. (2018)
Best Practices for Developing and
Validating Scales: A Primer
doi:[10.3389/fpubh.2018.00149](https://doi.org/10.3389/fpubh.2018.00149)



DeVellis and Thorpe (2021)
Scale development:
Theory and Applications
ISBN: 9781544379340

Scale development steps

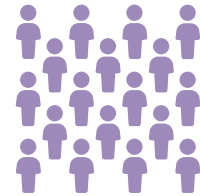


Define domain
Research what to ask

Scale development steps



Define domain
Research what to ask



Developing a scale

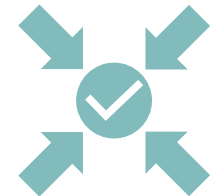
Scale development steps



Define domain
Research what to ask



Developing a scale

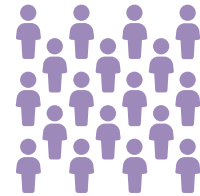


**Evaluating the
scale**

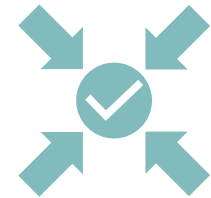
Scale development steps



Define domain
Research what to ask



Developing a scale



**Evaluating the
scale**

BeauVis Scale

To what extent do you agree that this visual representation is ... ?

	strongly disagree							strongly agree
enjoyable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
likable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
pleasing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
appealing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Recommended version

1. Define domain
Research what to ask



Identification of the domain

What are we trying to measure?

Osbert, Alphonse. (1908). *Soir Antique* [Oil on canvas]. Petit Palais, Paris.



Aesthetics

➔ Aesthetic Pleasure



Aesthetic Pleasure: Beauty

A pleasurable subjective experience that is directed toward an object and not mediated by intervening reasoning. [Reber et al., 2004]

Osbert, Alphonse. (1908). *Soir Antique* [Oil on canvas]. Petit Palais, Paris.



Aesthetic Pleasure: Beauty

not related to how understandable,
informative, or memorable it is

Item generation

**What could we ask to measure
our target variable?**

Literature review from VIS papers

Terms from 68 out of 3189 IEEE VIS, TVCG and CG&A papers



Spreadsheet for collecting terms

is and weaknesses of both images and rate each on a 5-point likert scale between poor and excellent. For each pair, participants also indicated which image was better for a task. The high image was more **aesthetic**, and which image was more useful. These questions were answered with one 5-point likert scale. The possible answers were *clearly left*, *rather left*, *neither*, *rather right*, *clearly right*.

[Jenny et al., 2021]

Statement	Mean	SD
The visualization was enjoyable	3.75 (2.09)	5.1
Using the visualization aid		

[Albo et al., 2016]

3.5 Aesthetic Requirements

Participants rated the display they were exposed to on semantic differential scales. Participants rated the display as predominantly **Attractive**, **Beautiful**, and **Interesting**, with no negative responses being listed in these categories. One participant rated the display they saw during the experiment as **Unattractive**.

[Rodgers and Bartram, 2011]

Selection criteria: (1) Ease/difficulty in understanding the visualization; (2) Ease/difficulty in using the visualization for comparison; (3) the **aesthetic appearance** of the visualization. For each visualization, we also asked for their general feedback.

Results

[Chen et al., 2021]

Results. We observed only two significant differences. Memex was rated faster in terms of performance time (average rating of 5.7 v. 4.8 for FacetMap) in terms of **aesthetic appeal**, FacetMap scored higher (average rating of 5.3 v. 4.1), $t(18)=1.9$, $p<0.05$. User satisfaction ratings are provided in Table 1.

[Smith et al., 2006]

Results. We asked the participants to rate each visualization on a 5-point scale for the following criteria: **redundant** – **informative**, **hindering** – **helpful**, **boring** – **entertaining**, **ugly** – **elegant**. The results are shown in Table 2. We asked the participants to rate each visualization on a 5-point scale for the following criteria: **redundant** – **informative**, **hindering** – **helpful**, **boring** – **entertaining**, **ugly** – **elegant**.

[Duncan et al., 2021]

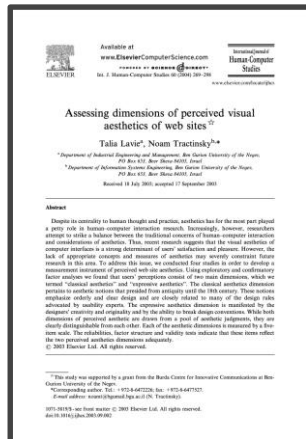
journal	title	doi link	filename	searchterm	likert term	adjective: To what extent do you agree that this visualization is ...	whether I saw the whole questionnaire(source code) only papers I found relevant now.	excluded term	participant feedback	participant feedback context	path
TVCG	Evaluating Cartogram Effectiveness	https://doi.org/10.1109/TVCG.2018.2642109	07792176.pdf	likert;aesthetic;	??? (5-point: entertaining ... boring); ??? (5-point: elegant ... drab); ??? (5-point: innovative ... conventional); ??? (5-point: easy to understand); ??? (5-point: showing magnitude clearly ... poorly)	entertaining;boring;elegant;drab;innovative;conventional;understandable;clear;			helpful/hindering;interested to use later;		..Vis-all_full_per_pdf-text-extraction-results/InfoVis-2020/11100a337.tex.xml
VIS	Chartem: Reviving Chart Images with Data Embedding	http://dx.doi.org/10.1109/TVCG.2020.3030351	111100a337	likert;aesthetic;	aesthetic appeal (7-point?);	aesthetic;	score the three streamgraphs based on their aesthetic preference and perceived readability;	no	like (the bridge metaphor);		..Vis-all_full_per_pdf-text-extraction-results/InfoVis-2020/06_infvis_smith.tex.xml
VIS	FacetMap: A Scalable Search and Browse Visualization	http://dx.doi.org/10.1109/TVCG.2008.142	06_infvis_smith	questionnaire;aesthetic;	aesthetic appearance (5-point: nice, more or less nice, neutral, somewhat ugly, ugly);	aesthetic;nice;ugly;	For the aesthetic appearance, the users preferred Co-Bridges more.	no			..Vis-all_full_per_pdf-text-extraction-results/InfoVis-2020/111100b612.tex.xml
VIS	Co-Bridges: Pair-wise Visual Connection and Comparison for Multi-item Data Streams	http://dx.doi.org/10.1109/TVCG.2020.3030411	111100b612	likert;questionnaire;interview;aesthetic;	aesthetic preference (7-point: totally disagree ... neutral ... totally agree);readability (7-point: totally disagree ... neutral ... totally agree);	aesthetic;readable;					..Vis-all_full_per_pdf-text-extraction-results/InfoVis-2020/111100b634.tex.xml
VIS	SineStream: Improving the Readability of Streamgraphs by Minimizing Sine Illusion Effects	http://dx.doi.org/10.1109/TVCG.2020.3030404	111100b634	likert;questionnaire;aesthetic;	aesthetic preferences(7-point: strongly preferred ...);	aesthetic;	streamgraph;				..Vis-all_full_per_pdf-text-extraction-results/InfoVis-2020/251_251vis1214_ucho-2345426.tex.xml
VIS	The Influence of Contour on Similarity Perception of Star Glyphs	http://dx.doi.org/10.1109/TVCG.2014.2345426	2251_20hvg12-fuchs-2345426	likert;questionnaire;aesthetic;		aesthetic;well-designed;					..Vis-all_full_per_pdf-text-extraction-results/InfoVis-2020/1313_vast_zhang.tex.xml
VIS	Vis4Heritage: Visual Analytics Approach on Circle Wall Painting Degradations	http://dx.doi.org/10.1109/TVCG.2013.219	13_vast_zhang	questionnaire;aesthetic;	aesthetic;visual design; aestheticness (5-point: worst ... best);expressiveness (5-point: worst ... best);understandability (5-point: worst ... best);	aesthetic;expressive;understandable;	tool;				..Vis-all_full_per_pdf-text-extraction-results/InfoVis-2020/11100a453.tex.xml
VIS	Callope: Automatic Visual Data Story Generation from a Spreadsheet	http://dx.doi.org/10.1109/TVCG.2020.3030403	111100a453	likert;questionnaire;interview;	aesthetics (11-point: 0 ... 10);visual design (11-point: 0 ... 10);	aesthetic;well-designed;	visualization - design of graphical user interface and visualization/visual design - of three views were suitable for analyzing domain tasks;		satisfied;nice design;thoughtful design;		..Vis-all_full_per_pdf-text-extraction-results/InfoVis-2020/11100a453.tex.xml
VIS	Vismate: Interactive Visual Analysis of Station-Based Observation Data on Climate Changes	http://dx.doi.org/10.1109/AST.2014.7042489	14_ast_7042489	questionnaire;aesthetic;	aesthetic;						..Vis-all_full_per_pdf-text-extraction-results/InfoVis-2020/11100a453.tex.xml
---	Augmented Reality Graph	https://doi.org/10.1109/M	---	questionnaire;aesthetic;	aesthetics (5-point: ...);	aesthetic;					..Vis-all_full_per_pdf-text-extraction-results/InfoVis-2020/11100a453.tex.xml

Literature from related field

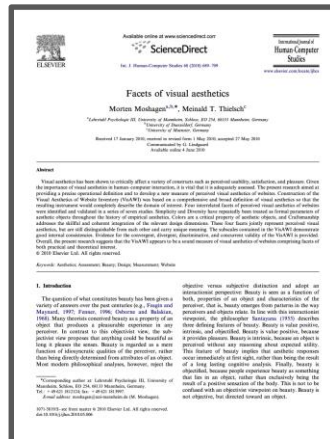
Terms from 4 aesthetics-related scales development papers



Spreadsheet for collecting terms



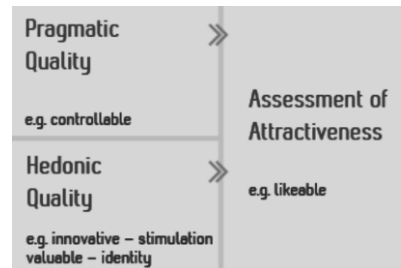
[Lavie & Tractinsky, 2003]



[Moshagen & Thielsch, 2010]



[Blijlevens et al., 2017]



[Hassenzahl et al., 2003]

	AttrakDiff	Blijlevens, 2017	Lavie, 2003	Moshagen, 2010
aesthetic			aesthetic	
appealing	repelling/appealing	appealing		appealing
attractive	attractive	attractive		attractively positioned; attractive; balanced; there are too many elements in o
balanced				
beautiful		beautiful	beautiful	
clean		clean	clean	
cluttered				
creative	creative	creative	creative	creative
elegant				
harmonious			harmonious	
inviting	inviting	inviting		
modern			modern	modern; contemporary
nice		nice		
organized			organized	pleasantly organized
overloaded		overloaded		overloaded
pretty		pretty		
tasteful			applies good taste	
well-designed		designed	skillfully designed	
artistic			artistic	
boring	dull		dull	boring
delightful		delightful		
engaging				
enjoyable			enjoyable	
entertaining				
exciting		exciting	exciting	
fascinating		fascinating	fascinating	
good		good		
color-harmonious				
interesting		interesting		interesting; lacks interesting design details
likable	likable	like		
motivating	motivating	motivating		
pleasing	pleasant	pleasant; pleasurable	pleasing; pleasant	a pleasant effect
professional	professional/unprofessional	professional	professional	professional
provoking				
satisfying				
sophisticated		sophisticated	sophisticated	
lovely				
dynamic		dynamic		dynamic
crowded		densely		crowded; too many elements
drab			monotonous	monotonous
high-quality		confers quality		
stylish	stylish			
well-proportioned				well proportioned
informative				
colorful			colorful	colorful; too few colors
eye-catching				

Expert suggestions

Invitation email sent to
57 visualization experts



Survey for collecting terms
(31 responses)

Subject: Survey Invitation - How to Judge the Aesthetics of Visualizations?

To: [email of an expert in visualization]

Dear [expert's name],

We are currently working on a research project about generating a validated scale for rating the aesthetics of a visualization. An important step in the generation of a scale is to elicit terms from experts. Given your expert status in our domain we would much appreciate it if you could spend 2-3 minutes of your time and fill out our short survey.

To participate, please access the survey here:[survey link]

Please notice that this study has a two-stage evaluation process. After this survey, we would like to contact you again for a second very short survey.

We thank you a lot in advance and would be happy to share the results of our work with you if you are interested! Just let us know.

Best regards,
Tingying He, Petra Isenberg, Raimund Dachzelt, and Tobias Isenberg

You decided that you want to ask people to rate the visualization using a 7-point scale like the one below.



Your task is to find good terms with which to capture this subjective opinion of participants about the **aesthetics** of the visualization. This means: you really only care about what the visualization looks like and NOT about how well people understand the data that it visualizes. So you decide to ask:

To what extent do you agree or disagree with the following statement:
This visualization is _____.

Which terms would you put in the blank? Give us as many alternative terms as you can think of, but please give us **at least 3 terms related to aesthetics**.

This is the last question in this survey. Please take at least a minute to think about your answer before clicking "Next". Thank you.

Please fill in at least 3 answers

Term 1	<input type="text"/>
Term 2	<input type="text"/>
Term 3	<input type="text"/>

209

terms

Content validity

What is the face validity of our candidate items?

209

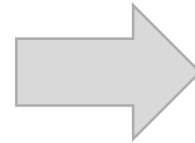
terms

aesthetic	emotion	cognitive	data-aesthetic
<i>a poor visual focus</i>	alienating	<i>a poor visual focus</i>	<i>expressive</i>
aesthetic	<i>appealing</i>	appropriate	<i>informative</i>
<i>appealing</i>	appreciating	attention-catching	suitable
artistic	averageness	categorizable	
asymmetrical	awe	challenging	
attractive	boring	clear	other
awesome	bring me closer to people/separates me from people	<i>cluttered</i>	a printing effect
balanced	<i>calm</i>	compelling	admirable
beautiful	comfortable	comprehensible	alive
bold	connective	conceptless	amateurish
<i>calm</i>	cool	confusing	bad
captivating	delightful	contemplative	botched
cautious	desirable	cumbersome	cheap
clean	disagreeable	easy to grasp	colorblind-safe
<i>cluttered</i>	disturbing	elicits associations	consistent
color-harmonious	dynamic	<i>informative</i>	convenient
colorful	elation	inspiring	convenient
complex	emotive	interpretable	<i>conventional</i>
conservative	energetic	intuitive	easy on eyes
contrastful	engaging	meaningful	easy orientation
<i>conventional</i>	enjoyable	memorable	easy to navigate
creative	entertaining	practical	easy to use
crisp	evocative	readable	faunist
crowded	evoking feelings	slick	fit together
discouraging	exciting	stimulating creativity	flowing
distinctive	fascinating	stimulating curiosity	fluent to process
drab	favorable	straightforward	good
elegant	fun	structured	hectic
<i>expressive</i>	gratifying	undemanding	<i>high-quality</i>
eye-catching	happy	understandable	human
familiar	hideous	<i>use of color is successful</i>	<i>innovative</i>
geometric	integrating		it is possible to discover new things even when looking at the page for a longer time.
harmonious	intense		manageable
has enough free space	interesting		noisy
<i>high-quality</i>	intriguing		one-sided
illuminating	intrusive		pleasantly animated
<i>innovative</i>	isolating		premium
inventive	likable		professional
inviting	motivating		restless
just eye-candy	moved		romantic
lack imagination	perfection		shows complete ignorance of human visual perception
looks great, but does not enable to get insight	pleasing		some elements seem out of place
lovely	positive		sophisticated
made with care	powerful		static
modern	predictable		stucco
nice	preferable		technology
novel	provoking		the control instructions are too static
old-fashioned	relaxed		the number of images is adequate
orderly	satisfying		the page contains too much text
ordinary	stimulating		too little happens on the page
organized	striking		unique
original	sublime		unruly
overloaded	the page changes too little due to user actions		uses special effects
painterly	thrills or chills		varied
patchy	touched		versatile
presentable	touched		well-combined
pretty	warm feeling		well-finished
realistic appearance			wretched
rejecting			
simple			
streamlined			
stunning			
stylish			
symmetrical			
tacky			
tasteful			
thoughtful			
thrown together			
ugly			
unimaginative			
unique			
up-to-date			
<i>use of color is successful</i>			
vulgar			
well-crafted			
well-designed			
well-proportioned			

Filtering on occurrence and semantics

6 criteria we applied

1. The terms needed to be **related to *aesthetic pleasure*** rather than *understanding* or *comprehension* of a visual representation or its data (e. g., we excluded “informative,” “clear,” or “confusing”).
2. The terms had to have **appeared at least twice** in one of the three resources we used for our item generation: visualization papers, other relevant aesthetics scale papers, or expert suggestions.
3. The terms should be **usable in a rating scale** and have a **clearly good or bad connotation** (e. g., we excluded “complex” because a complex aesthetic could be seen as positive or as negative).
4. The terms should be **easy to understand** (e. g., we excluded “consistent” because it would be unclear according to what aspect a visual appearance would be consistent) and their **interpretation should be clear** (e. g., we excluded “novel” because it would require people to know what “old” visualizations look like; we also excluded “drab” as a rare term that is not easily understood by many non-native speakers of English).
5. The terms had to **clearly apply to an assessment of a visual representation** (e. g., we excluded “dynamic” because, within visualization, the term may be read as referring to the property of being animated or interactive, rather than a dynamic aesthetic).
6. The terms should **not be pairs of opposite adjectives**. We only retained negative terms that did not have a clear positive opposite (e. g., we excluded “ugly” as the opposite of “beautiful”).



37
terms

Expert review

Invitation email sent to 56 visualization experts



Survey for collecting terms (25 responses)

Subject: Invitation for new short 4min survey - Terms to judge the aesthetics of a visualization

To: [\[email of an expert in visualization\]](#)

Dear [\[expert's name\]](#),

You have previously received an email from us about a first quick survey regarding how to judge the aesthetics of a visualization. If you had a chance to participate, thank you very much! We received a lot of useful input and comments that we will address! If not - don't worry - you still have a chance to participate in this second survey.

To clarify, our project is about developing a simple instrument to gauge the aesthetic pleasure of a visualization – meant to provide a few simple rating questions that can accompany other types of experiments (quantitative or qualitative).

In the first phase of our work we asked you to provide a few terms that you consider to be usable in an aesthetic rating. In addition to terms provided by experts like you, we have also assessed the literature and come up with a final list of 37 terms; narrowed down from a list of > 200 terms. An important second phase in scale development is to ask experts to rate the appropriateness of the terms we collected. As such, we would much appreciate it if you could spend around 3-4 minutes of your time to fill out our second survey. We hope that at least as a small reward seeing the list of terms may already be useful or inspiring to you.

To participate, please access the survey here: [\[survey link\]](#)

We thank you a lot in advance and would be happy to share the results of our work with you if you are interested! Just let us know.

Best regards,
Tingying He, Petra Isenberg, Raimund Dachsel, and Tobias Isenberg

The aesthetic pleasure of visualization is the pleasure people derive from looking at a visualization for its own sake, as a source of immediate experiential pleasure in itself, and not essentially for its utility in producing insight or knowledge gain or something else that is either useful or pleasurable.

The table below includes terms that have been suggested or used in the literature by visualization experts like you for studying the aesthetic pleasure of a visualization. Imagine that these terms would later be used in a rating scale that asks participants to select to what extent a visualization is ..[term].

Below we would like you to rate these different terms according to how relevant you consider them for actually judging the aesthetic pleasure of a visualization. Please note that we only care about aesthetic pleasure in terms of what a visualization looks like and not how well people can comprehend the data that it shows.

How relevant do you think the following terms are for judging or describing the aesthetic pleasure of a visualization?

	1 = not at all relevant	2	3	4	5 = very relevant
sophisticated	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
beautiful	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
appealing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
likable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
cluttered	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
enjoyable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tasteful	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
modern	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
aesthetic	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
clean	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
color-harmonious	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
boring	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
satisfying	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
delightful	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
entertaining	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
exciting	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

31

terms

aesthetic	emotion	cognitive	other
<i>appealing</i>	<i>appealing</i>	<i>cluttered</i>	professional
artistic	delightful		sophisticated
attractive	engaging		
balanced	enjoyable		
beautiful	exciting		
clean	fascinating		
<i>cluttered</i>	interesting		
color-harmonious	likable		
creative	motivating		
elegant	pleasing		
harmonious	provoking		
inviting	satisfying		
lovely			
nice			
organized			
pretty			
tasteful			
well-designed			

2. Developing the scale



Sampling and survey administration

How to establish a representative sample for data collection?

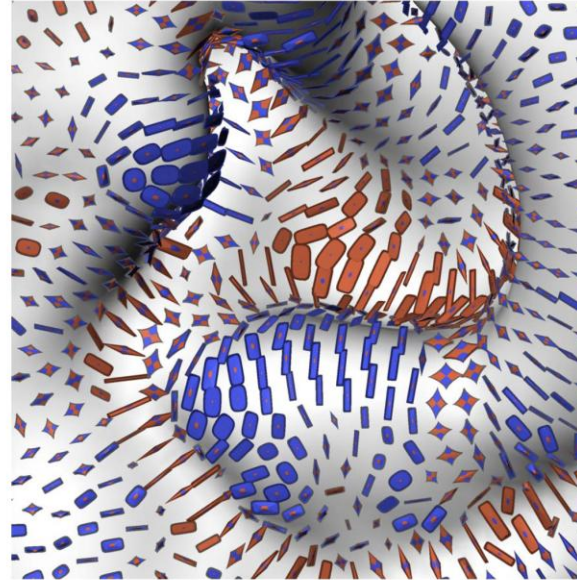
→ what visualizations should we test this on?

Crowdsourced Experiment

- ▶ 1001 participants
- ▶ 15 data representations
- ▶ Each participant rated 3 representations

★ To what extent do you agree or disagree with the following statement:

The visualization is ____.



	Strongly disagree	Disagree	Slightly disagree	Neutral	Slightly agree	Agree	Strongly agree
motivating	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
provoking	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
organized	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
engaging	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
creative	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
clean	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
aesthetic	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
beautiful	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
pretty	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
pleasing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
fascinating	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
elegant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Strongly		Slightly		Slightly		Strongly

Previous

Next

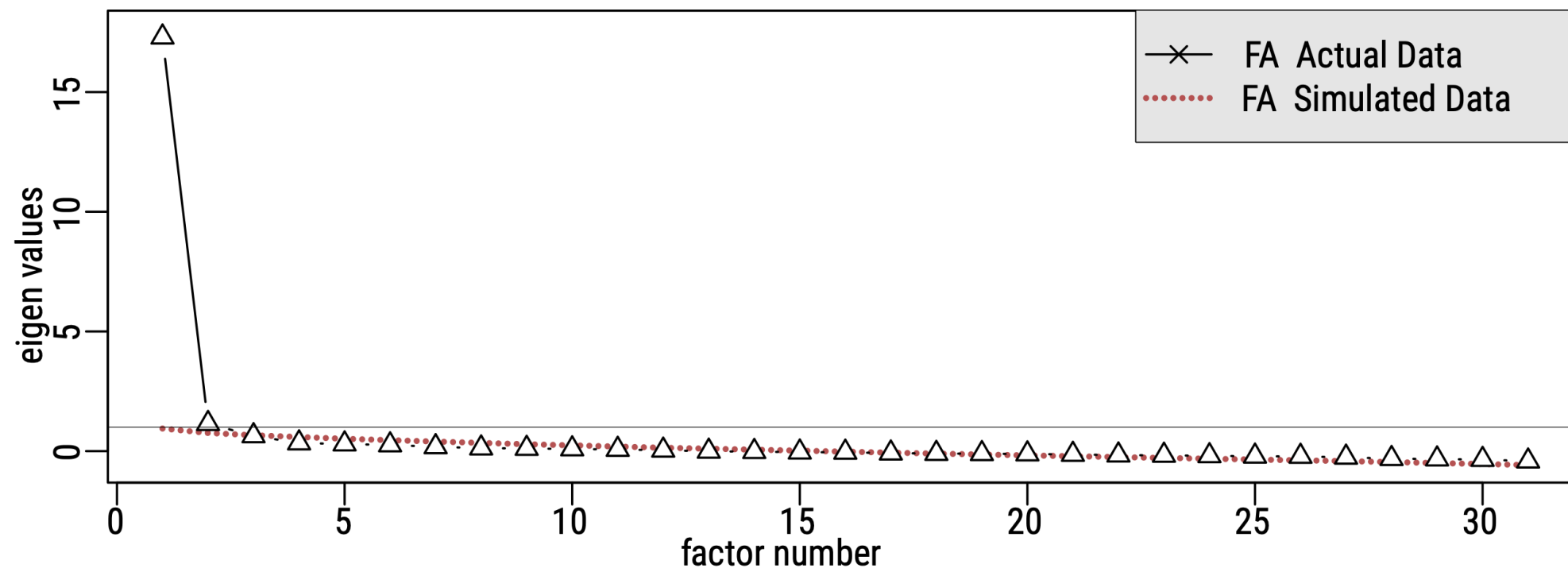
Exploratory experiment screenshot

Extraction of factors
**Identifying components
of aesthetic pleasure**

Potential factors

aesthetic	emotion	cognitive	other
<i>appealing</i>	<i>appealing</i>	<i>cluttered</i>	professional
artistic	delightful		sophisticated
attractive	engaging		
balanced	enjoyable		
beautiful	exciting		
clean	fascinating		
<i>cluttered</i>	interesting		
color-harmonious	likable		
creative	motivating		
elegant	pleasing		
harmonious	provoking		
inviting	satisfying		
lovely			
nice			
organized			
pretty			
tasteful			
well-designed			

Parallel Analysis shows 1 factor



Scree plot for Image 1 (3D surface glyphs), see our paper for details

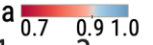
Item reduction

**Which candidate items
should we eliminate?**

Reliability: Cronbach's Alpha

Alpha > 0.7 : Reliable

[Boateng et al., 2018]

		alpha 																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
Enjoyable Likable Pleasing	enjoyable-likable-pleasing	0.92	0.86	0.89	0.91	0.91	0.90	0.94	0.92	0.88	0.92	0.91	0.93	0.91	0.92	0.93	0.91	0.91
	enjoyable-likable-nice	0.93	0.87	0.90	0.90	0.91	0.89	0.93	0.92	0.88	0.91	0.91	0.92	0.91	0.92	0.93	0.91	
	likable-nice-pleasing	0.93	0.87	0.88	0.90	0.92	0.90	0.93	0.91	0.86	0.91	0.91	0.92	0.91	0.91	0.92	0.91	
Enjoyable Likable Pleasing Nice	enjoyable-likable-pleasing-nice	0.94	0.90	0.91	0.92	0.93	0.92	0.95	0.94	0.90	0.93	0.93	0.94	0.93	0.93	0.95	0.93	0.93
	enjoyable-likable-appealing-pleasing	0.94	0.89	0.91	0.93	0.93	0.92	0.95	0.94	0.91	0.94	0.92	0.94	0.93	0.93	0.94	0.93	
	enjoyable-likable-appealing-nice	0.94	0.90	0.91	0.92	0.93	0.92	0.95	0.94	0.91	0.93	0.92	0.94	0.93	0.93	0.95	0.93	
Enjoyable Likable Pleasing Nice Appealing	enjoyable-likable-nice-pleasing-appealing	0.95	0.92	0.92	0.94	0.94	0.94	0.96	0.95	0.92	0.94	0.94	0.95	0.95	0.94	0.96	0.94	0.94
	appealing-nice-enjoyable-likable-pleasing	0.94	0.91	0.92	0.94	0.94	0.93	0.96	0.94	0.92	0.95	0.94	0.95	0.94	0.94	0.95	0.94	
	attractive-enjoyable-likable-nice-pleasing	0.95	0.91	0.92	0.93	0.94	0.94	0.96	0.94	0.92	0.94	0.94	0.95	0.94	0.94	0.95	0.94	

Cronbach's alpha for each image on the most reliable 3-, 4-, and 5-item subsets

BeauVis Scale

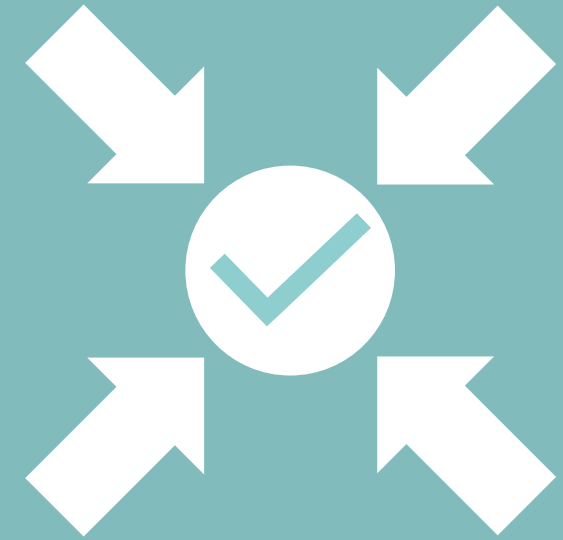
To what extent do you agree that this visual representation is ... ?

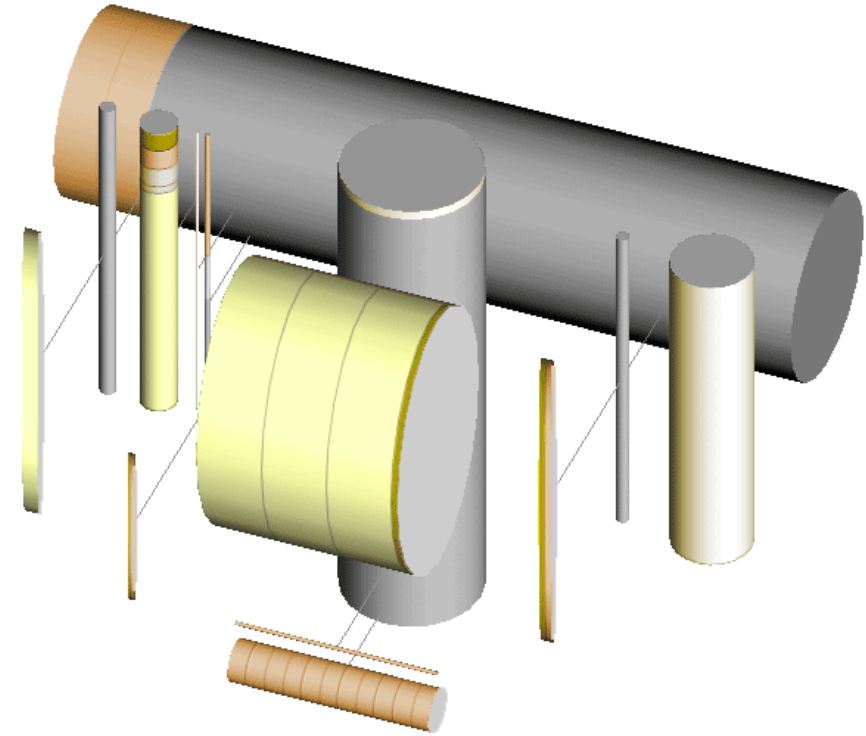
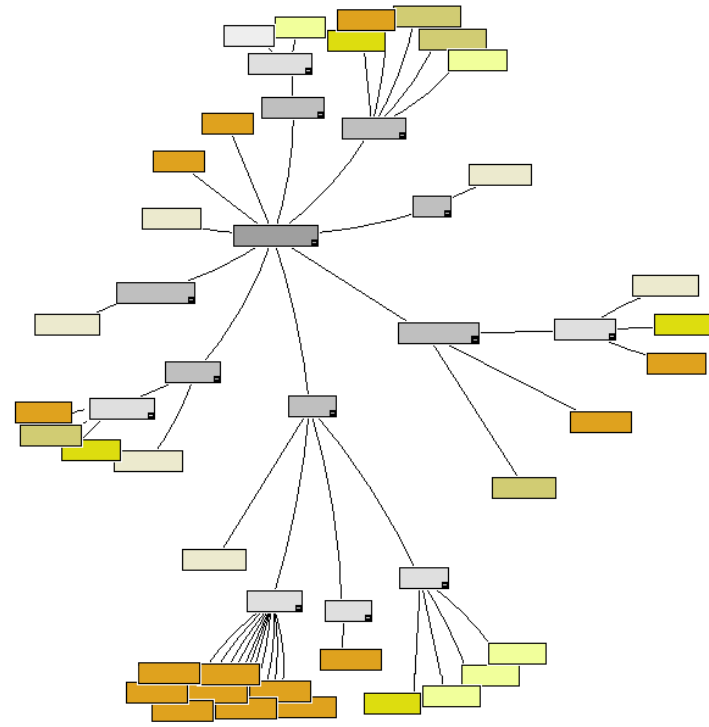
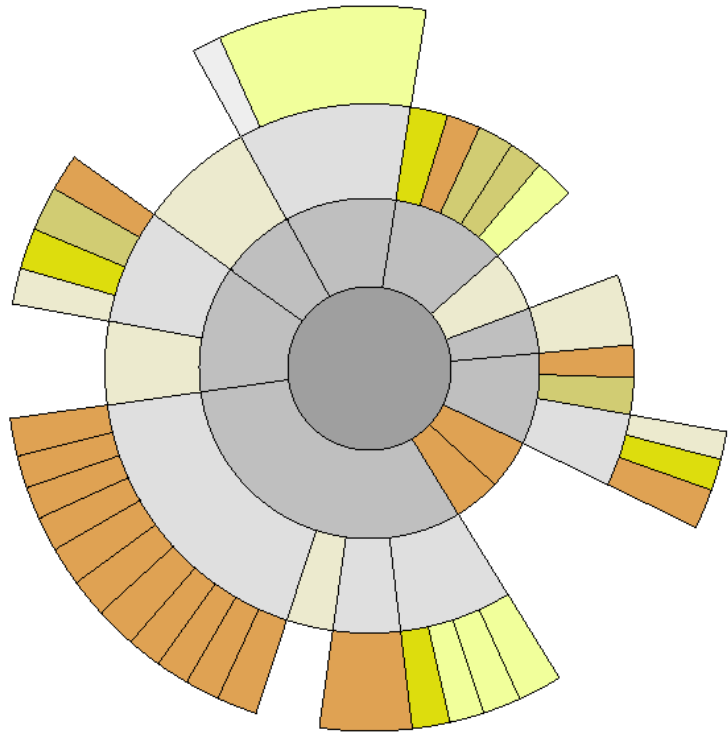
strongly disagree strongly agree

enjoyable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
likable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
pleasing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
appealing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

BeauVis scale in its recommended version

3. Evaluating the scale





201 Prolific participants

Good news

- ▶ The scale still had only 1 factor
- ▶ The results were still reliable
- ▶ We have convergent validity

Convergent Validity

	SunBurst	StarTree	BeamTree	mTree
Cronbach's Alpha	0.84	0.88	0.87	.95

Pearson correlation

How should you use BeauVis

- ▶ Ideally at the start of an experiment
- ▶ Give all 5 items, calculate the average
- ▶ Use it in conjunction with other measures

Please answer the following questions regarding the visualization used in this part of the experiment:

To what extent do you agree that this visualization is ... ?

	Strongly Disagree						Strongly Agree
Enjoyable	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Likable	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pleasing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Nice	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Appealing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Progress bar: 40%

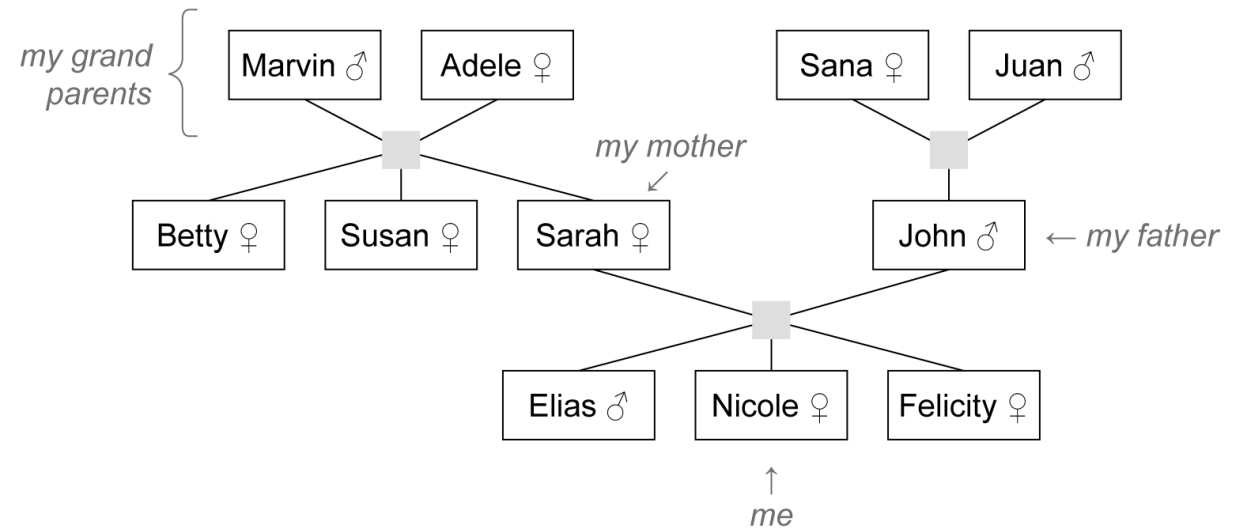
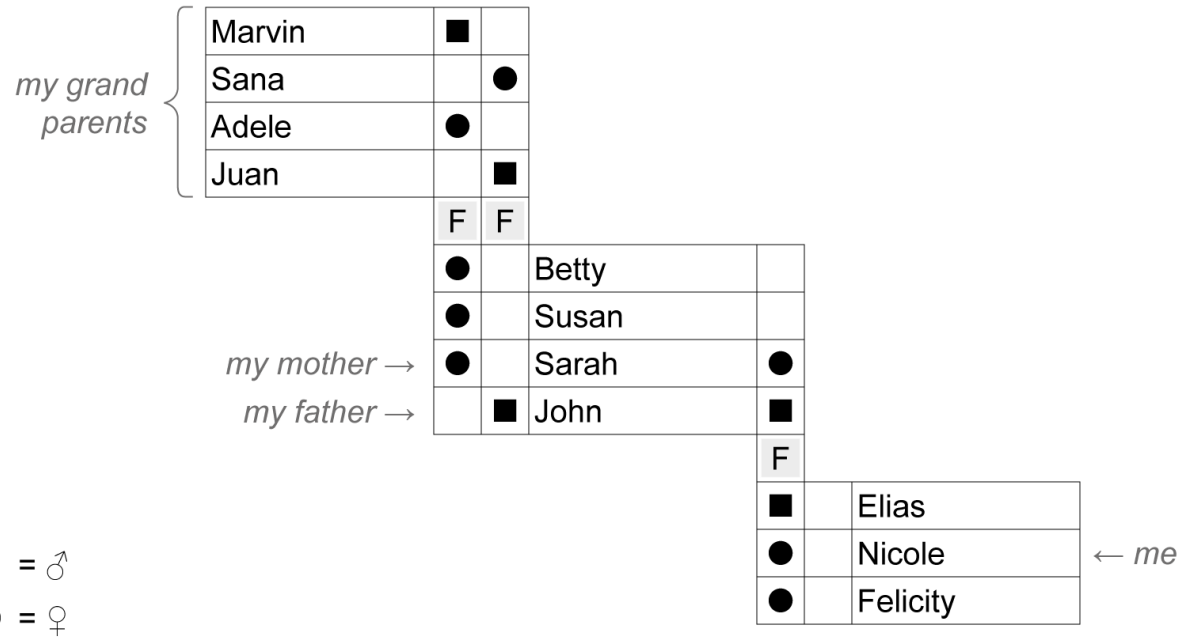
PREVis

Perceived Readability Evaluation for Visualizations



Anne-Flore Cabouat, Tingying He, Petra Isenberg, Tobias Isenberg
Université Paris-Saclay, CNRS, Inria, LISN

... readable?



PREVis: 11 questions across 4 scales

- ◆ **Understand** obvious + represent + understand
- ◆ **Layout** messy + crowded + distracting
- ◆ **DataRead** find + identify + inform
- ◆ **DataFeat** visible + clear

Strongly
disagree

Disagree

Slightly
disagree

Neutral

Slightly
agree

Agree

Strongly
agree

I don't know / Not applicable (if you wish, feel free to elaborate below)

PREVis: 11 questions across 4 scales

◆ Understand obvious + represent + understand

- ◆ It is **obvious** for me how to read this visualization
- ◆ I can easily understand **how the data is represented** in this visualization
- ◆ I can **easily understand** this visualization

Strongly
disagree

Disagree

Slightly
disagree

Neutral

Slightly
agree

Agree

Strongly
agree



I don't know / Not applicable (if you wish, feel free to elaborate below)

PREVis: 11 questions across 4 scales

◆ **Layout** messy + crowded + distracting

- ◆ I **don't** find this visualization **messy**
- ◆ I **don't** find this visualization **crowded**
- ◆ I **don't** find **distracting parts** in this visualization

Strongly
disagree

Disagree

Slightly
disagree

Neutral

Slightly
agree

Agree

Strongly
agree

I don't know / Not applicable (if you wish, feel free to elaborate below)

PREVis: 11 questions across 4 scales

◆ **DataRead** *find + identify + inform*

- ◆ I can easily **find specific elements** in this visualization
- ◆ I can easily **identify relevant information** in this visualization
- ◆ I can easily **retrieve information** in this visualization

Strongly
disagree

Disagree

Slightly
disagree

Neutral

Slightly
agree

Agree

Strongly
agree



I don't know / Not applicable (if you wish, feel free to elaborate below)

PREVis: 11 questions across 4 scales

◆ DataFeat *visible + clear*

- ◆ I find data features (for example, a minimum, or an outlier, or a trend) **visible** in this visualization
- ◆ I can **clearly see** data features (for example, a minimum, or an outlier, or a trend) in this visualization

Strongly
disagree

Disagree

Slightly
disagree

Neutral

Slightly
agree

Agree

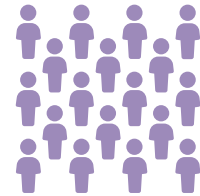
Strongly
agree

I don't know / Not applicable (if you wish, feel free to elaborate below)

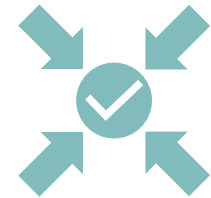
Scale development steps



Creating candidate questions



Developing a questionnaire



Evaluating the questionnaire

1. Creating candidate questions



Domain definition

Candidate questions generation & validation

Text readability

a

Text readability

thanks

a

[Letters

Text readability

So long and thanks
for all the fish.

a

thanks



Letters

Words

Text readability

The last ever dolphin message was misinterpreted as a surprisingly sophisticated attempt to do a double-backwards-somersault through a hoop whilst whistling the 'Star Spangled Banner,' but in fact the message was this: 'So long and thanks for all the fish.'

a

thanks

So long and
thanks for all
the fish.

Letters

Words

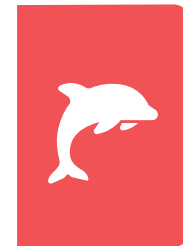
Sentences

Text readability

a

thanks

So long and
thanks for all
the fish.



Paragraphs,
chapters,
books!

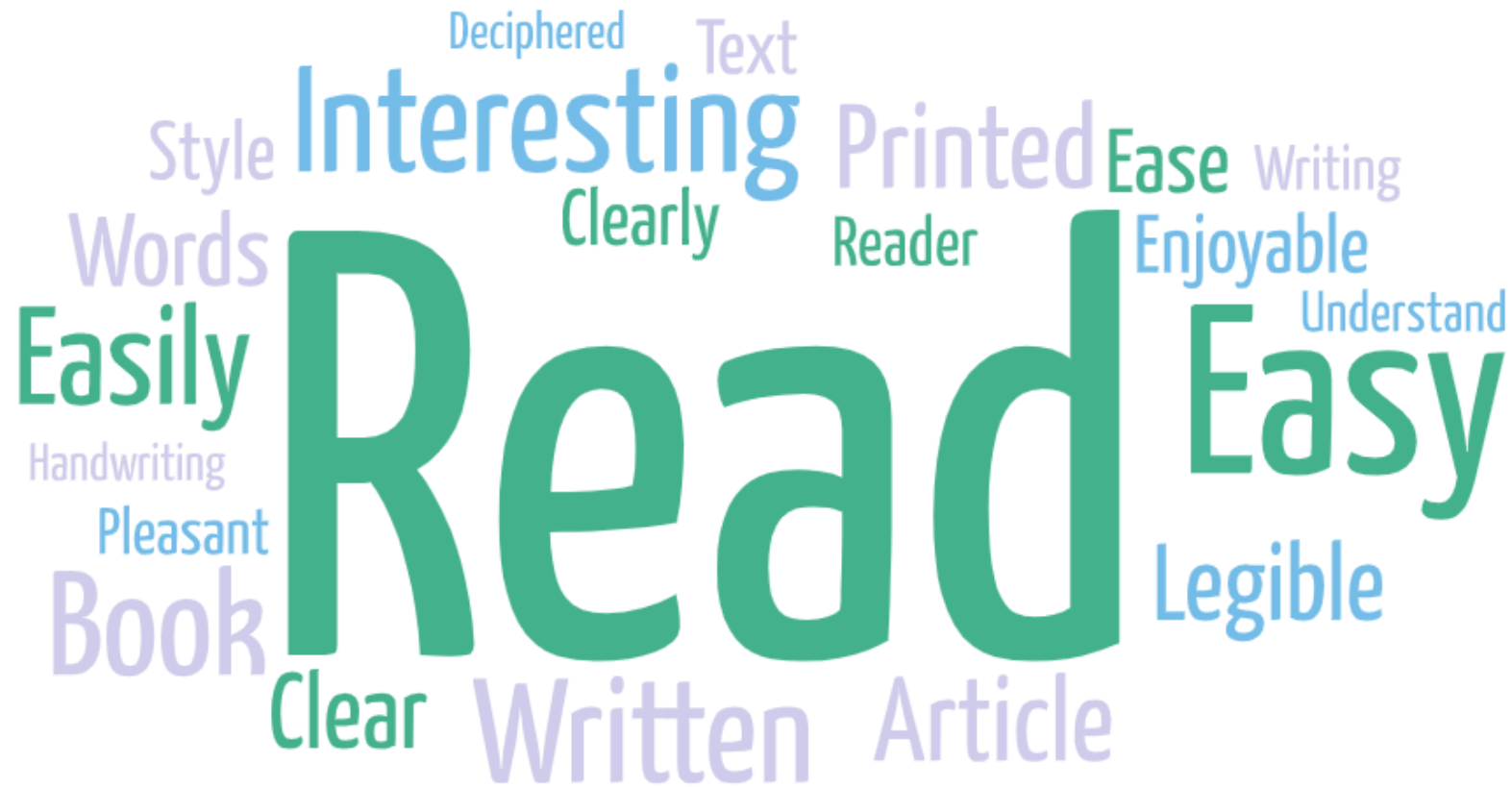
The last ever dolphin message was misinterpreted as a surprisingly sophisticated attempt to do a double-backwards-somersault through a hoop whilst whistling the 'Star Spangled Banner,' but in fact the message was this: 'So long and thanks for all the fish.'

Letters

Words

Sentences

Easy to read



Visualization readability?

“The relative ease with which the user finds the information he is looking for.” [Ghoniem et al.]

“The ability to make direct observations from the visualizations.” [Ruchikachorn and Mueller]

“The extent to which a visualization supports the graphical perception of the information it contains.” [Thudt et al., 2016]

“the visual decoding of information encoded on graphs” [Cleveland and McGill, 1984]

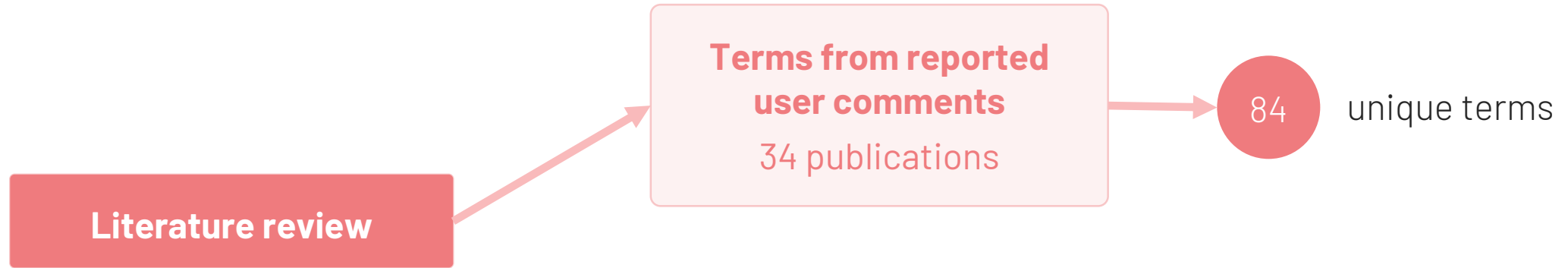
Item generation: deductive method

Literature review

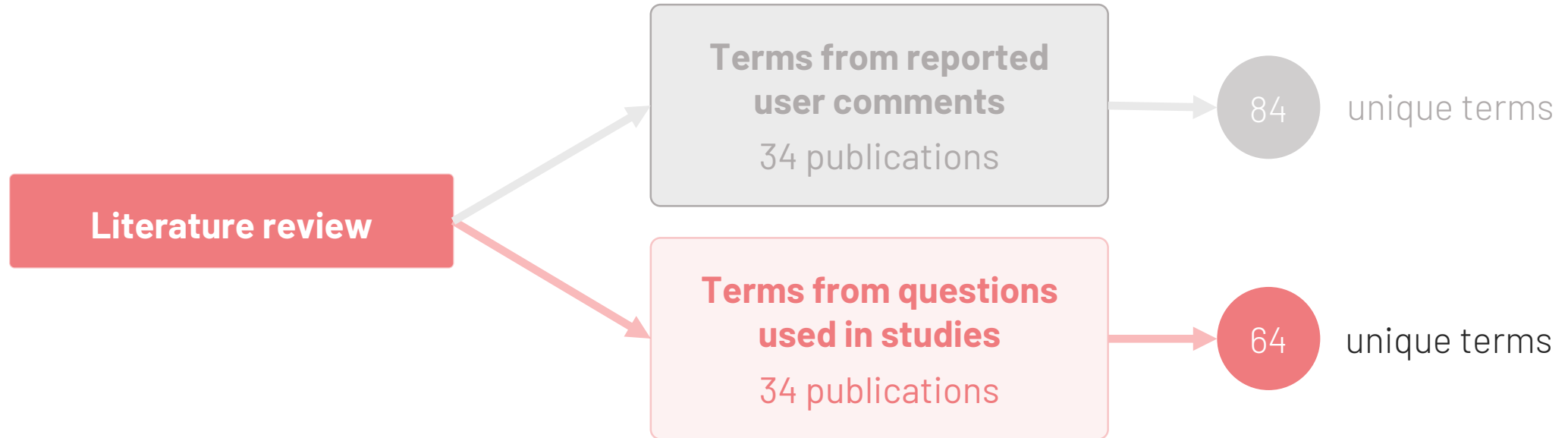
```
"likert" AND ("readab" OR "legib" OR "deciph")
```

IEEE VIS papers (1990-2022) + TVCG and CG&A journal
papers presented at IEEE VIS (2011-2021)

Item generation: deductive method



Item generation: deductive method



Not just different terms!

This graph has good readability

The visualization looked cluttered

It was easy to gain a good overview of the data

It was easy to see the link between two components

Is it clear what information is being presented and why?

Which of these charts is quicker and easier to read?

Visualization quality vs user activity

This graph has good readability

The visualization looked cluttered

It was easy to gain a good overview of the data

It was easy to see the link between two components

Is it clear what information is being presented and why?

Which of these charts is quicker and easier to read?

Different objects

This **graph** has good readability

The **visualization** looked cluttered

It was easy to gain a good overview of **the data**

It was easy to see the **link between two components**

Is it clear what **information** is being presented and why?

Which of these **charts** is quicker and easier to read?

Different granularities: high-level vs low-level

This graph has good readability

The visualization looked cluttered

It was easy to gain a good overview of the data

It was easy to see the link between two components

Is it clear what information is being presented and why?

Which of these charts is quicker and easier to read?

Item generation: inductive method

Literature review

Item generation: inductive method



Merging inductive and deductive methods

unique terms from
user comments

84

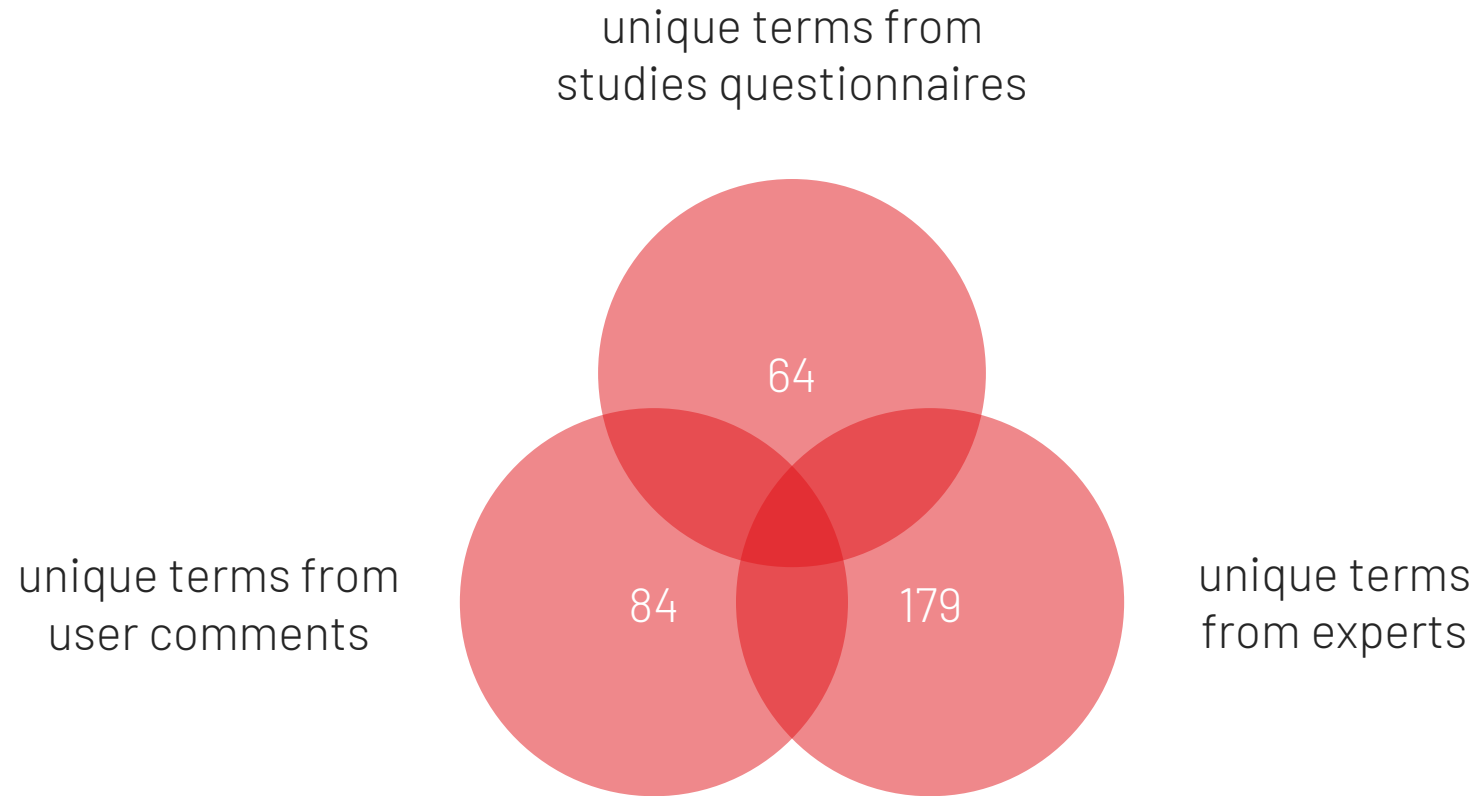
unique terms from
studies questionnaires

64

unique terms
from experts

179

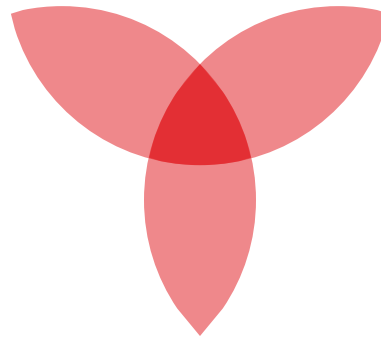
Merging inductive and deductive methods



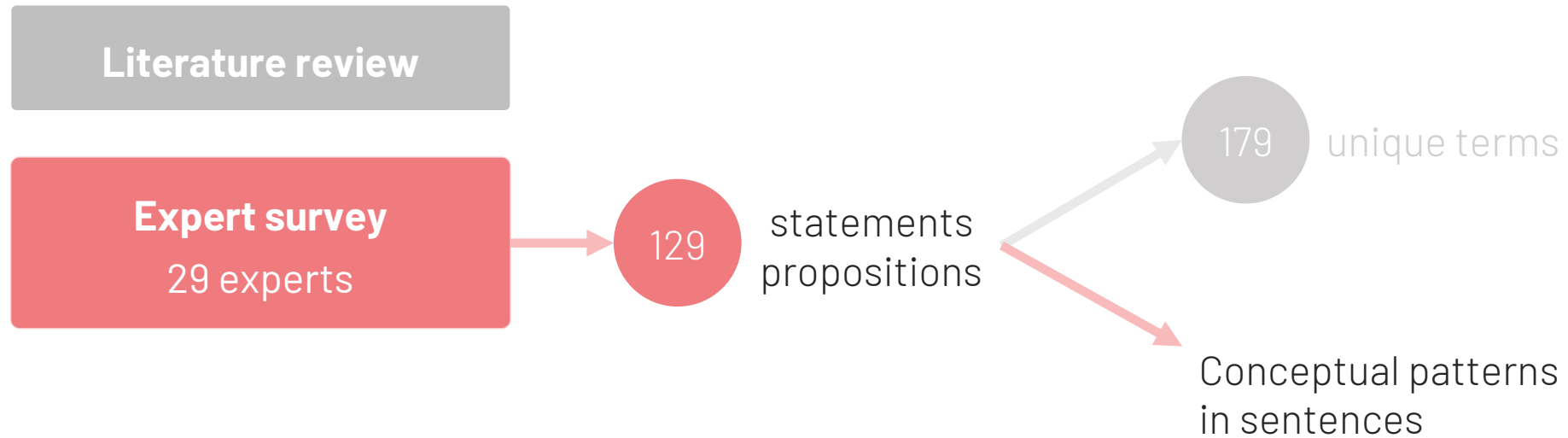
Merging inductive and deductive methods

59

terms common
to at least 2 pools



Item generation: phrasing patterns



Conceptual patterns to help write items from terms

clear

What kind of objects are referred to as being “clear” (e.g., data, patterns, information...)?

understand

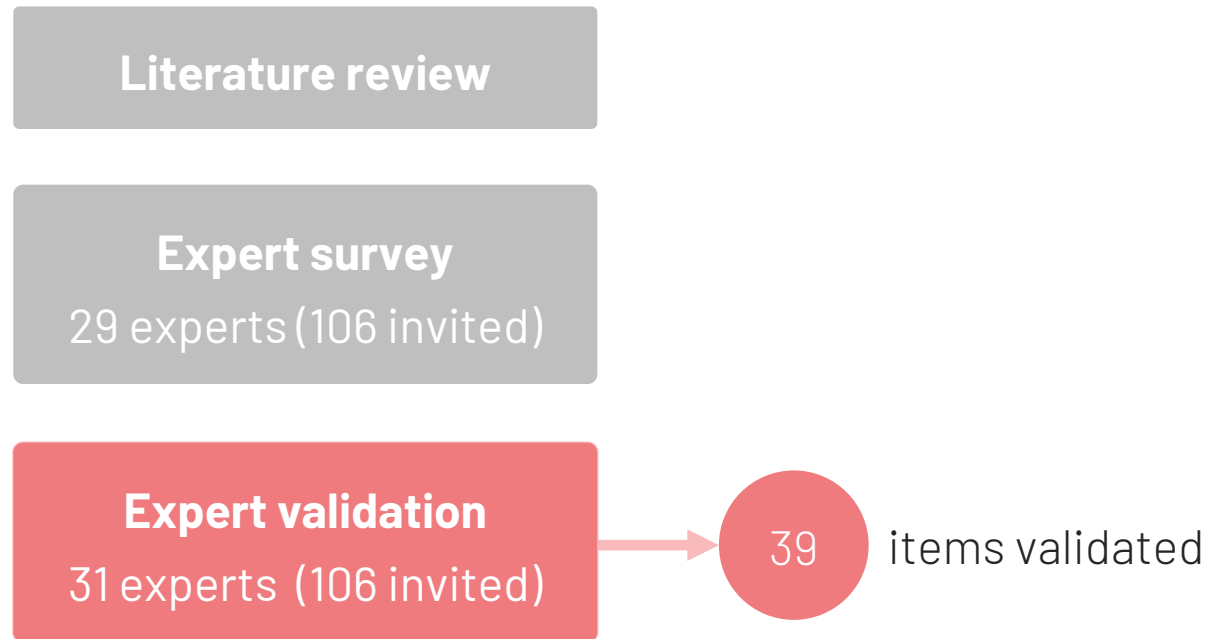
Should we write “This visualization is understandable” or “I can understand this visualization”?

Are there frequent descriptions of how people should be able to “understand” readable visualizations (e. g., quickly, easily, clearly)?

39 items

Item statement	Primary term
I can answer questions about the data after reading this visualization	answer
The visual encodings (e.g., color, shapes...) in this visualization make the information clear to me	clear
This visualization shows information in a clear way for me	clear
I find this visualization cluttered [reverted item]	clutter
I find this visualization complex to read [reverted item]	complex
I am confident in my interpretation of this visualization	confid
I find this visualization confusing [reverted item]	confus
This visualization allows me to correctly interpret the data	correct
I find the visual encodings (e.g., color, shapes...) difficult to decipher in this visualization [reverted item]	deciph
I can determine key information in this visualization	determin
This visualization allows me to discern elements of the data	discern
I find parts of the visualization distracting [reverted item]	distract
This visualization effectively shows the data to me	effect
I can understand how the data is encoded in this visualization	encod
I can easily find specific elements or information in this visualization	find
I find this visualization easy to follow	follow
This visualization helps me understand the data	help
I can identify relevant information in this visualization	identifi
I can easily retrieve information from this visualization	inform
I find this visualization informative	inform
I find this visualization easy to interpret	interpret
I find this visual design intuitive	intuit
The visual encodings (e.g., color, shapes) in this visualization make sense to me	make sense
I can easily interpret the overall meaning of the data visualization	mean
I can understand what the visual components of the visualization mean	mean
It is obvious for me how to read this visualization	obvious
I find this visualization well organized	organiz
This visualization provides me with a good overview of the data	overview
I can easily perceive data features (e.g., trends, minimums, outliers...) in this visualization	perceiv
I can easily read this visualization	read
I find this visualization readable	readabl
I can recognize data features (e.g., trends, minimums, outliers...) in this visualization	recogn
I can clearly see data features (e.g., trends, minimums, outliers...) in this visualization	see
This visualization shows the data in an appropriate manner for me	show
I find this visualization simple to read	simpl
I can easily understand this visualization	understand
I can quickly understand this visualization	understand
I can extract data values from this visualization	valu
I find data features (e.g., trends, minimums, outliers...) visible in this visualization	visibl

Item generation: content validity



2. Developing a questionnaire



Pre-testing with non-expert visualization readers

- ▶ Think-aloud procedure
- ▶ 11 participants over 3 rounds (saturation)
- ▶ Respondent problem matrix coding to identify difficulties in answering the questions
[Conrad, 1996]

EXAMPLE: ITEM DROPPED

This visualization shows the data in an appropriate manner for me


"So, uh, I have difficulty to answer this (...) It's hard to know what's an appropriate manner. For me." [1042447091]

Pre-testing with non-expert visualization readers

- ▶ Think-aloud procedure
- ▶ 11 participants over 3 rounds (saturation)
- ▶ Respondent problem matrix coding to identify difficulties in answering the questions [Conrad, 1996]


I can **easily find** specific elements in this visualization.

1 2 3 4 5 6 7



I find this visualization **messy**.

1 2 3 4 5 6 7



Real target

Selected answer

Pre-testing with non-expert visualization readers

- ▶ Think-aloud procedure
- ▶ 11 participants over 3 rounds (saturation)
- ▶ Respondent problem matrix coding to identify difficulties in answering the questions
[Conrad, 1996]

I can **easily find** specific elements in this visualization.



Strongly disagree



Disagree



Slightly disagree



Neutral



Slightly agree



Agree



Strongly agree

I find this visualization **messy**.



Strongly disagree



Disagree



Slightly disagree



Neutral



Slightly agree



Agree



Strongly agree

Pre-testing with non-expert visualization readers

- ▶ Think-aloud procedure
- ▶ 11 participants over 3 rounds (saturation)
- ▶ Respondent problem matrix coding to identify difficulties in answering the questions
[Conrad, 1996]

I can **easily find** specific elements in this visualization.

Strongly disagree

Disagree

Slightly disagree

Neutral

Slightly agree

Agree

Strongly agree

I **don't** find this visualization **messy**.

Strongly disagree

Disagree

Slightly disagree

Neutral

Slightly agree

Agree

Strongly agree

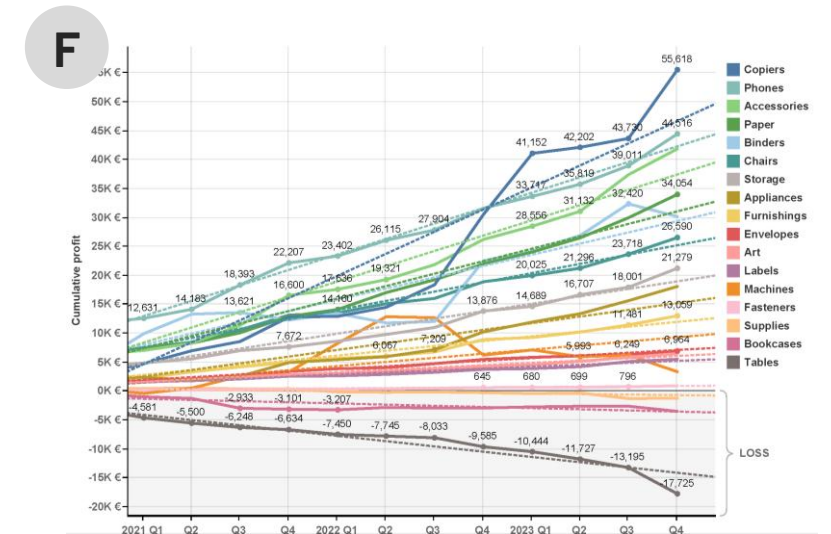
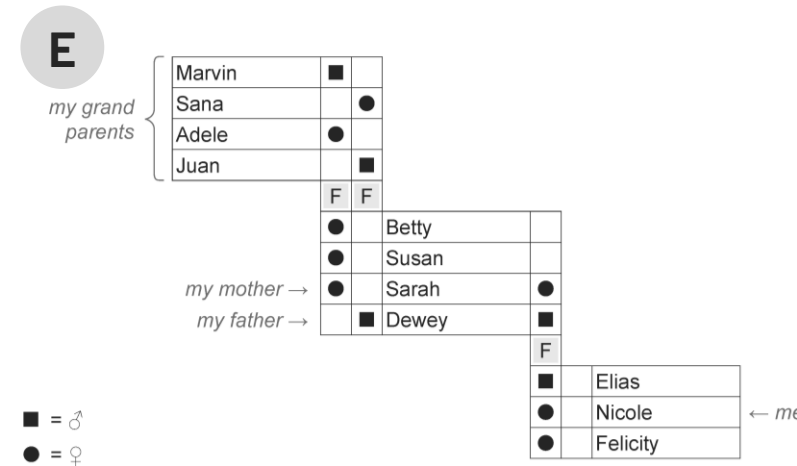
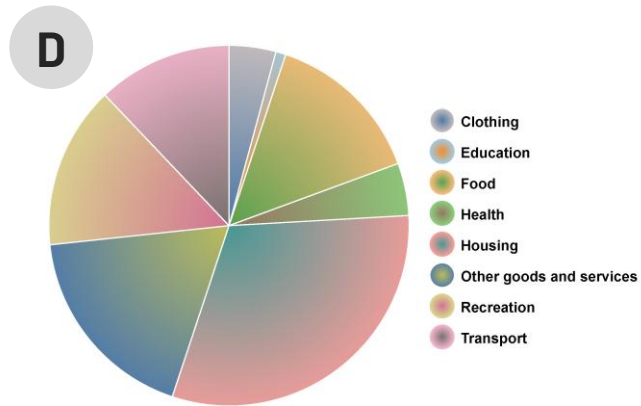
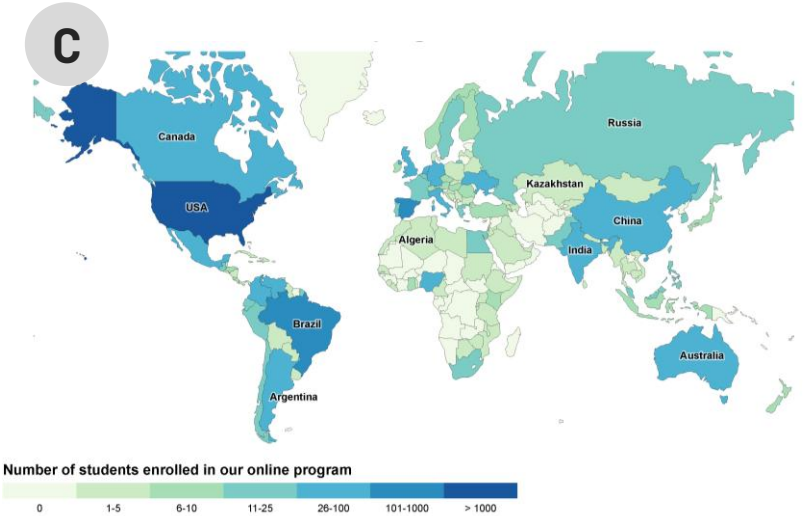
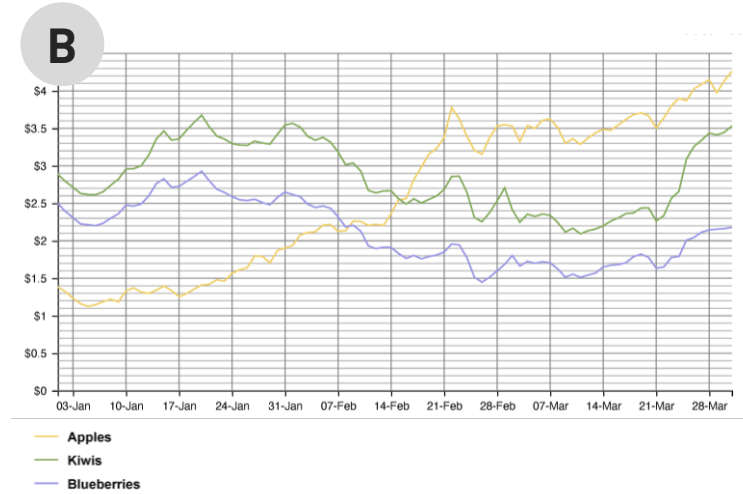
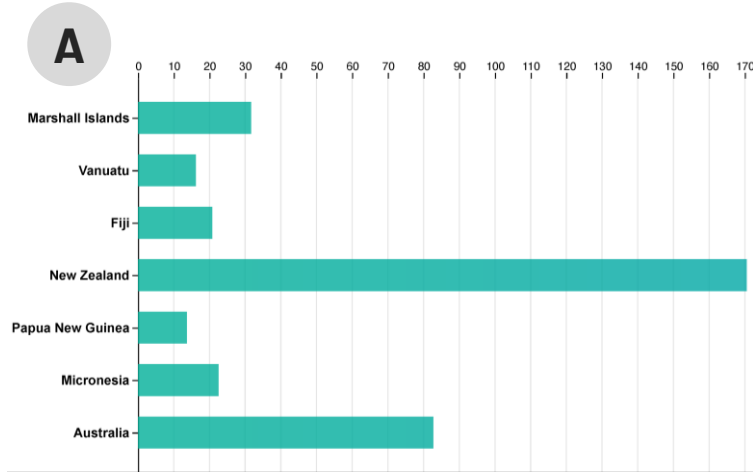
Pre-testing with non-expert visualization readers

- ▶ 11 participants over 3 rounds (saturation)
 - ▶ Think-aloud procedure
 - ▶ Respondent problem matrix coding to identify difficulties in answering the questions
- [Conrad, 1996]



29 items left

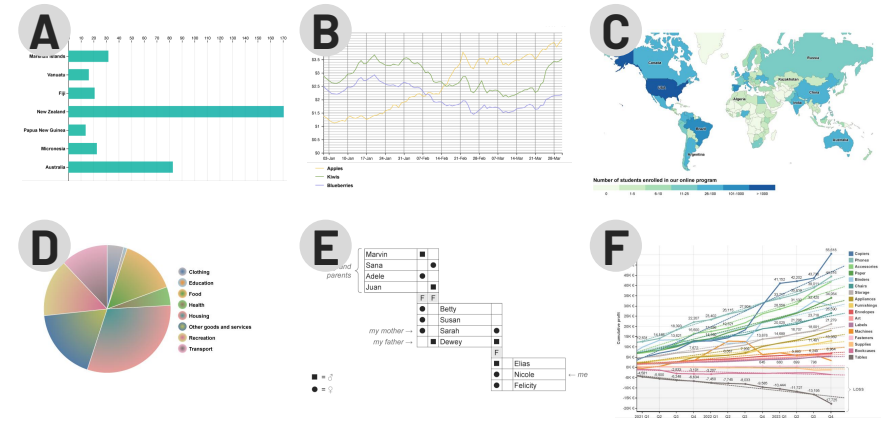
Exploratory survey



A. Bezerianos, P. Dragicevic, J.-D. Fekete, J. Bae, and B. Watson.
 GeneaQuilts: A system for exploring large genealogies. doi: [10/btb4js](https://doi.org/10/btb4js)

Exploratory survey

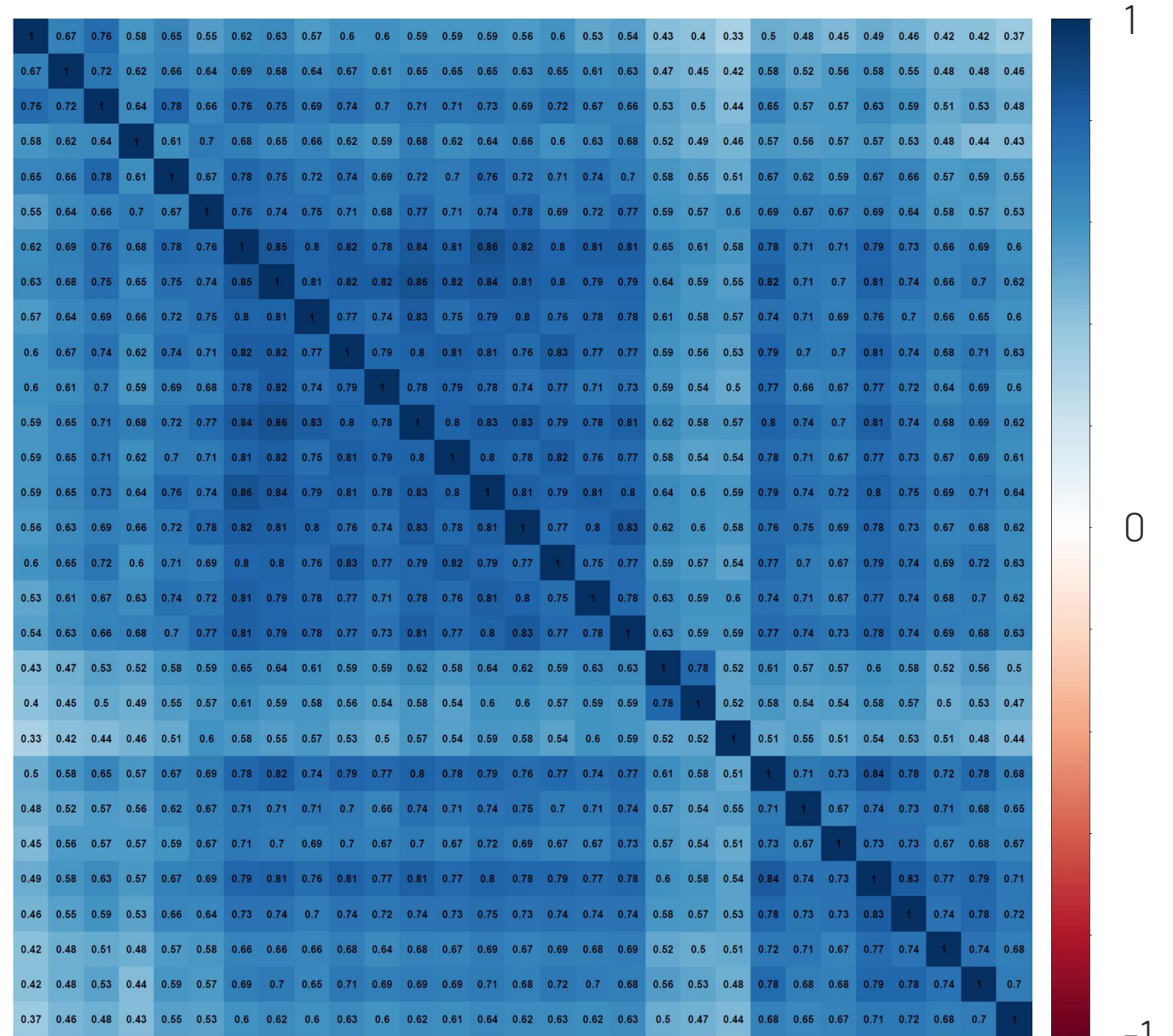
- ▶ 6 visualizations
- ▶ Each participant saw one visualization
- ▶ 2 reading questions (low-level tasks)
- ▶ 29 candidate readability questions (fully randomized order)



Stimulus	N
A	288
B	288
C	287
D	309
E	293
F	293
Total	1758

Exploratory Factor Analysis

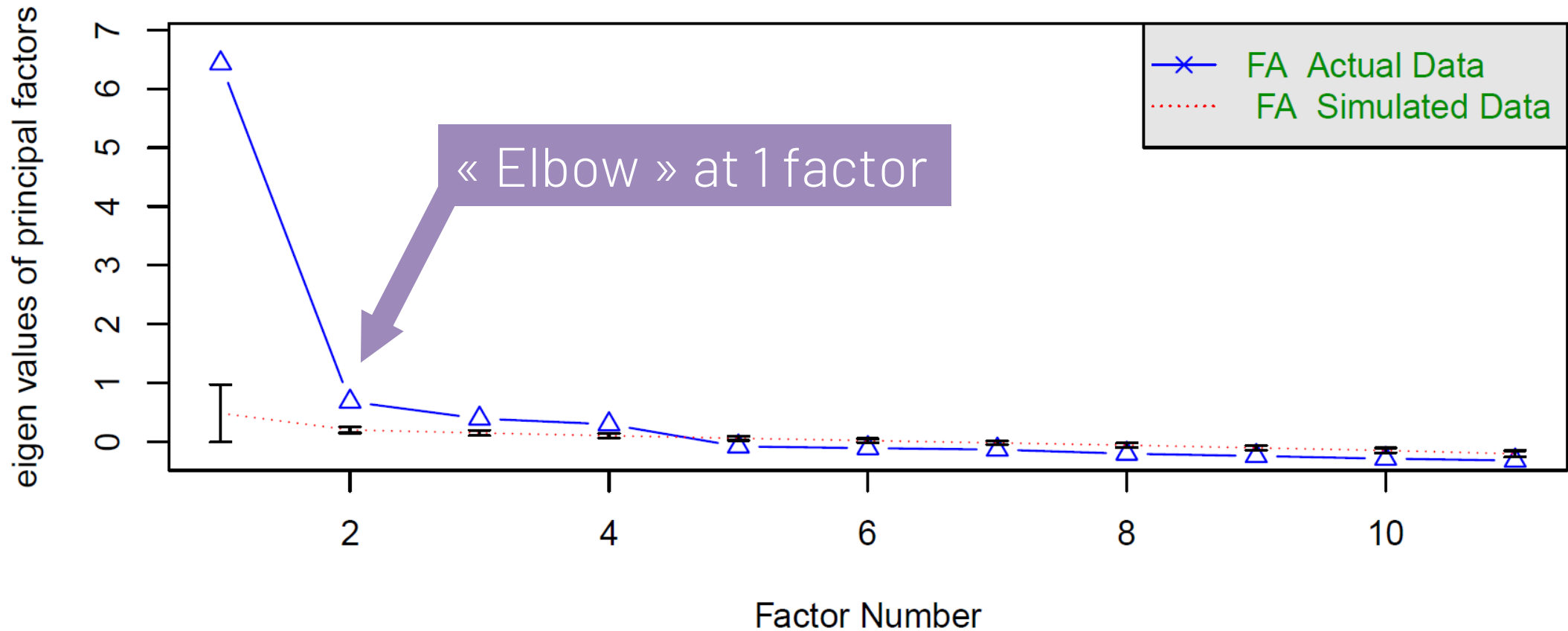
Identifying meaningful co-variance patterns (i.e., dimensions, or factors)



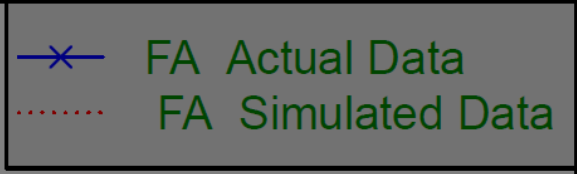
Correlation matrix from our exploratory study data

How many factors?

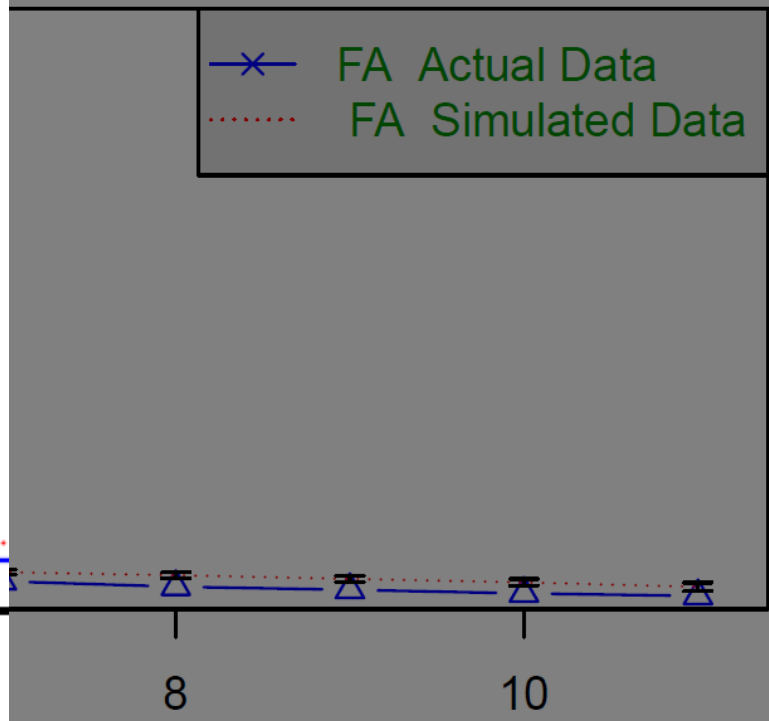
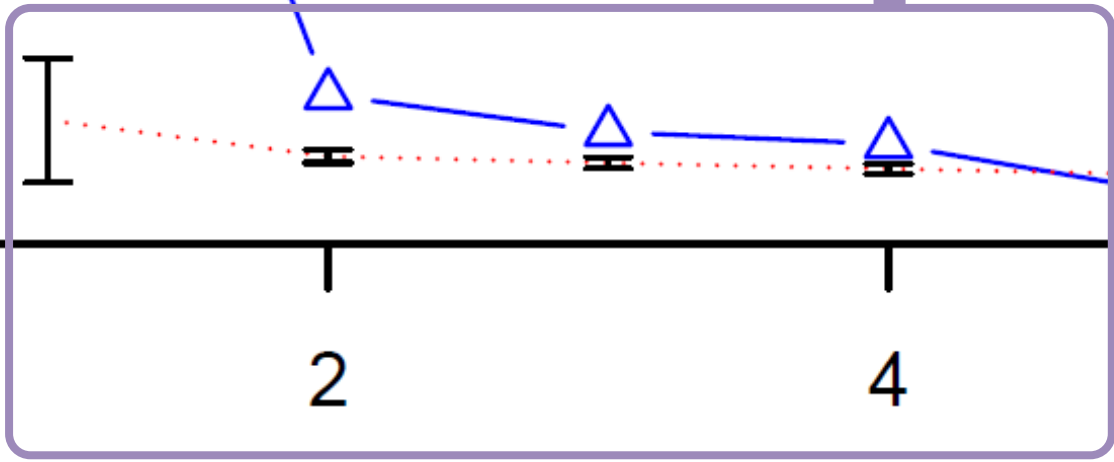
Scree plots with parallel analysis



analysis



Up to a 4-factor solution, the signal is stronger than at random



4-factors solution

1	0.7	0.74	0.78	0.79	0.78	0.68	0.42	0.53	0.48	0.59	0.57	0.68	0.48	0.68	0.65	0.53	0.56
0.7	1	0.68	0.72	0.71	0.68	0.67	0.37	0.48	0.46	0.55	0.54	0.63	0.44	0.62	0.61	0.48	0.5
0.74	0.68	1	0.74	0.77	0.72	0.67	0.42	0.51	0.48	0.57	0.58	0.69	0.5	0.67	0.66	0.5	0.52
0.78	0.72	0.74	1	0.83	0.78	0.73	0.46	0.59	0.55	0.66	0.64	0.74	0.53	0.73	0.7	0.57	0.58
0.79	0.71	0.77	0.83	1	0.84	0.73	0.49	0.63	0.58	0.67	0.69	0.78	0.53	0.78	0.76	0.58	0.6
0.78	0.68	0.72	0.78	0.84	1	0.73	0.5	0.65	0.58	0.67	0.69	0.76	0.51	0.76	0.74	0.58	0.61
0.68	0.67	0.67	0.73	0.73	0.73	1	0.45	0.57	0.56	0.59	0.67	0.73	0.5	0.69	0.69	0.54	0.57
0.42	0.37	0.42	0.46	0.49	0.5	0.45	1	0.76	0.67	0.65	0.55	0.54	0.32	0.56	0.57	0.4	0.42
0.53	0.48	0.51	0.59	0.63	0.65	0.57	0.76	1	0.72	0.78	0.66	0.66	0.44	0.69	0.69	0.5	0.53
0.48	0.46	0.48	0.55	0.58	0.58	0.56	0.67	0.72	1	0.66	0.64	0.63	0.42	0.64	0.64	0.45	0.47
0.59	0.55	0.57	0.66	0.67	0.67	0.59	0.65	0.78	0.66	1	0.67	0.7	0.51	0.71	0.72	0.55	0.58
0.57	0.54	0.58	0.64	0.69	0.69	0.67	0.55	0.66	0.64	0.67	1	0.77	0.6	0.78	0.75	0.57	0.59
0.68	0.63	0.69	0.74	0.78	0.76	0.73	0.54	0.66	0.63	0.7	0.77	1	0.59	0.83	0.78	0.59	0.63
0.48	0.44	0.5	0.53	0.53	0.51	0.5	0.32	0.44	0.42	0.51	0.6	0.59	1	0.58	0.57	0.52	0.52
0.68	0.62	0.67	0.73	0.78	0.76	0.69	0.56	0.69	0.64	0.71	0.78	0.83	0.58	1	0.8	0.6	0.62
0.65	0.61	0.66	0.7	0.76	0.74	0.69	0.57	0.69	0.64	0.72	0.75	0.78	0.57	0.8	1	0.58	0.61
0.53	0.48	0.5	0.57	0.58	0.58	0.54	0.4	0.5	0.45	0.55	0.57	0.59	0.52	0.6	0.58	1	0.78
0.56	0.5	0.52	0.58	0.6	0.61	0.57	0.42	0.53	0.47	0.58	0.59	0.63	0.52	0.62	0.61	0.78	1

4-factors solution

◆ Reader's understanding

It is obvious for me how to read this visualization

I can easily understand the overall meaning of this data visualization

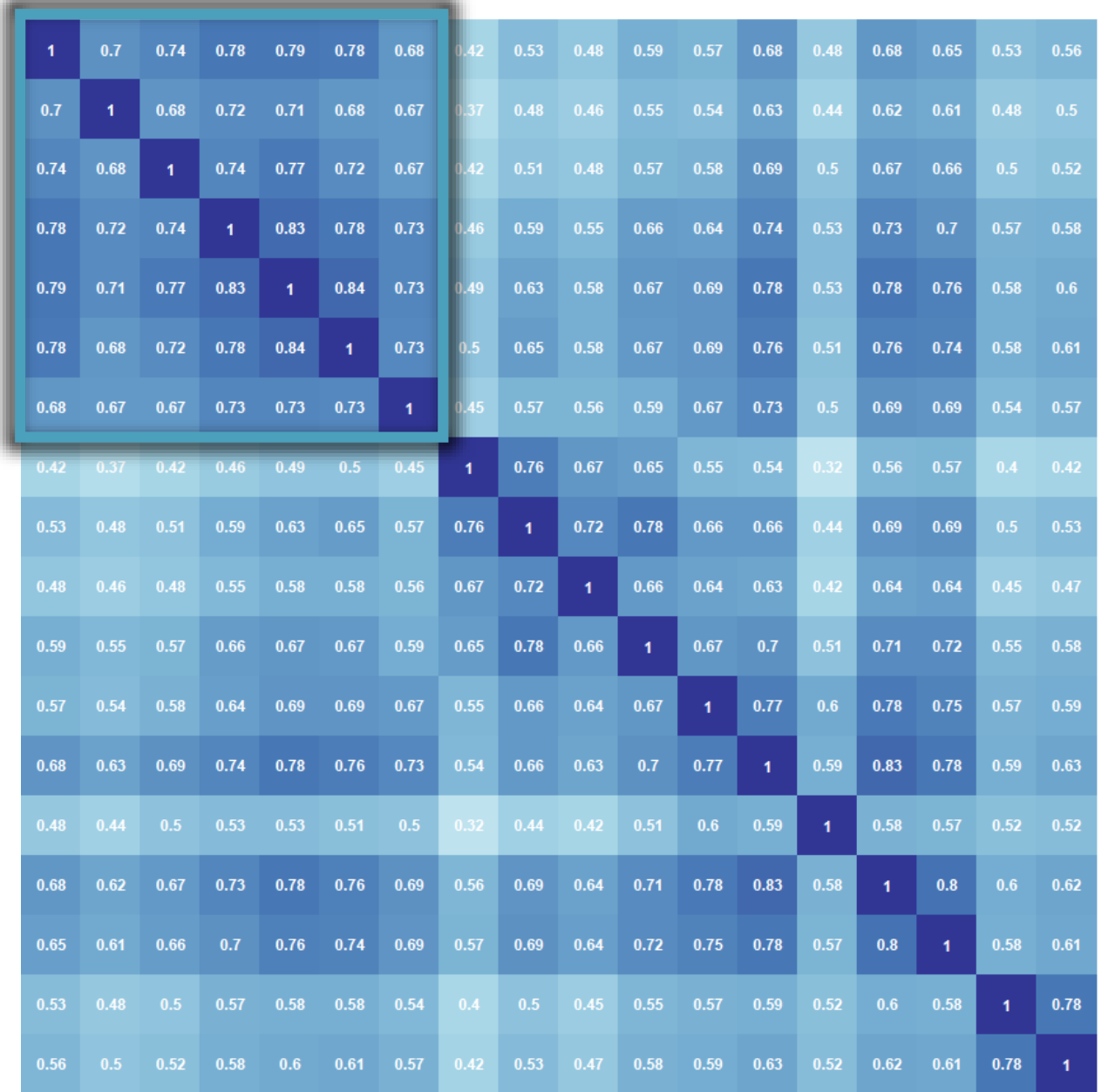
I am confident in my understanding of this visualization

I can easily understand how the data is represented in this visualization

I can easily understand this visualization

I can quickly understand this visualization

I can easily understand what the different elements of the visualization mean



4-factors solution

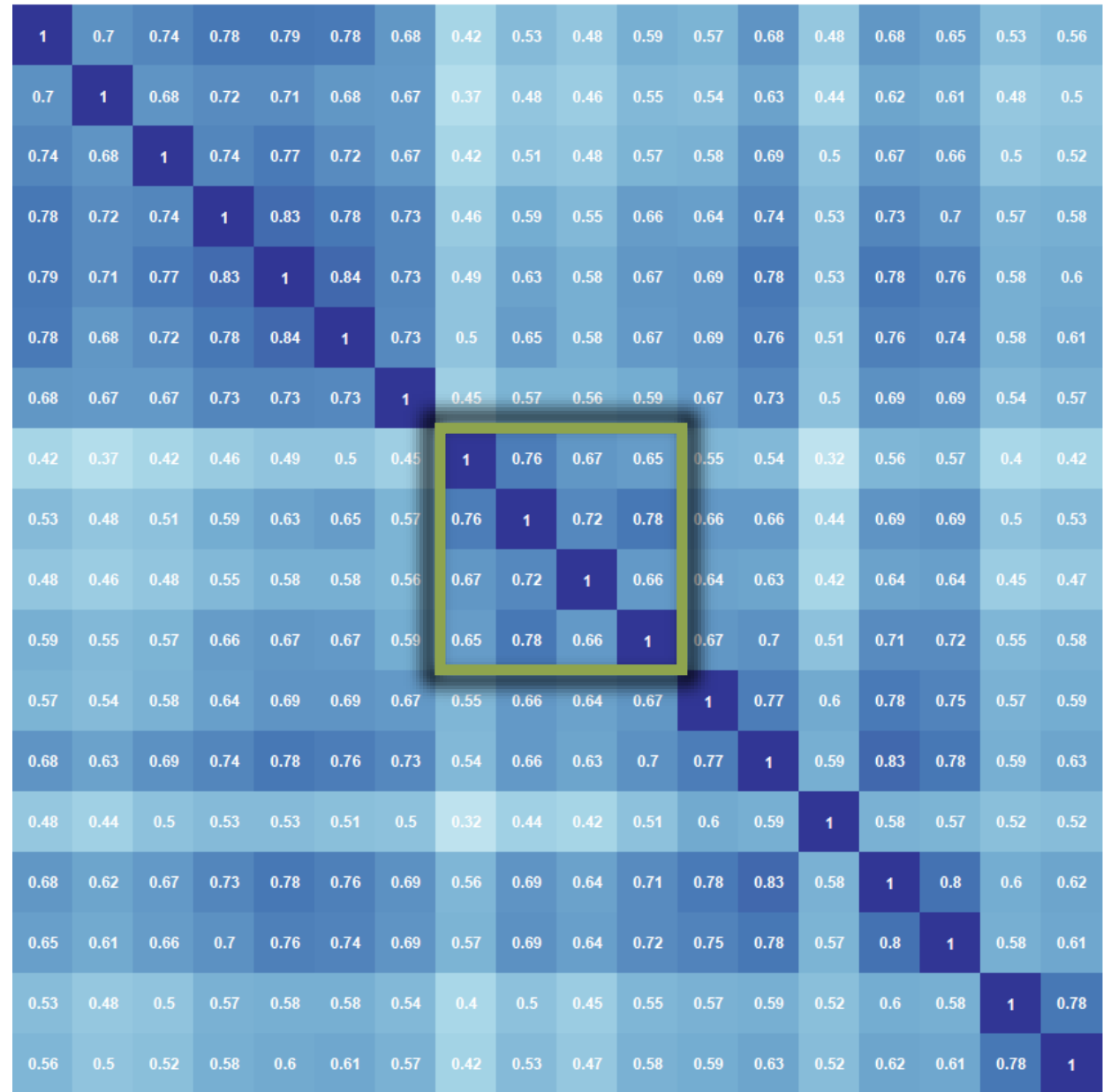
◆ Layout clarity

I don't find this visualization crowded

I don't find this visualization messy

I don't find distracting parts in this visualization

I find this visualization well organized



4-factors solution

◆ "Reading the data"

[Curcio, 1987]

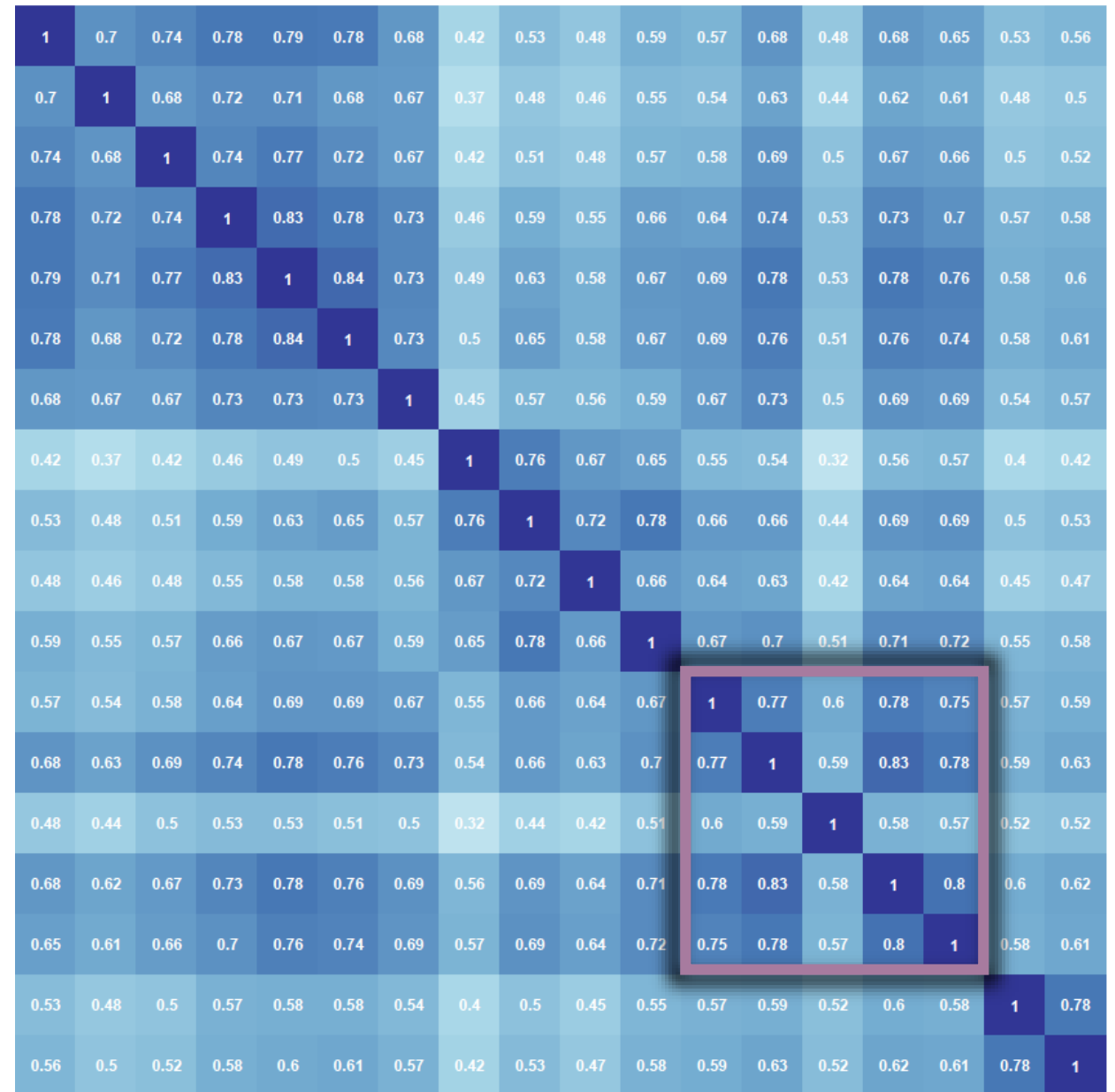
I can easily find specific elements in this visualization

I can easily identify relevant information in this visualization

I can read data values from this visualization

I can easily retrieve information from this visualization

I find this visualization readable

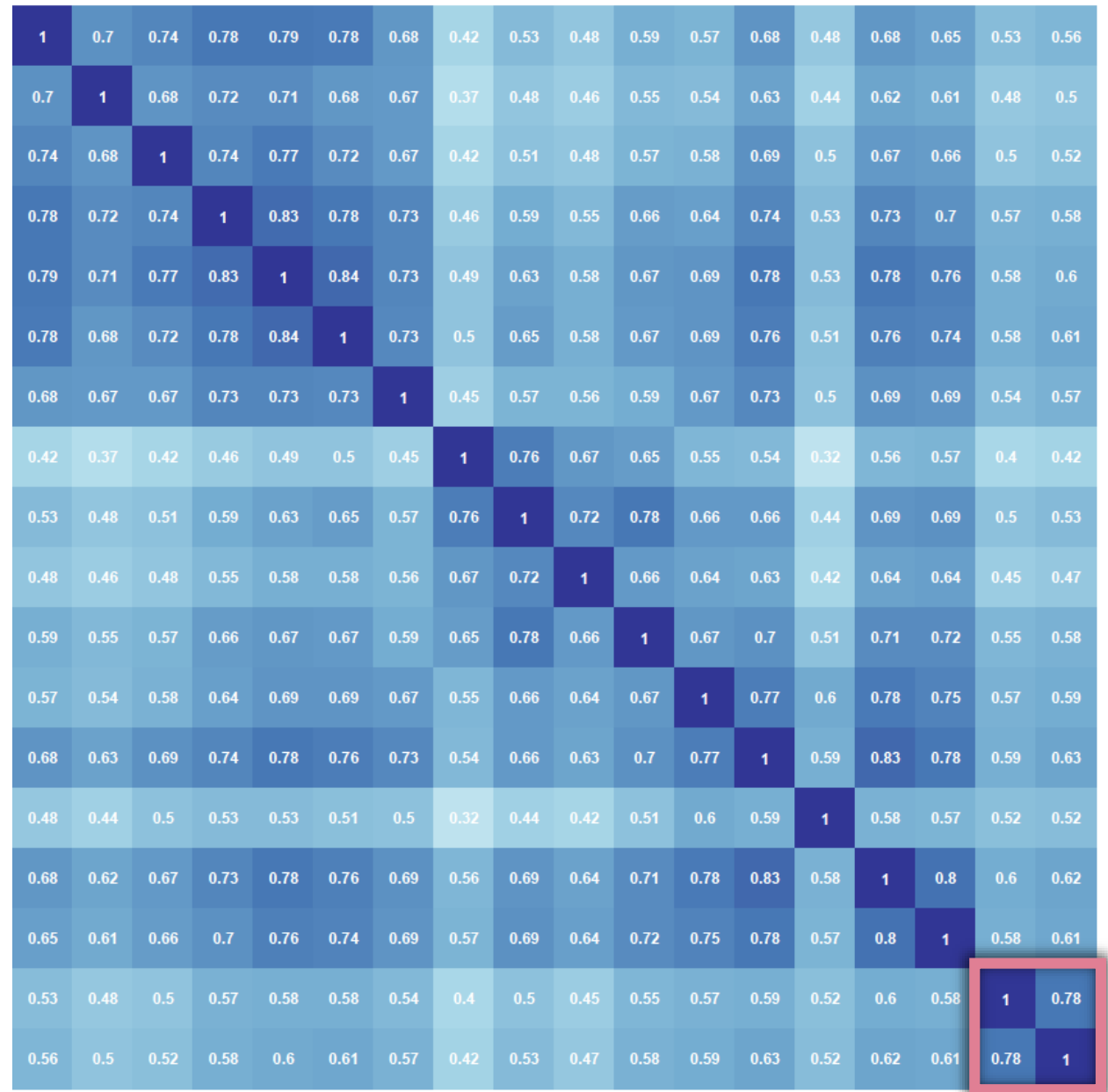


4-factors solution

◆ "Reading between the data" [Curcio, 1987]

I find data features (for example, a minimum, or an outlier, or a trend) visible in this visualization

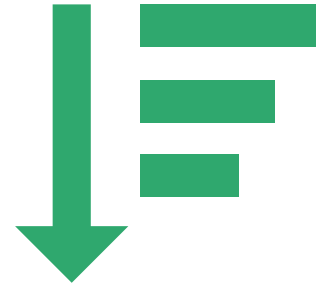
I can clearly see data features (for example, a minimum, or an outlier, or a trend) in this visualization



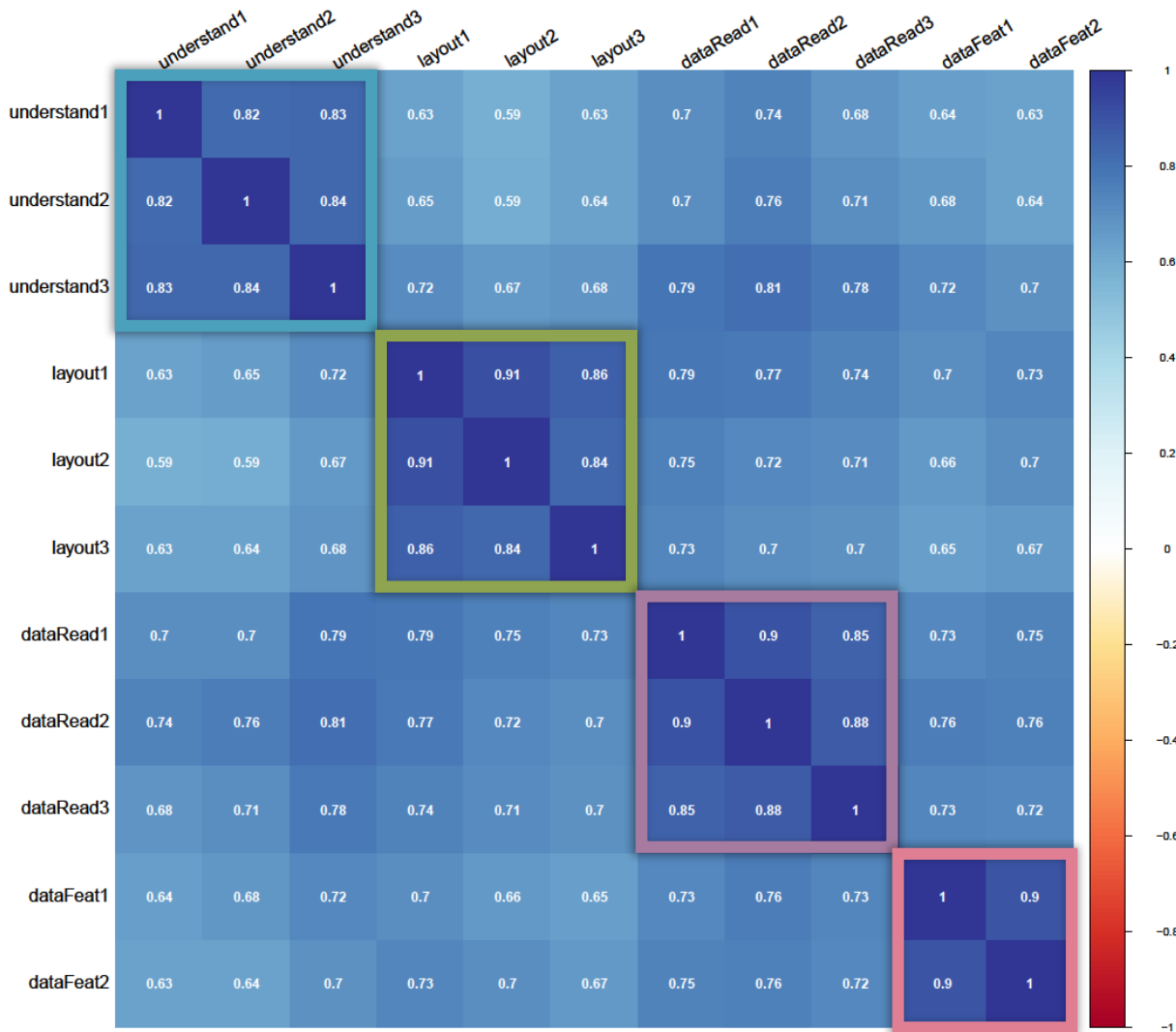
Selecting final items for each scale



- ✘ Poor factor loading
- ✘ Cross-factors loading
(contributes to 2 factors)



- ✓ Highest ranking items,
based on their contribution
to the scale's
- ✓ Congruent phrasing



◆ Understand subscale

understand1 It is **obvious** for me how to read this visualization

understand2 I can easily understand **how the data is represented** in this visualization

understand3 I can easily understand this visualization

◆ Layout subscale

layout1 I **don't** find this visualization **messy**

layout2 I **don't** find this visualization **crowded**

layout3 I **don't** find **distracting parts** in this visualization

◆ DataRead subscale

dataRead1 I can easily **find specific elements** in this visualization

dataRead2 I can easily **identify relevant information** in this visualization

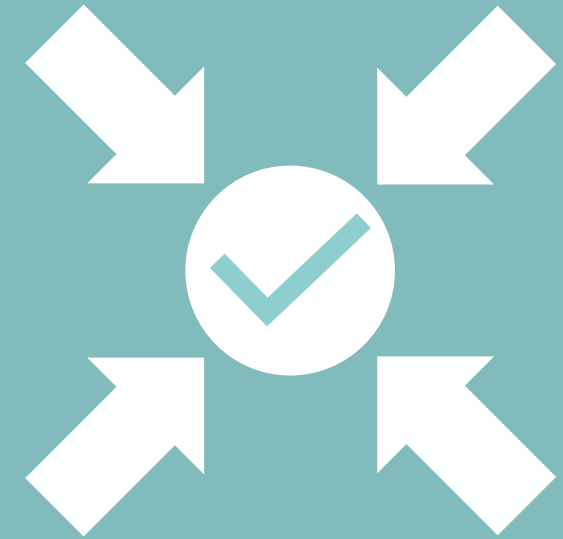
dataRead3 I can easily **retrieve information** from this visualization

◆ DataFeat subscale

dataFeat1 I find data features (for example, a minimum, or an outlier, or a trend) **visible** in this visualization

dataFeat2 I can **clearly see** data features (for example, a minimum, or an outlier, or a trend) in this visualization

3. Evaluating the questionnaire

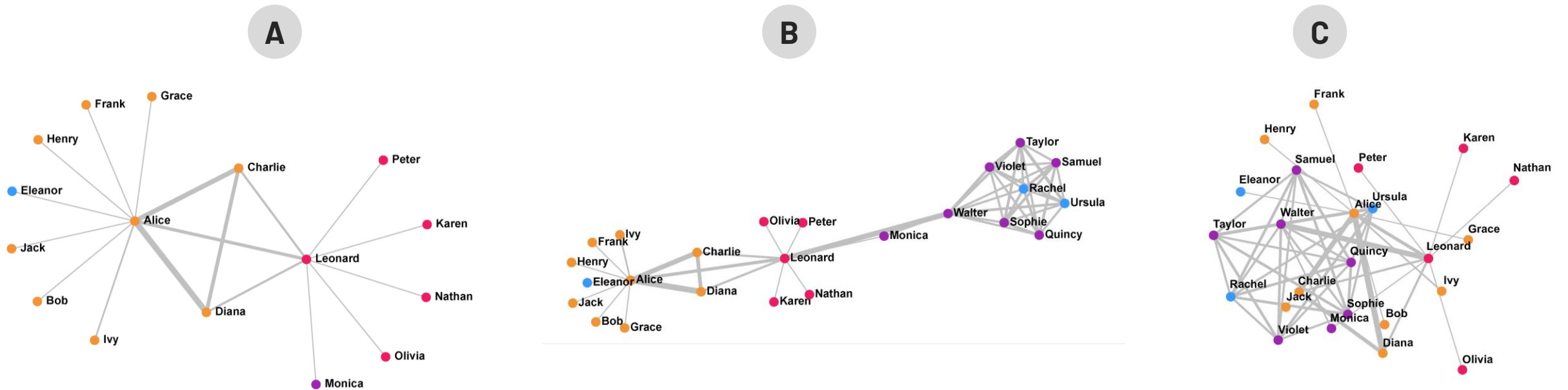


Can we confirm the 4-factors structure?

How reliable are the scales with an independent sample?

What evidence of construct validity can we establish?

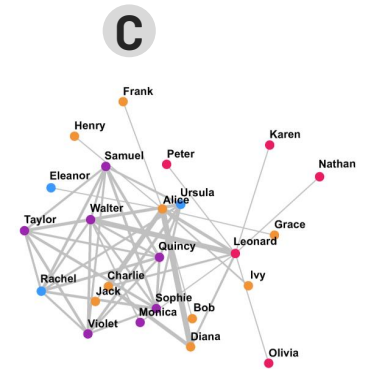
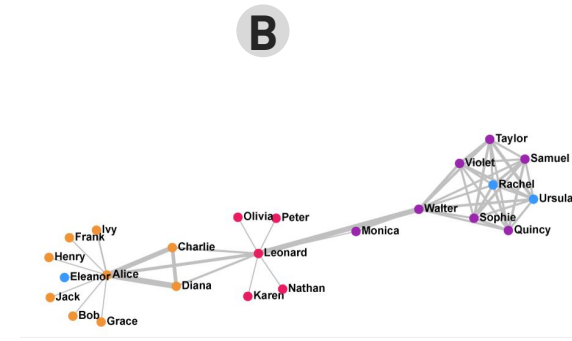
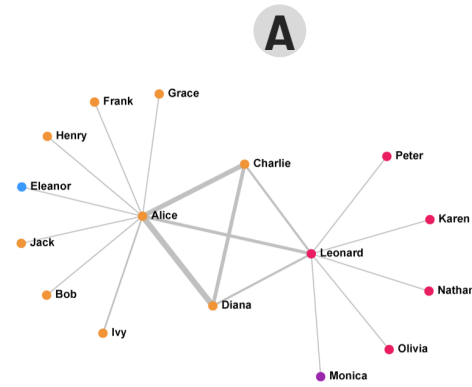
Validation study



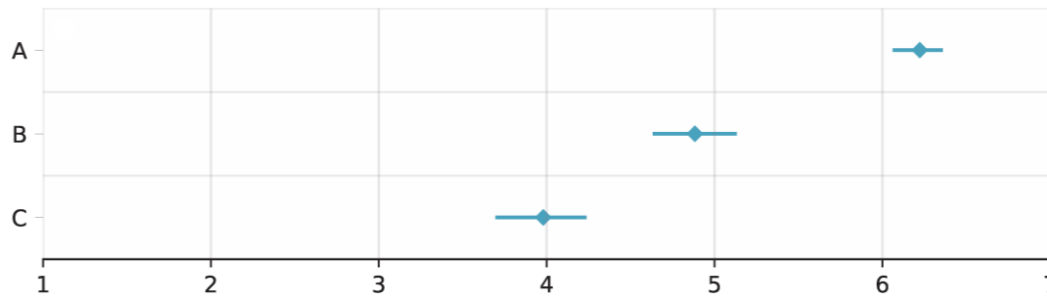
- ▶ N=147 on Prolific
- ▶ All participants saw A, B, and C in random order
- ▶ People answered 2 reading questions then 11 PREVis questions

Results: ratings

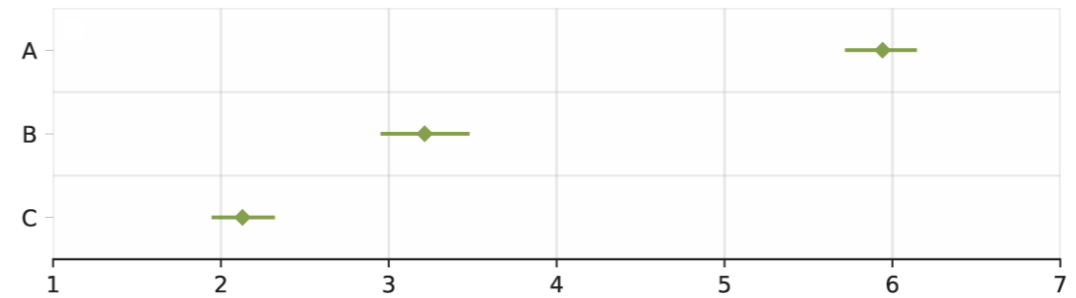
Point estimates with 95% CI



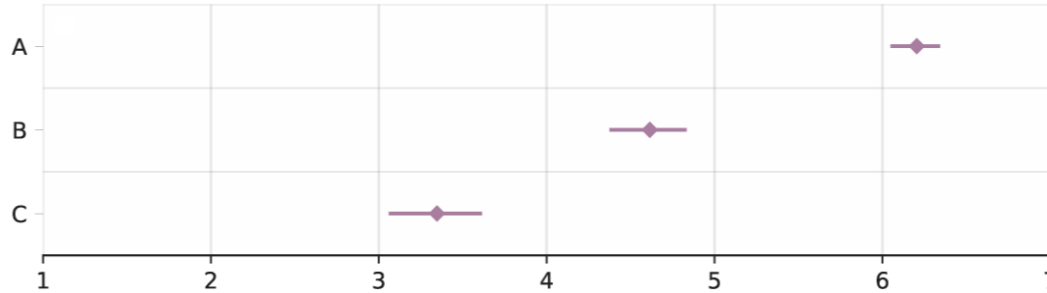
◆ Understand obvious + represent + understand



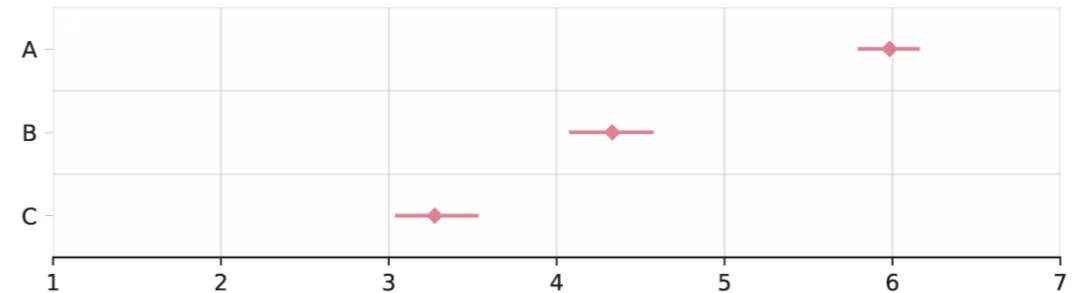
◆ Layout messy + crowded + distracting



◆ DataRead find + identify + inform



◆ DataFeat visible + clear



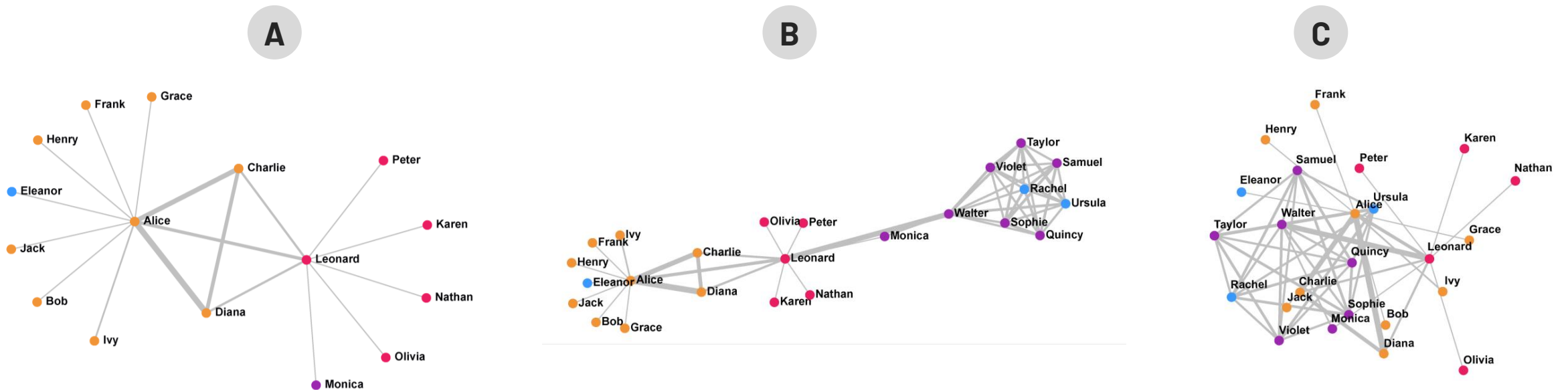
Results: convergent construct validity


Computed reference layout measures from greadability.js (Gove, 2018)

Stimulus A 0.872

Stimulus B 0.786

Stimulus C 0.679





PREVis implementation guidelines

7-points Likert agreement labels

To what extent do you agree or disagree with the following statement:

(...)

Strongly
disagree

Disagree

Slightly
disagree

Neutral

Slightly
agree

Agree

Strongly
agree

I don't know / Not applicable (if you wish, feel free to elaborate below)

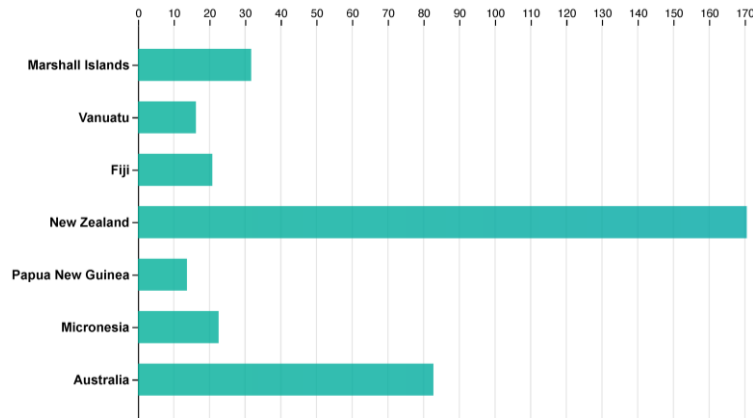


Perceived readability

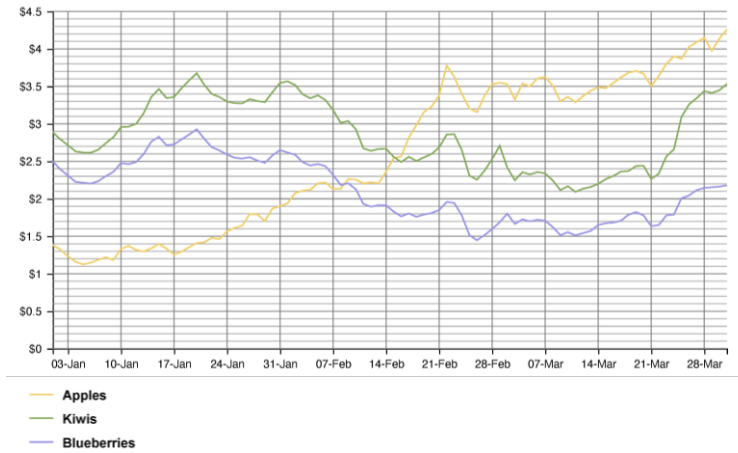
- ◆ **Understand** obvious + represent + understand
- ◆ **Layout** messy + crowded + distracting
- ◆ **DataRead** find + identify + inform
- ◆ **DataFeat** visible + clear

Exploratory survey: 6 independent samples (N=287 to 309 per sample)

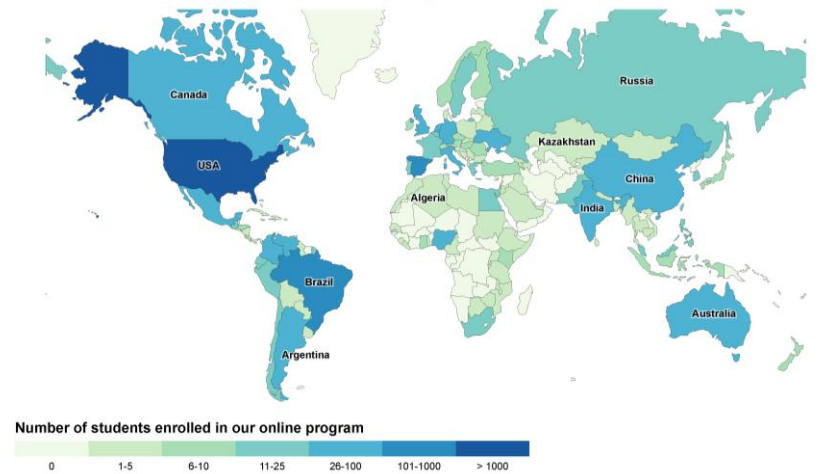
A



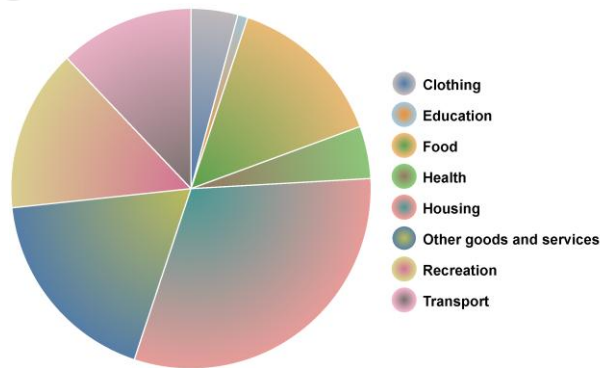
B



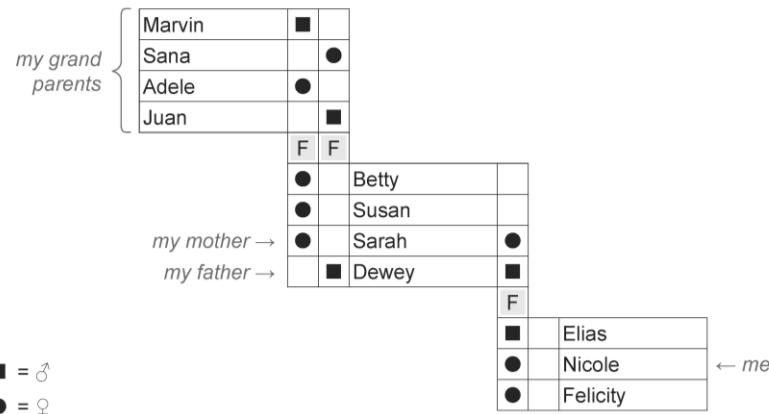
C



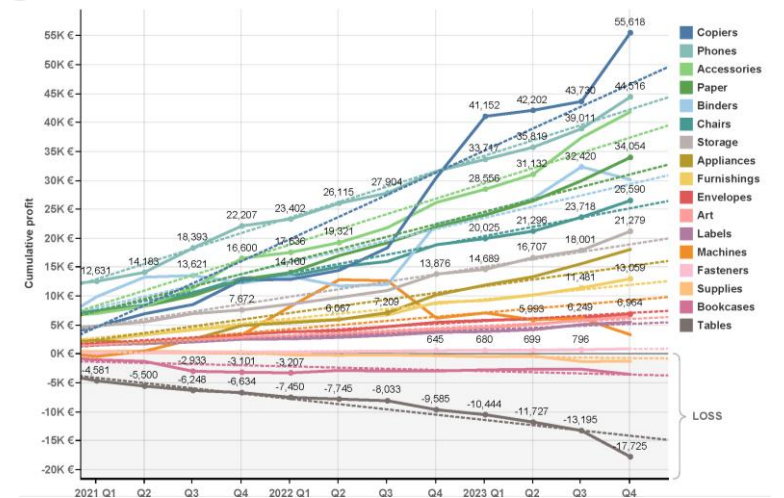
D

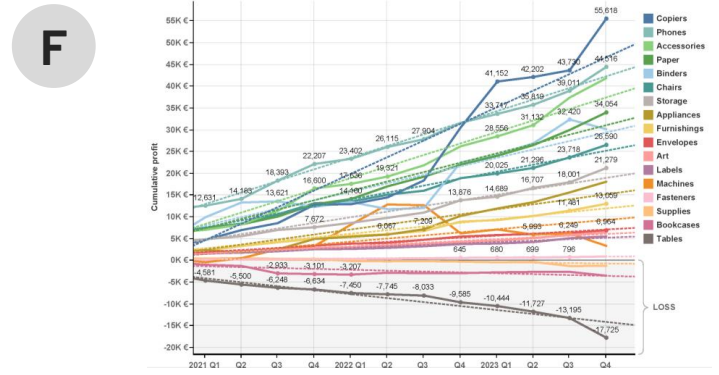
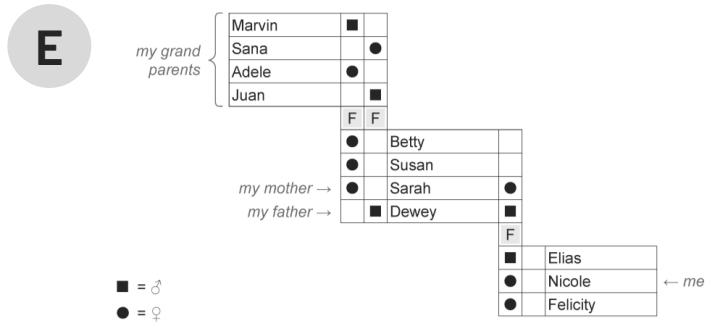
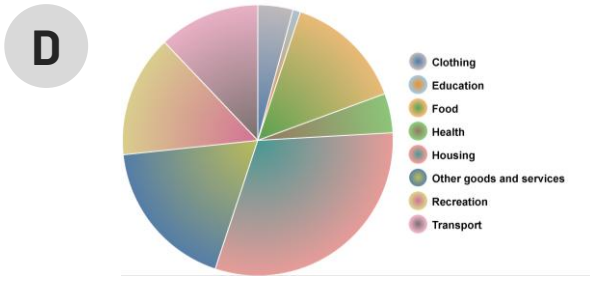


E

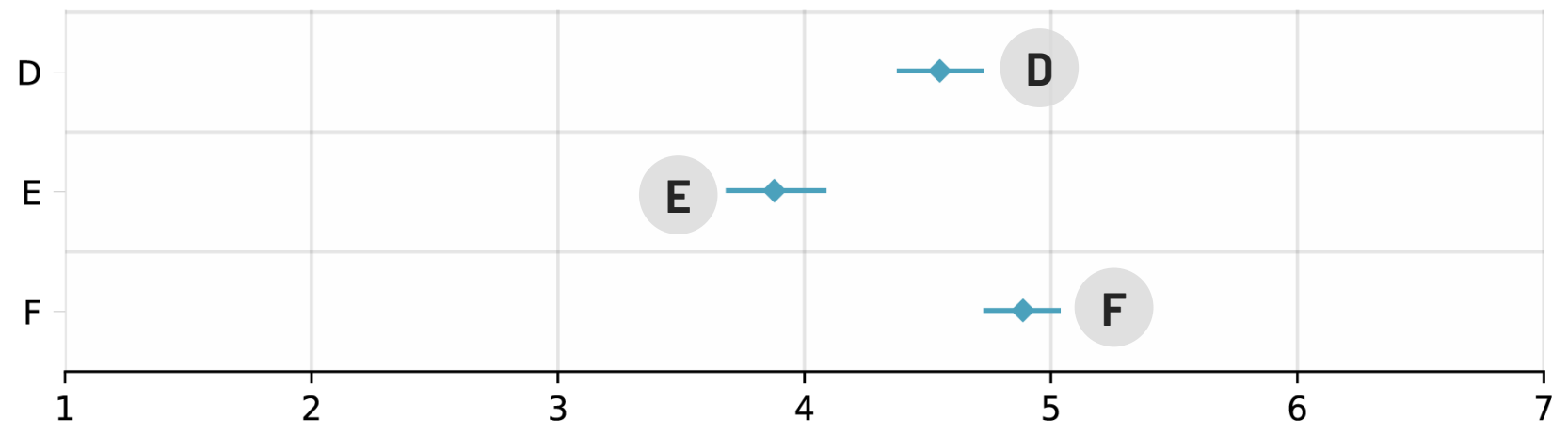


F

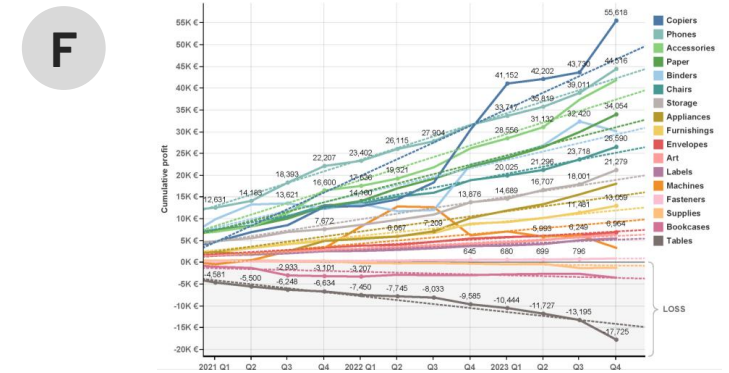
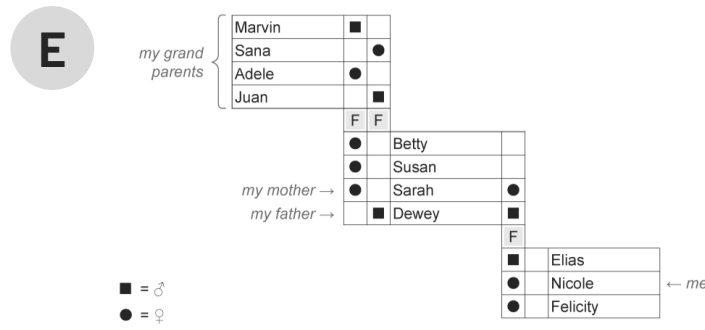
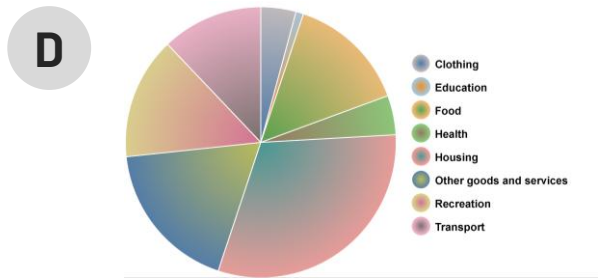




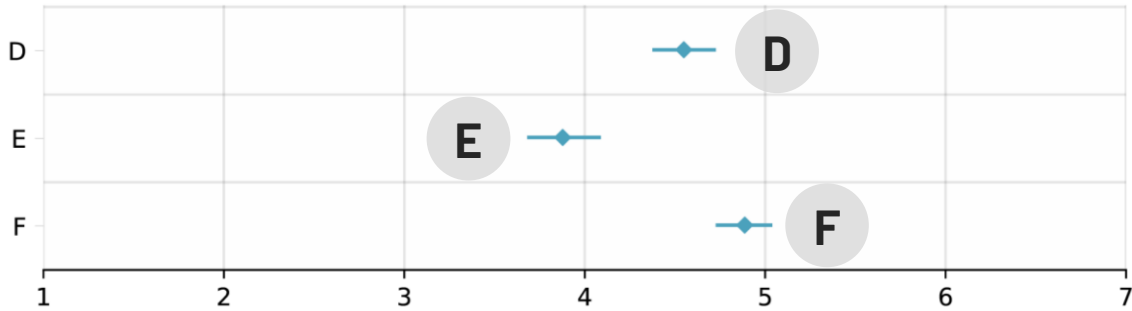
◆ Understand obvious + represent + understand



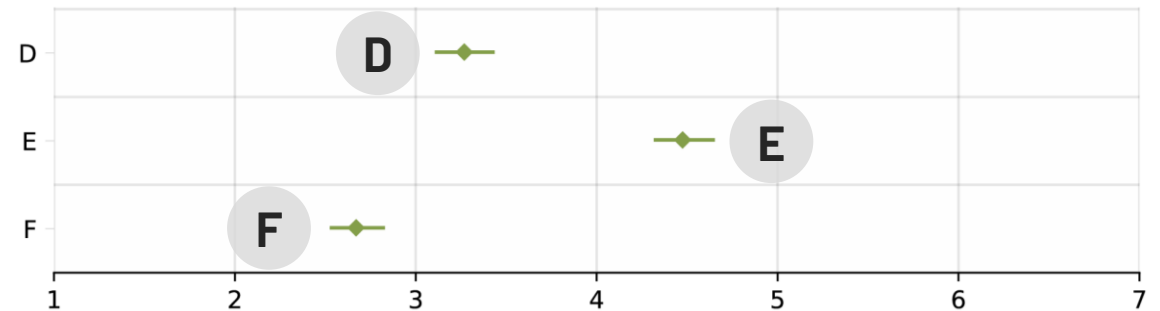
Mean score and 95% confidence interval



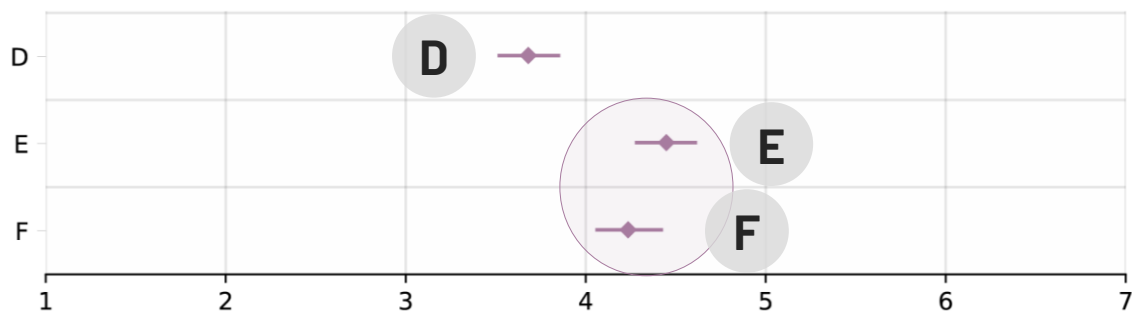
◆ **Understand** obvious + represent + understand



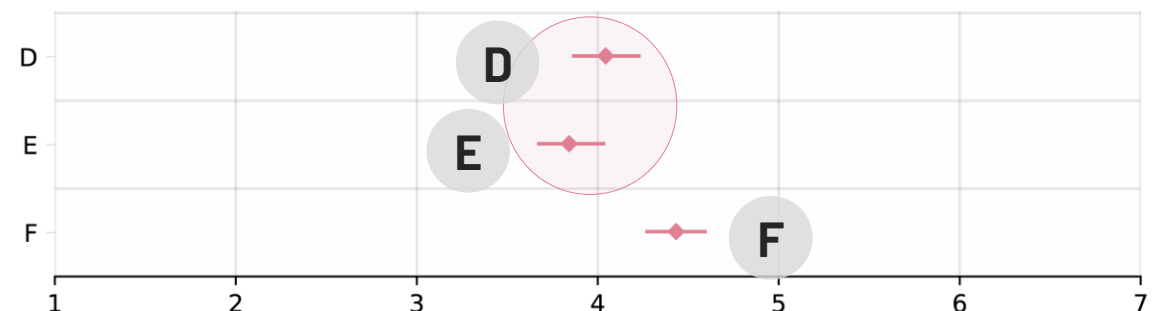
◆ **Layout** messy + crowded + distracting



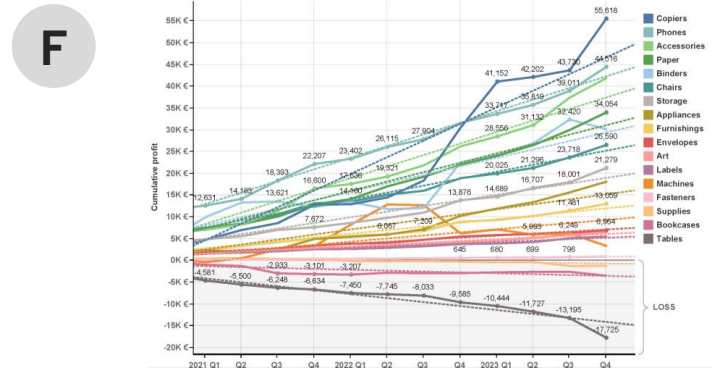
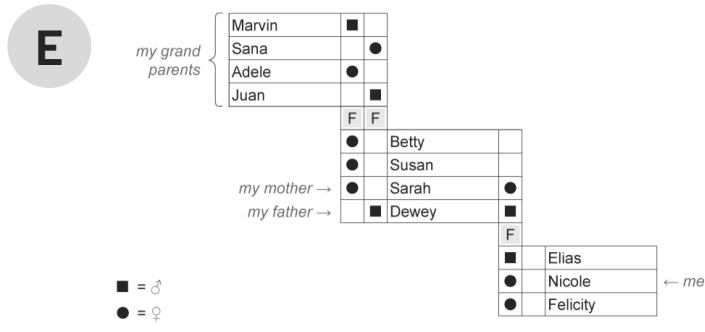
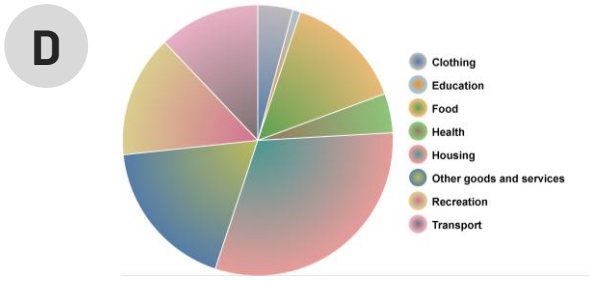
◆ **DataRead** find + identify + inform



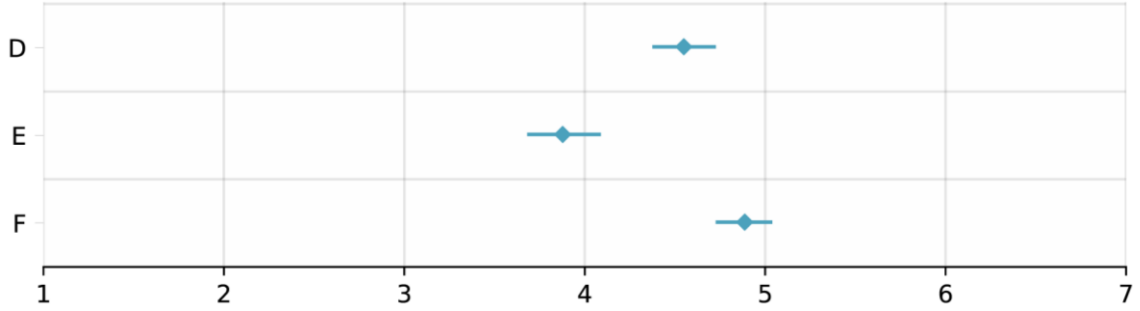
◆ **DataFeat** visible + clear



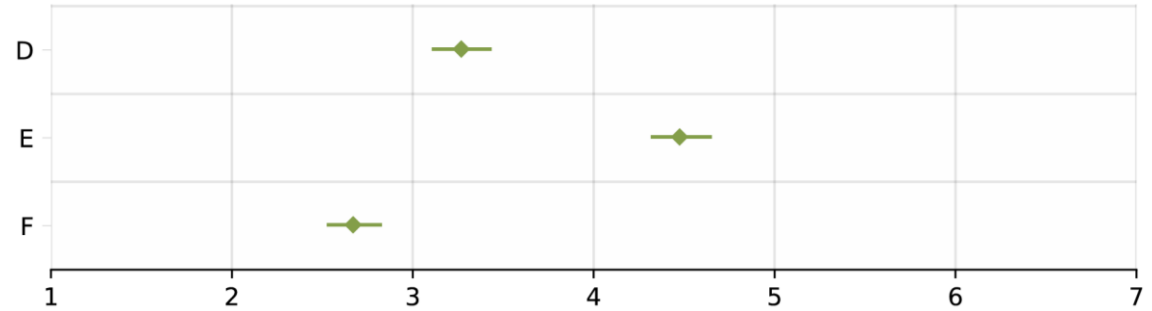
For each scale: mean score and 95% confidence interval



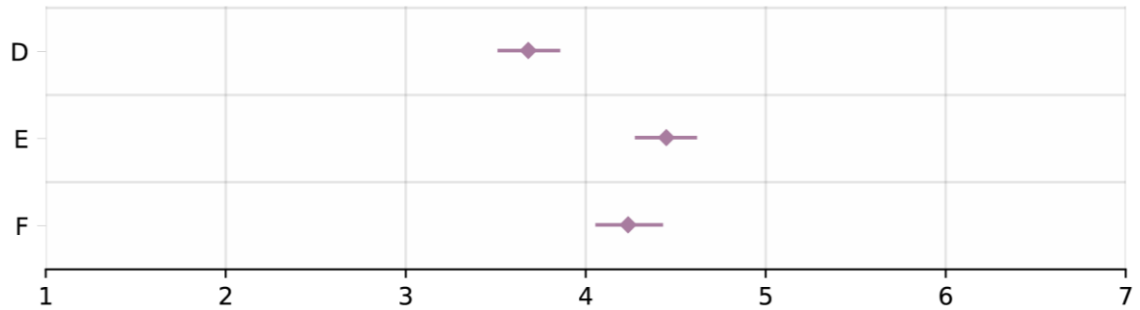
Understand obvious + represent + understand



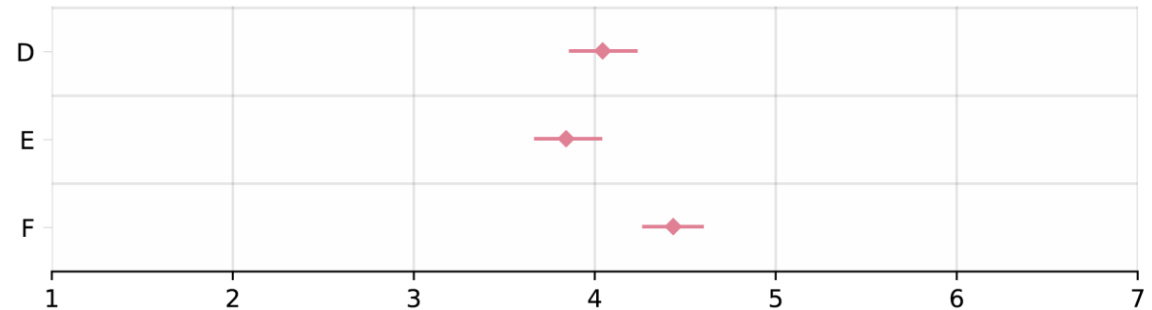
Layout messy + crowded + distracting



DataRead find + identify + inform



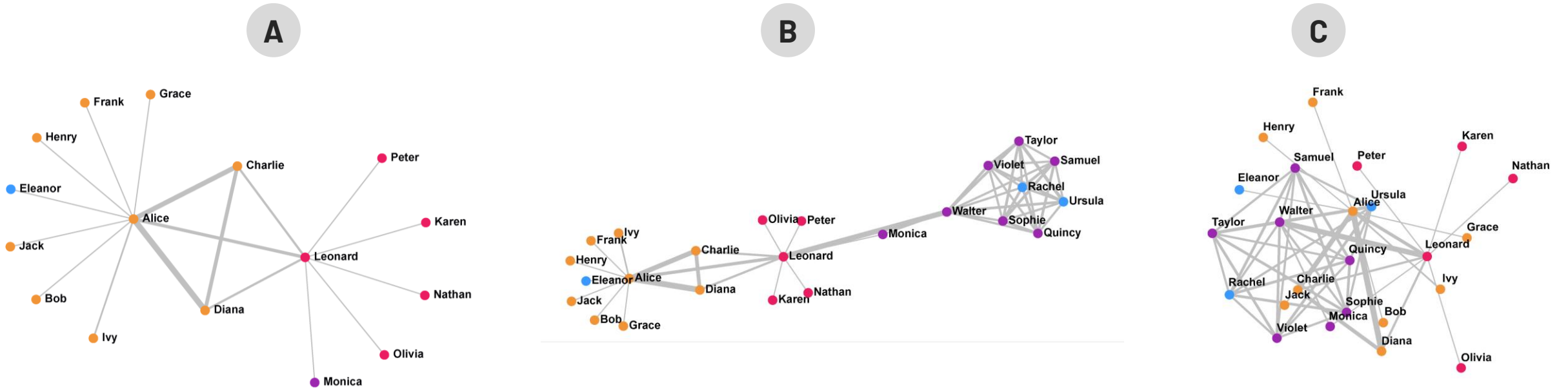
DataFeat visible + clear



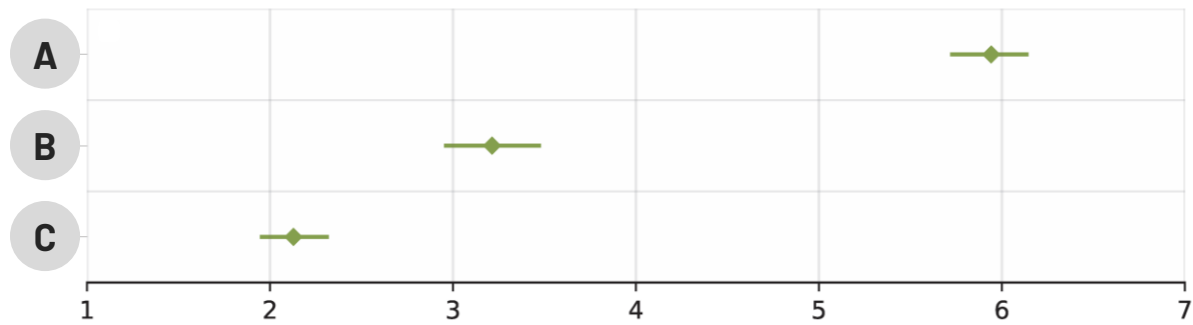
For each scale: mean score and 95% confidence interval

Individual subscales

- ◆ **Understand** obvious + represent + understand
- ◆ **Layout** messy + crowded + distracting
- ◆ **DataRead** find + identify + inform
- ◆ **DataFeat** visible + clear



◆ **Layout**

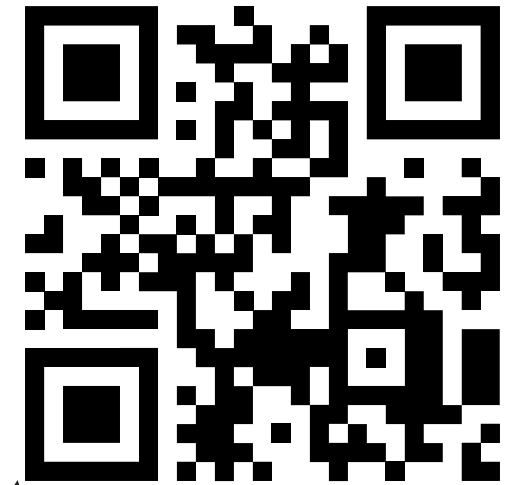


Mean score and 95% confidence interval (N=147 participants from PREVis validation study)

Filling PREVis out takes less than 1 min 30

- ◆ **Understand** obvious + represent + understand
- ◆ **Layout** messy + crowded + distracting
- ◆ **DataRead** find + identify + inform
- ◆ **DataFeat** visible + clear

Templates
FAQ
Study protocols
Data analysis code

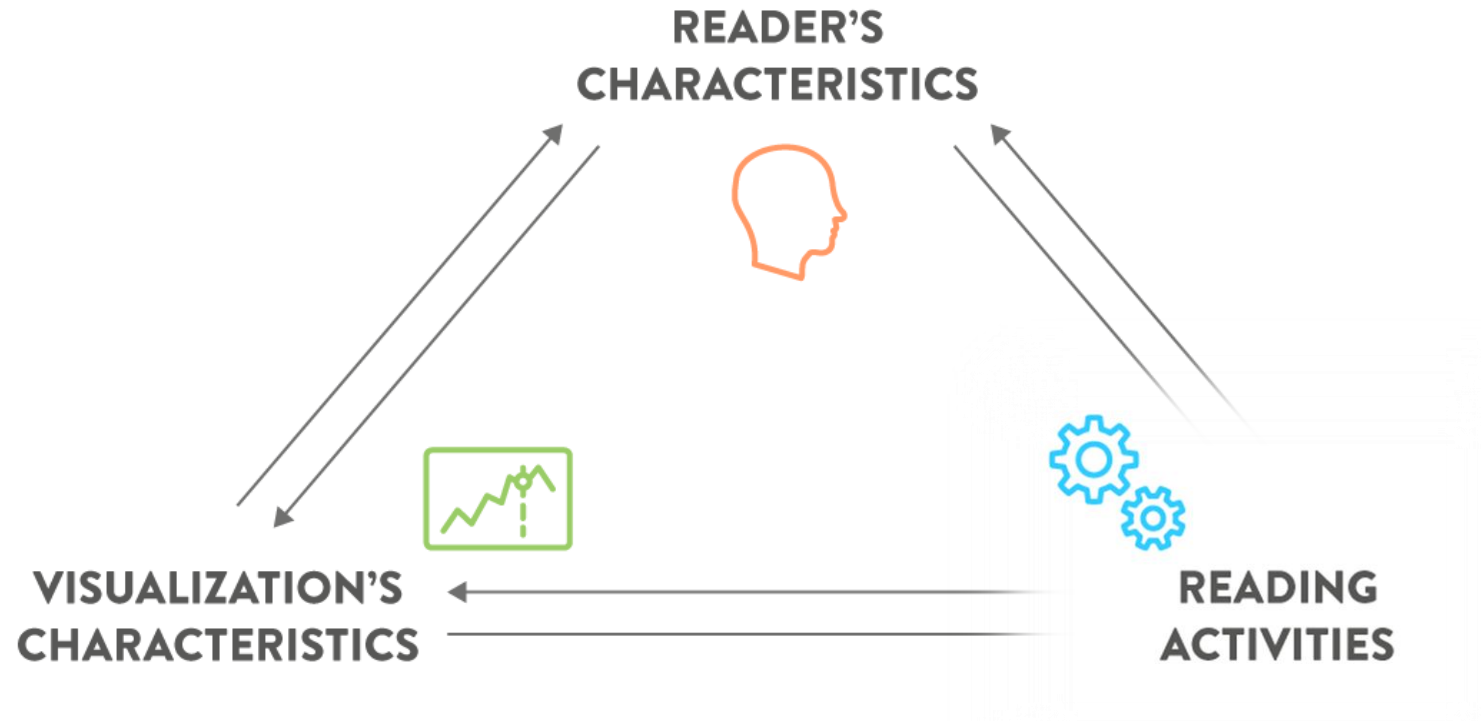


aviz.fr/PREVis

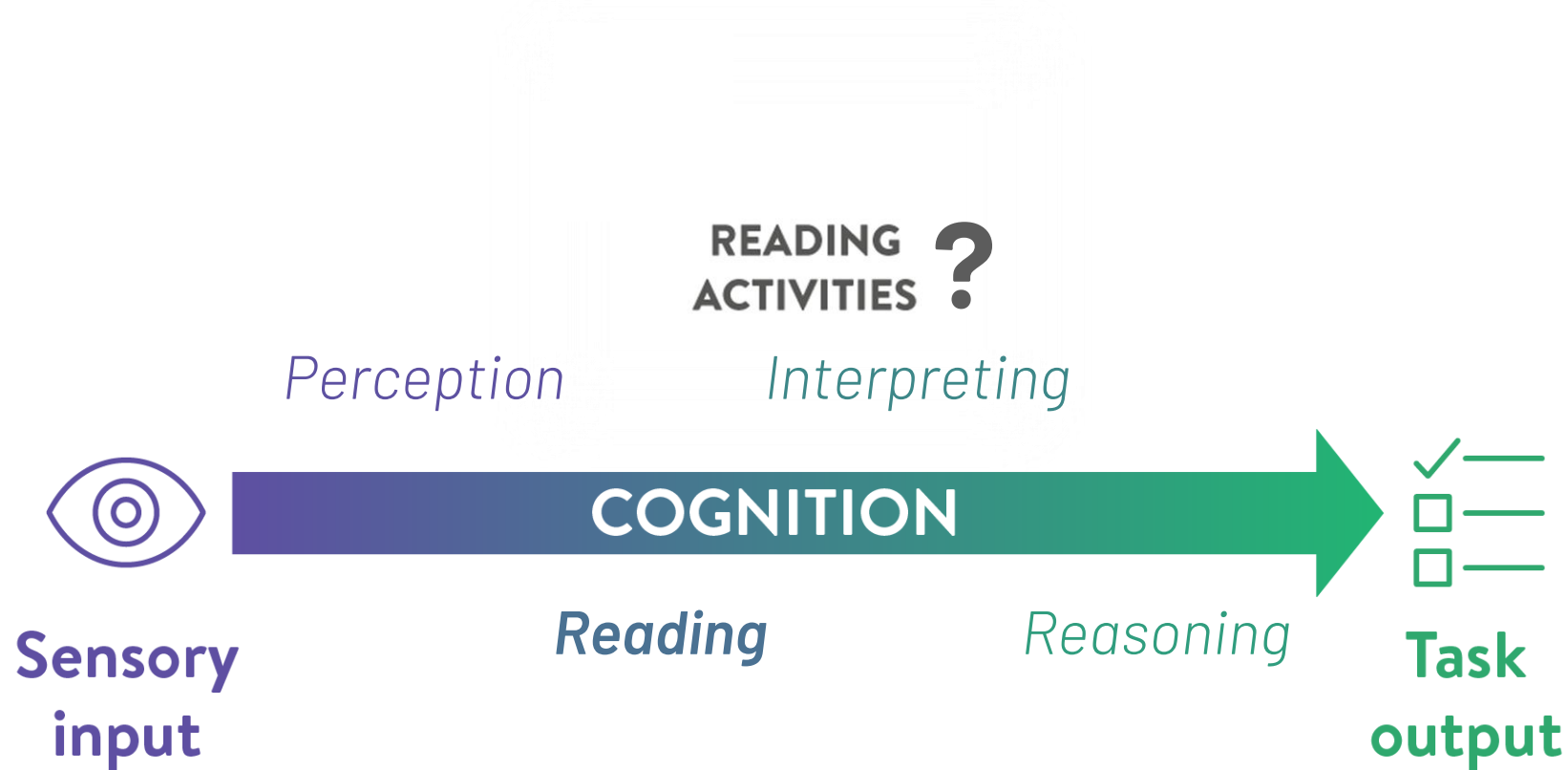
[github.com/AF-Cabouat/
PREVis-scales/](https://github.com/AF-Cabouat/PREVis-scales/)

Why seek subjective
evaluation of readability?

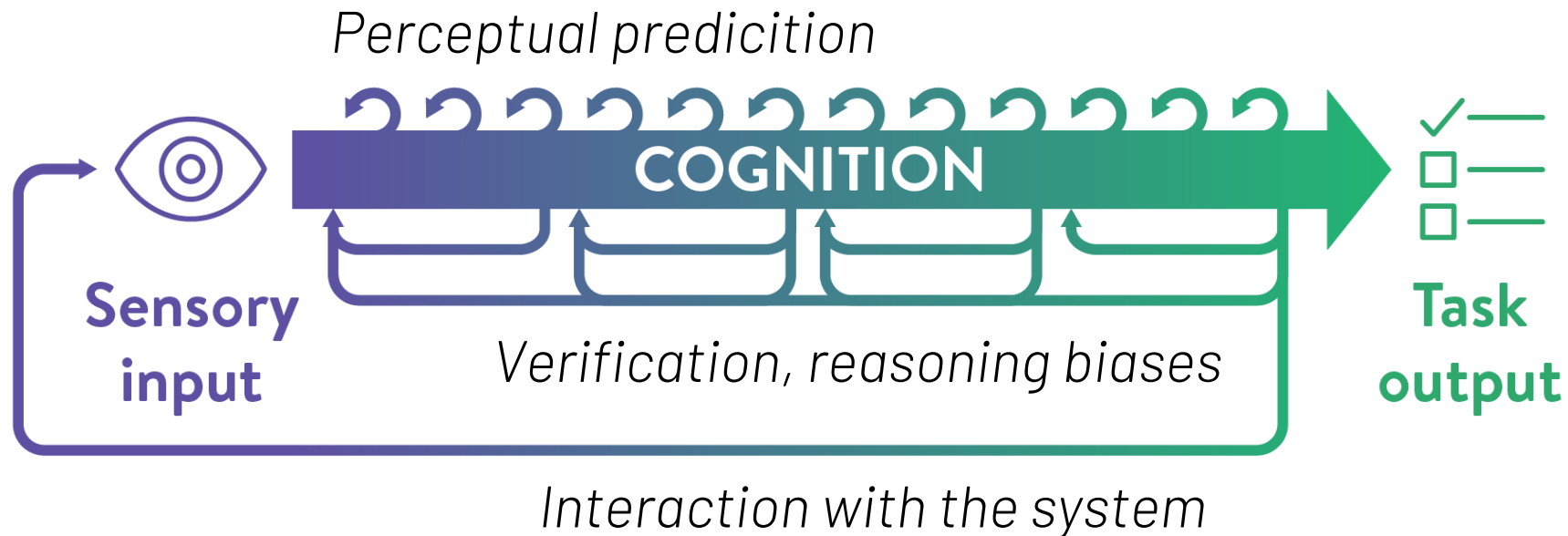
3 components influencing readability: the visualization, the task, and the reader



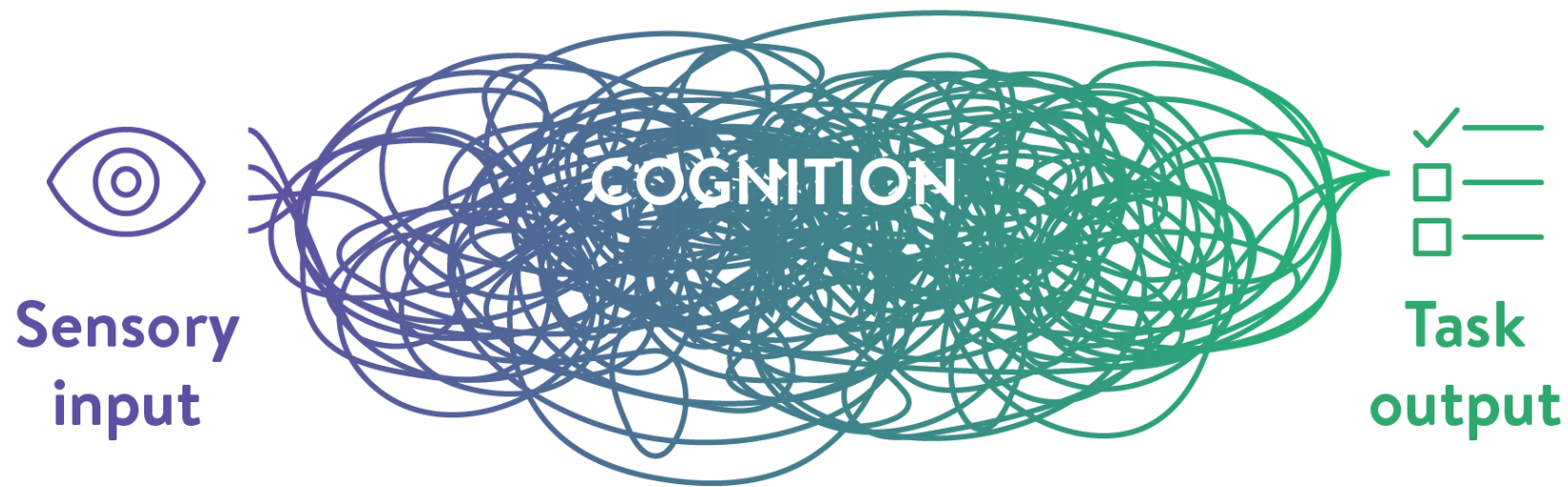
Viewer's cognitive processes in data visualization



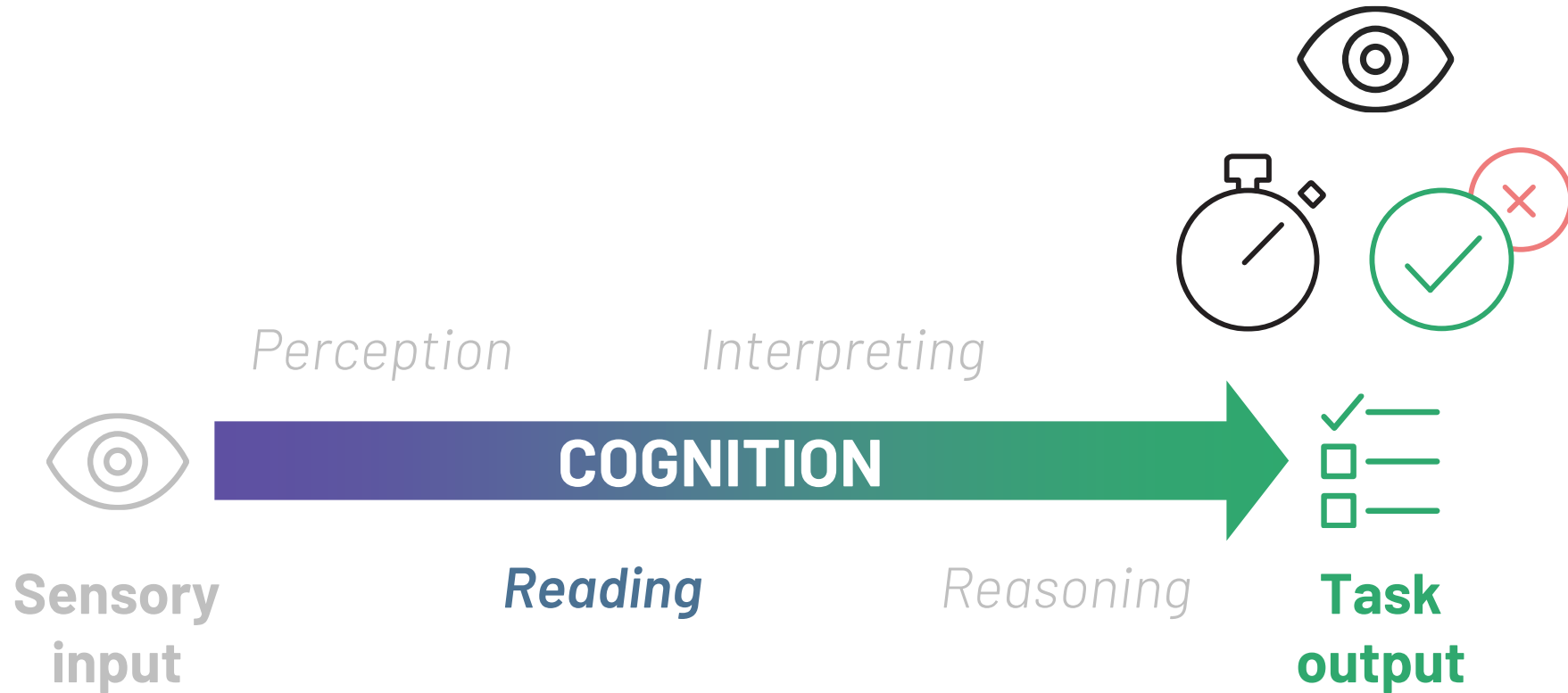
Viewer's cognitive processes in data visualization



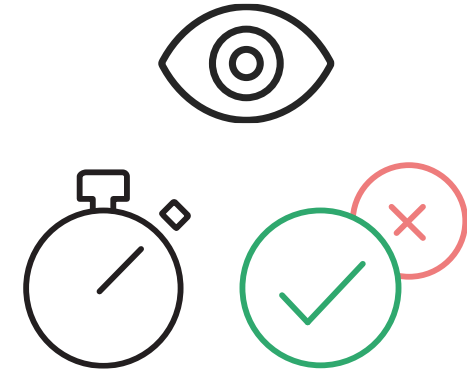
Viewer's cognitive processes in data visualization



Measuring readability?



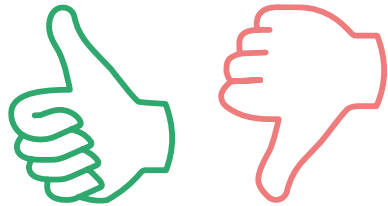
Measuring readability?



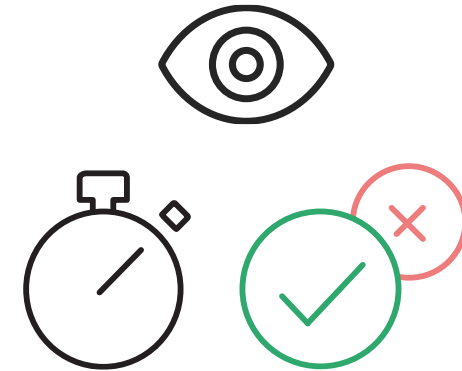
Behavioral
measurements

Can be influenced by other factors (reader's bias, interactive features...)

Measuring readability?



Subjective
evaluations



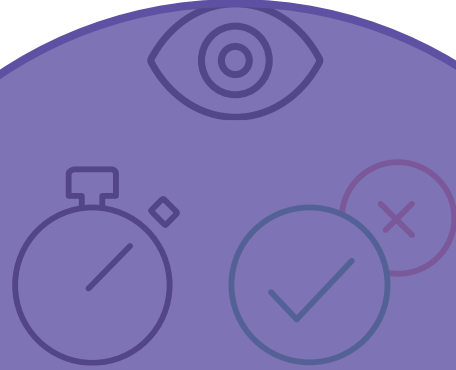
Behavioral
measurements

Can be influenced by
other factors (interactive
features, reader's bias...)

Studying cognitive activities

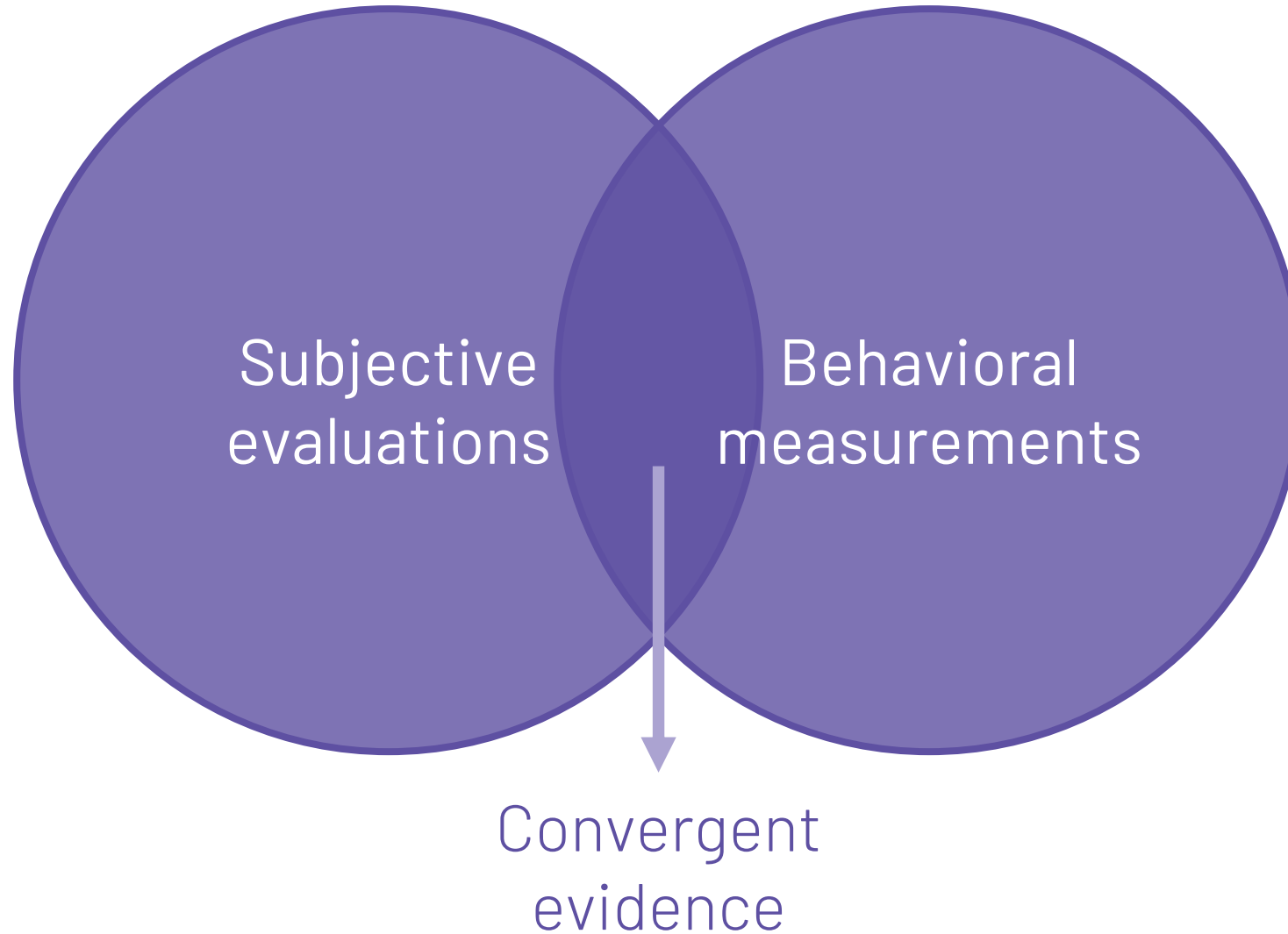


Subjective
evaluations



Behavioral
measurements

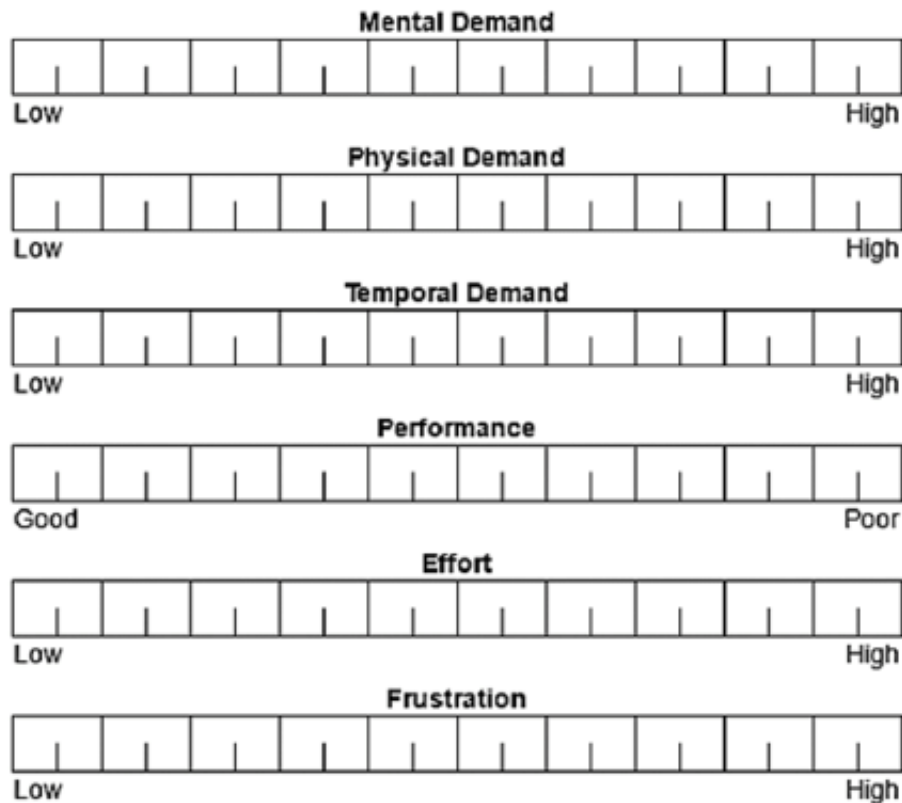
Studying cognitive activities



FUTURE WORK

User experience evaluation scales for visual analytics systems

NASA-TLX



System Usability Score

1. I think that I would like to use this system frequently.	1	2	3	4	5
2. I found the system unnecessarily complex.	1	2	3	4	5
3. I thought the system was easy to use.	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use this system.	1	2	3	4	5
5. I found the various functions in this system were well-integrated.	1	2	3	4	5
6. I thought there was too much inconsistency in this system.	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly.	1	2	3	4	5
8. I found the system very cumbersome to use.	1	2	3	4	5
9. I felt very confident using the system.	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system.	1	2	3	4	5

ADDITIONAL SLIDES

ITEM DEVELOPMENT

Diversity of Likert answer options

► Agreement (7 points)

The visualization looked cluttered.

Strongly Disagree Strongly Agree

Diversity of Likert answer options

- ▶ Agreement (7 points)
- ▶ Bipolar numbered agreement (5 points)

It was easy to understand the information displayed.

-2	-1	0	+1	+2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Disagree				Agree

Diversity of Likert answer options

- ▶ Agreement (7 points)
- ▶ Bipolar numbered agreement (5 points)
- ▶ Bipolar quality (7 points)

Is it easy or hard to interpret?

Very easy Very hard

The visualization was...

Well-organized Poorly-organized

Diversity of Likert answer options

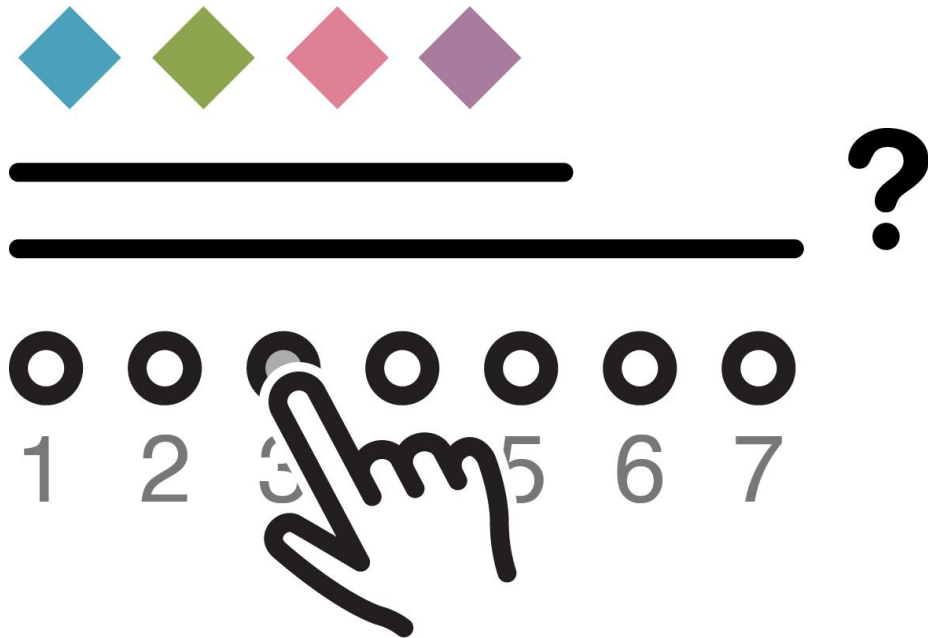
- ▶ Agreement (7 points)
- ▶ Bipolar numbered agreement (5 points)
- ▶ Bipolar quality (7 points)
- ▶ Bipolar statements (5 points)

Is it clear what information is being presented and why?

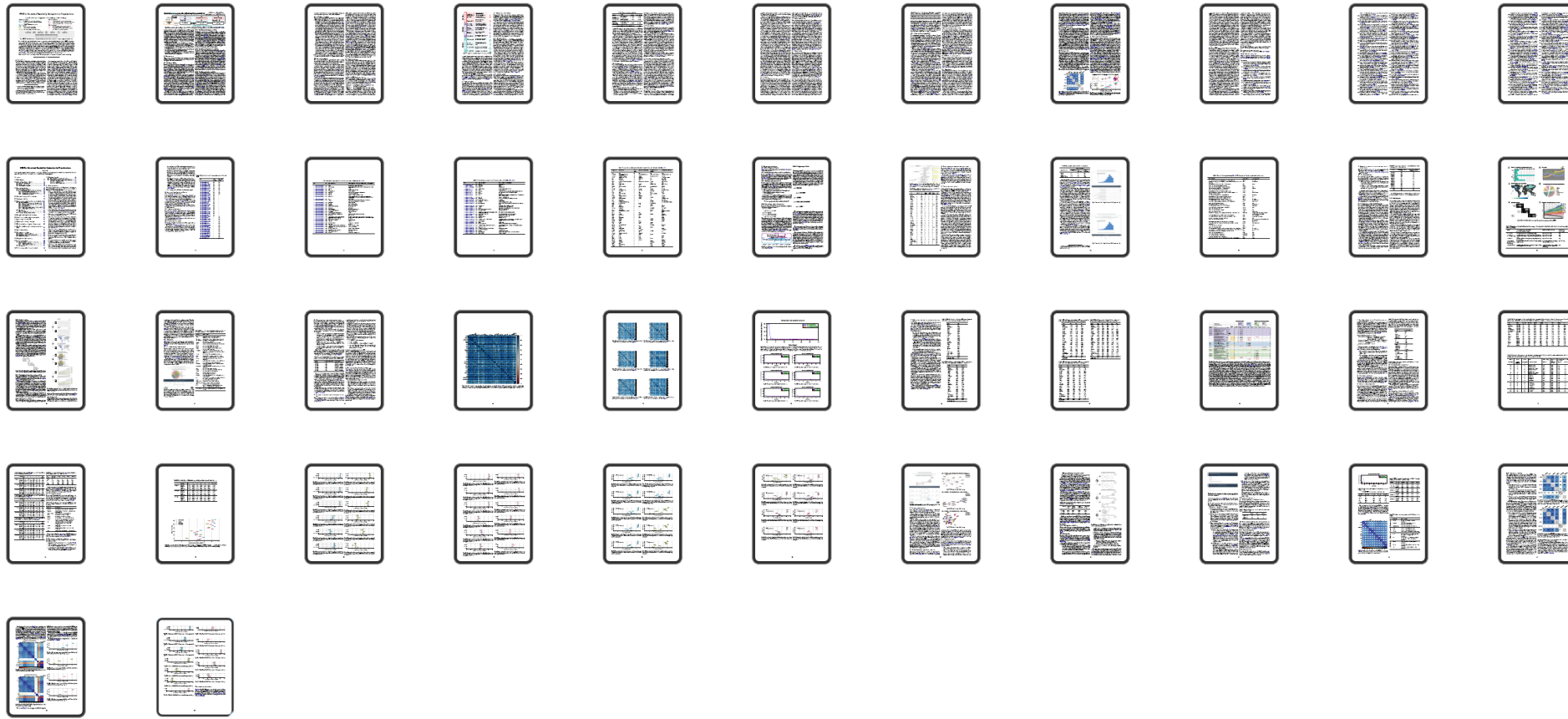
I am utterly confused It makes perfect sense

EXPLORATORY STUDY

PREVis: a standardized questionnaire to evaluate perceived readability



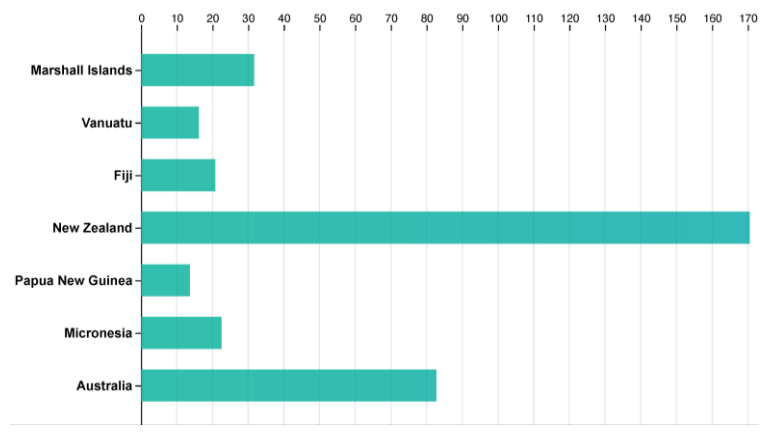
Development and validation methodology



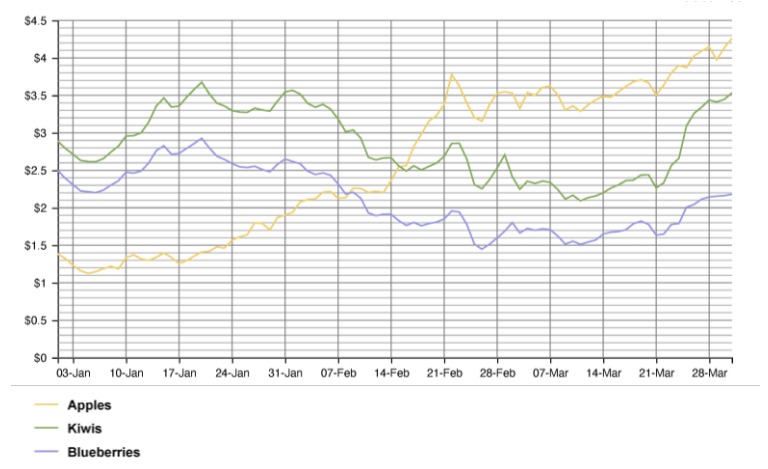
Paper

Appendices
(35 pages)

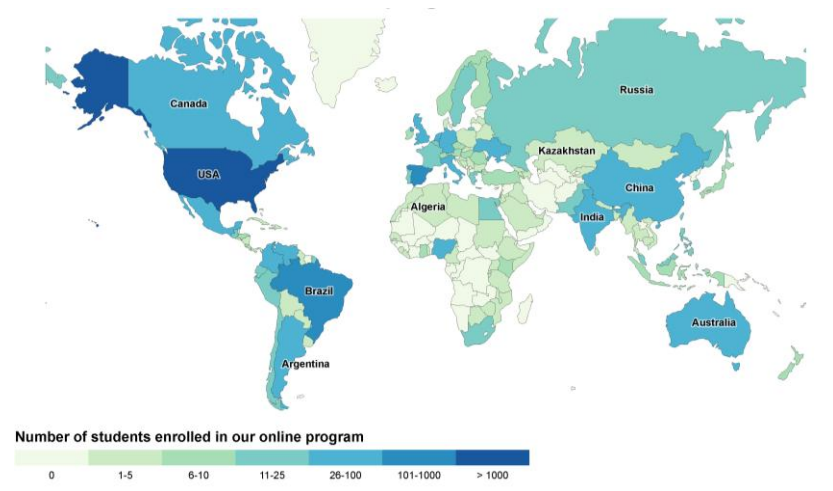
A



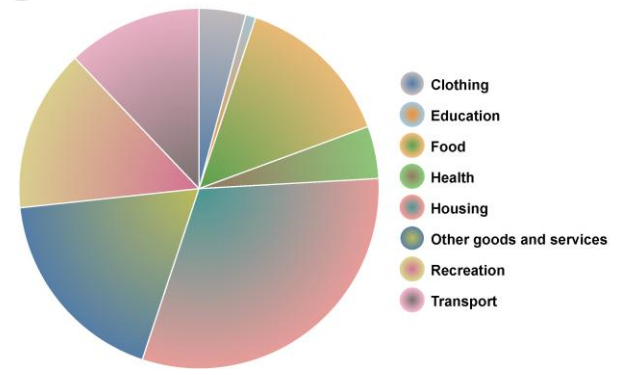
B



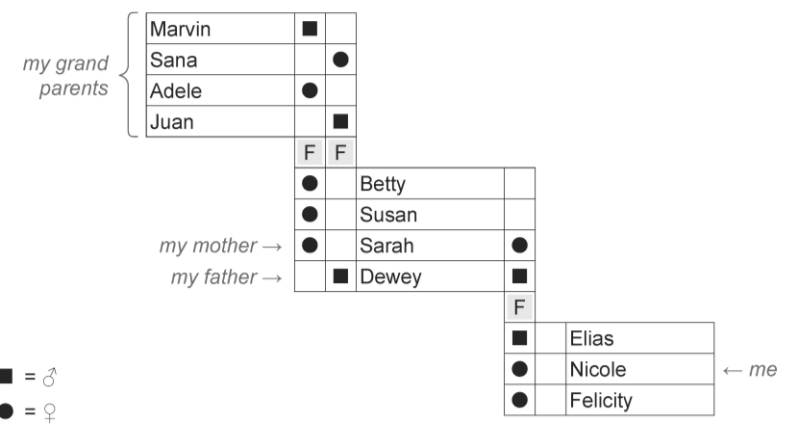
C



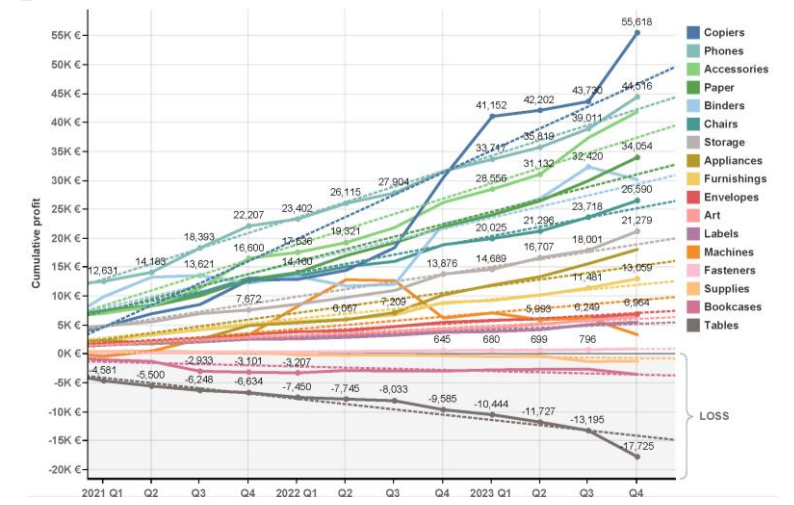
D



E



F



Exploratory survey (N=280 to 300 participants, independent samples)

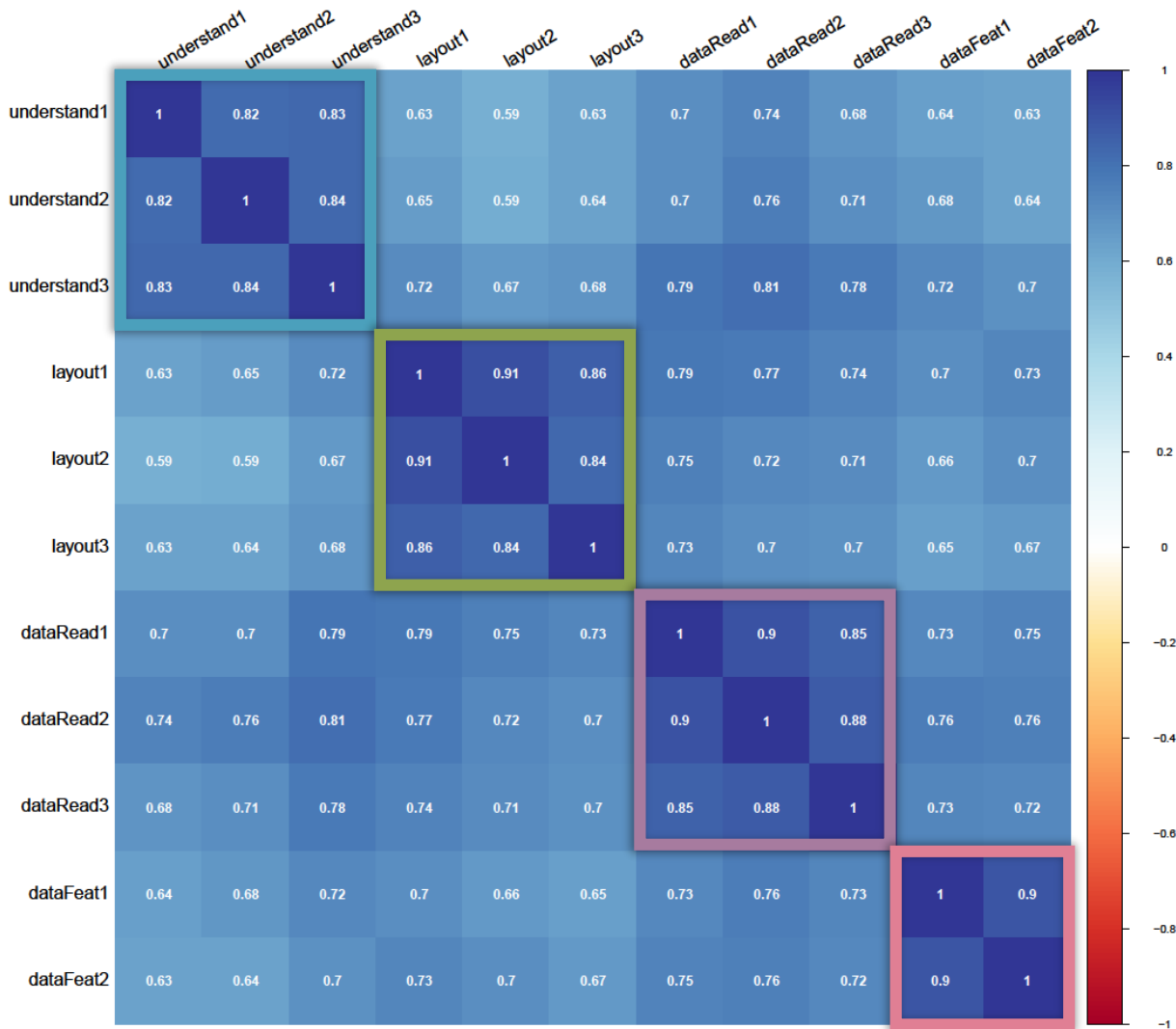
Stimuli	Display: encodings, data attributes <i>(on a canvas of max 800 x 500px)</i>	Individual: idiom familiarity	Task difficulty
A - Bar chart	2 visual variables, few entities (7 bars), clear layout	Very familiar/intuitive	Easy
B - Line chart	3 visual variables, few entities (3 lines), cluttered layout (grid)	Very familiar/intuitive	Easy
C - World choropleth map	4 visual variables, many entities (all countries), simple encodings	Very familiar/intuitive	Easy
D - Pie chart	2 visual variables, few entities (8 categories), messy encodings (colors, labels,)	Very familiar/intuitive	Easy
E - GeneaQuilts (family tree in a matrix style)	3 visual variables (4 visual encodings), reasonable amount of entities (11 people), simple encodings	Not familiar, mapping inference possible from labels with effort	- (not in VLAT)
F - Many lines chart	4 visual variables, many entities (18 categories x 2 types of lines), overplotting (too much information)	Somewhat familiar/intuitive (general trend dashed lines)	Easy

Reducing the number of questions

Exclusion criteria:

- ▶ Weak factor loading
- ▶ Cross-factors loadings
- ▶ Final selection on reliability metrics

	conceptual meaning >				understanding, meanings		visual clarity (layout quality)		reading the data		reading between the data
	Factor variable name in R >				understand		layout		dataRead		dataFeat
	Variance proportion explained >				0.43		0.25		0.25		0.07
variable name in R in full model	statement	item	Cross-loadings in Agg	factor rank in PA1 Agg	PA1 Agg	factor rank in PA2 Agg	PA2 Agg	factor rank in PA4 Agg	PA4 Agg	factor rank in PA3 Agg	PA3 Agg
understand1	It is obvious for me how to read this visualization	obvious	0	1	1.062337821						
understand2	I can easily understand the overall meaning of this data visualization	meanOverall	0	2	0.900123582						
understand3	I am confident in my understanding of this visualization	confid	0	3	0.88850186						
understand4	I can easily understand how the data is represented in this visualization	represent	0	4	0.882837581						
understand5	I can easily understand this visualization	understandEas	0	5	0.845610731						
understand6	I can quickly understand this visualization	understandQui	0	6	0.768301622						
understand7	I can easily understand what the different elements of the visualization mean	meanElem	0	7	0.584213463						
	I can easily answer some questions about the represented data with this visualization	answer	1	8	0.559170953				0.388855235		
	I don't feel lost trying to read this visualization	lost	1	9	0.53136155		0.362853201				
	I don't find this visualization confusing	confus	1	10	0.514550271		0.412462743				
	I don't find this visualization complex to read	complex	1	11	0.503407051		0.395441902				
	I don't find the presentation of the data difficult to decipher in this visualization	deciph	1	12	0.472499253		0.331068973				
	The representation of the data makes the information clear to me in this visualization	clearRepresent	0	13	0.401080111						
layout1	I don't find this visualization crowded	crowd	0			1	1.002355239				
layout2	I don't find this visualization messy	messi	0			2	0.97629751				
layout3	I don't find distracting parts in this visualization	distract	0			3	0.716152421				
layout4	I find this visualization well organized	organiz	0			4	0.601772217				
	I find this visualization simple to read	simpl	1		0.380395624	5	0.393556178				
dataFeat1	I find data features (for example, a minimum, or an outlier, or a trend) visible in this visualization	visibl	0							1	0.801035197
dataFeat2	I can clearly see data features (for example, a minimum, or an outlier, or a trend) in this visualization	see	0							2	0.745991576
dataRead1	I can easily find specific elements in this visualization	find	0					1	0.813425997		
	I can easily distinguish individual elements of the represented data (for example individual lines, or dots, or areas, or colors...)	distinguish	1				0.355062313	2	0.655425343		
dataRead2	I can easily identify relevant information in this visualization	identifi	0					3	0.603689212		
dataRead3	I can read data values from this visualization	valu	0					4	0.600135661		
dataRead4	I can easily retrieve information from this visualization	inform	0					5	0.599757226		
dataRead5	I find this visualization readable	readabl	0					6	0.505646165		
	I can easily read this visualization	read	1		0.333193092			7	0.484235378		
	This visualization effectively shows the data to me	effect	1		0.35791681			8	0.399574603		
	This visualization shows the data in a clear way for me	clearData	1				0.347569591	9	0.36121256		



◆ Understand subscale

understand1 It is **obvious** for me how to read this visualization

understand2 I can easily understand **how the data is represented** in this visualization

understand3 I can easily understand this visualization

◆ Layout subscale

layout1 I **don't** find this visualization **messy**

layout2 I **don't** find this visualization **crowded**

layout3 I **don't** find **distracting parts** in this visualization

◆ DataRead subscale

dataRead1 I can easily **find specific elements** in this visualization

dataRead2 I can easily **identify relevant information** in this visualization

dataRead3 I can easily **retrieve information** from this visualization

◆ DataFeat subscale

dataFeat1 I find data features (for example, a minimum, or an outlier, or a trend) **visible** in this visualization

dataFeat2 I can **clearly see** data features (for example, a minimum, or an outlier, or a trend) in this visualization

Ratings with PREVis scales on exploratory stimuli

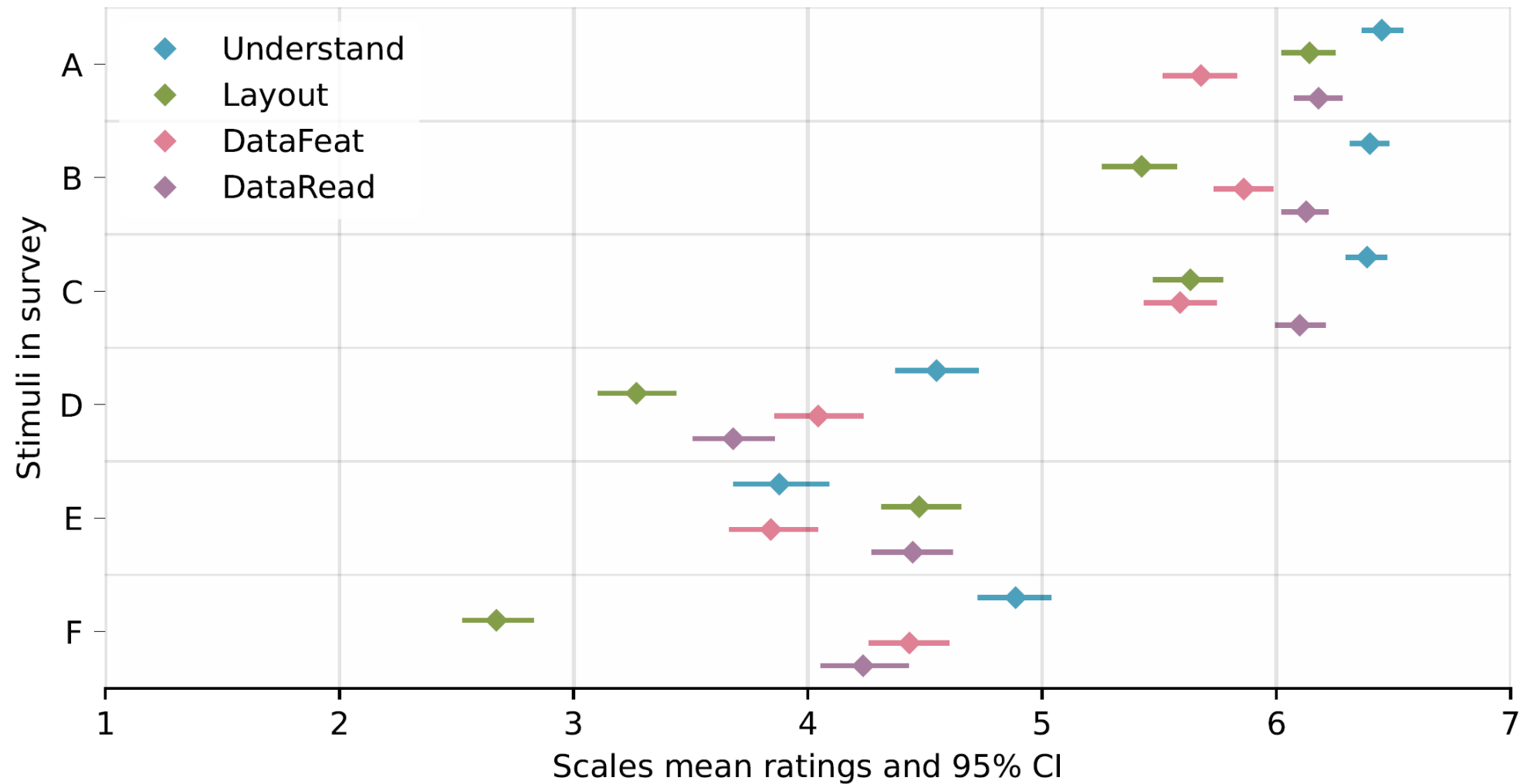
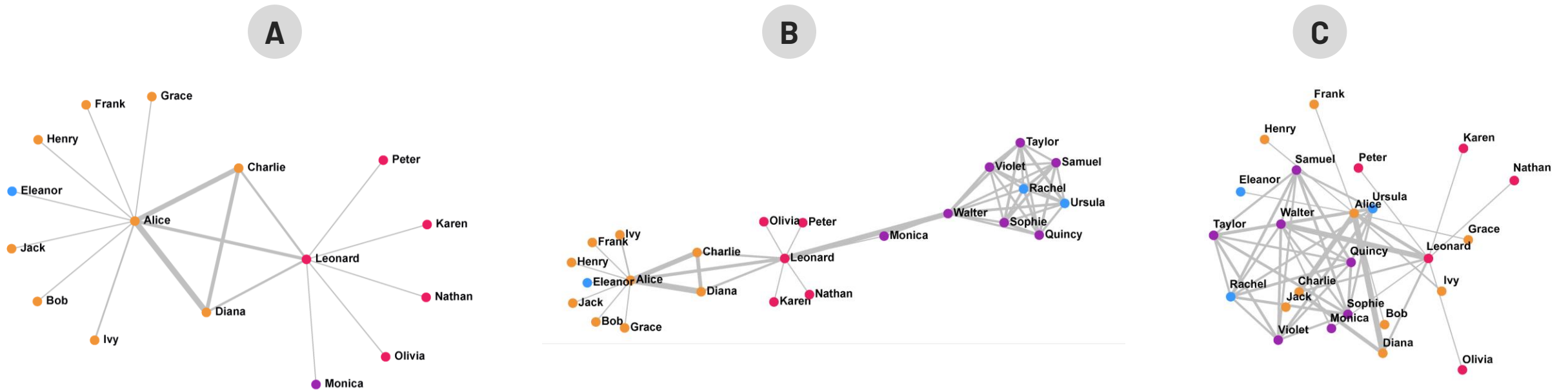


Fig. 29: Average ratings (from 1 = “Strongly disagree” to 7 = “Strongly agree”) using the four PREVis subscales \blacklozenge \blacklozenge \blacklozenge \blacklozenge on 6 visualizations of different readability: (A / B / C) > (D / E / F). For a given subscale, ratings and 95% CI do not overlap and fit the readability ranking.

VALIDATION STUDY

Validation study

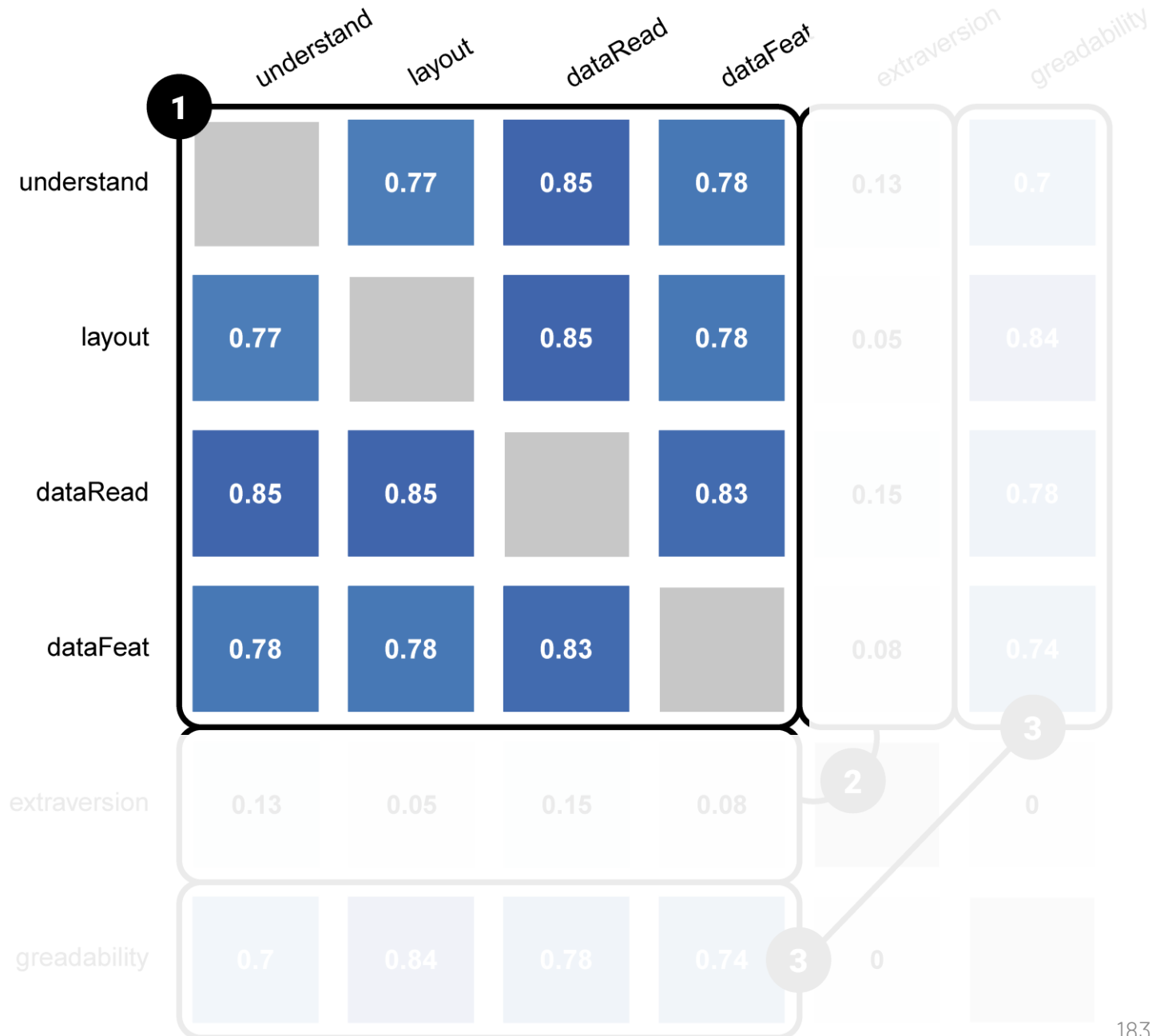


- ▶ N=147 on Prolific - All participants saw A, B, and C in random order
- ▶ People answered 2 reading questions then 11 PREVis questions
- ▶ Exclusion criteria: 1 comprehension check and 1 attention check

Results: construct validity

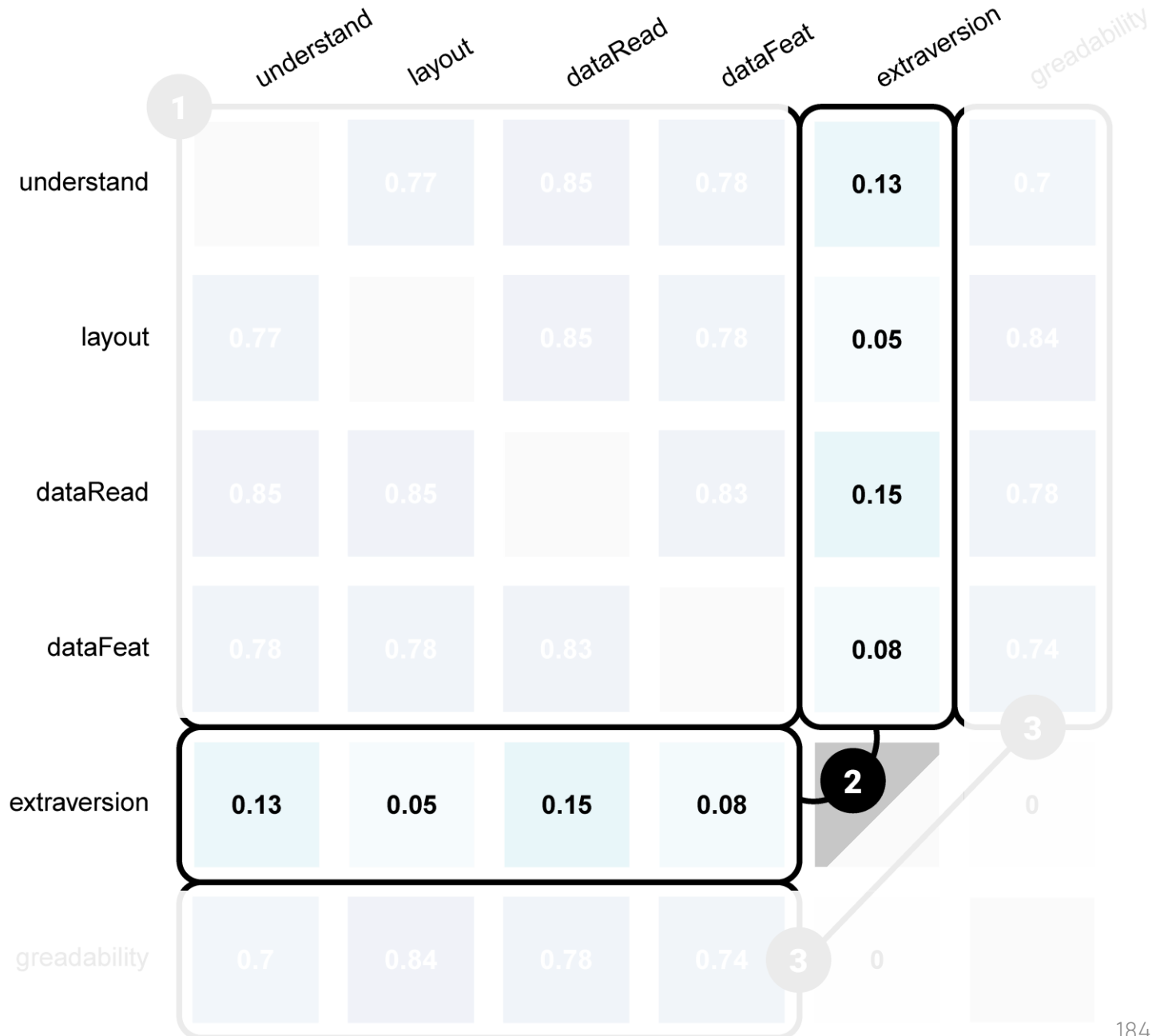
1. Inter-scales reliability:

each scale should highly correlate with the others, because they are all related to the same underlying construct: *readability*



Results: construct validity

1. Inter-scales reliability
2. **Discriminant validity:**
the results should not correlate with measures of an unrelated construct (extraversion scale from the Big Five Inventory)



Results: construct validity

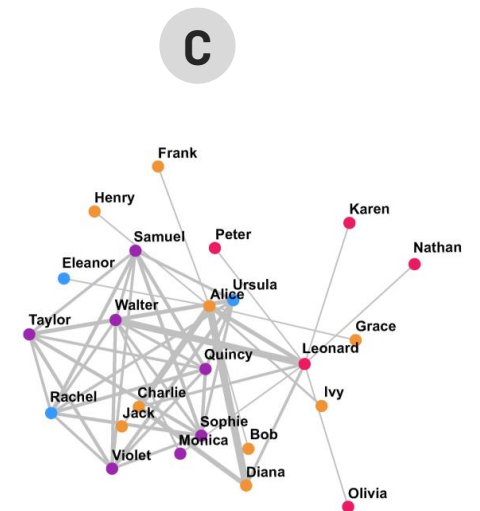
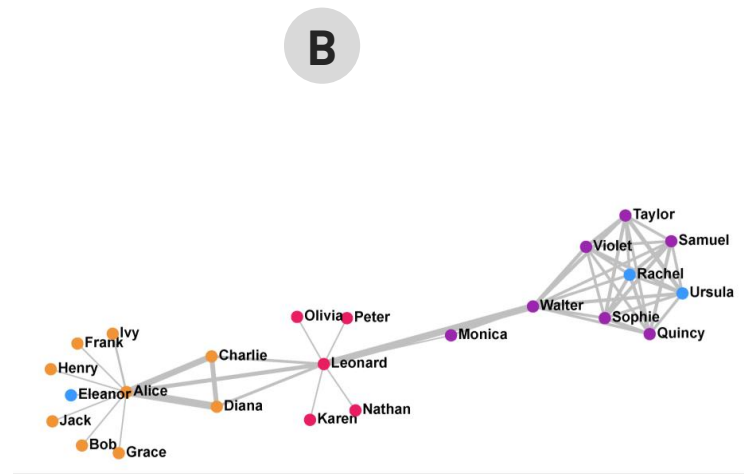
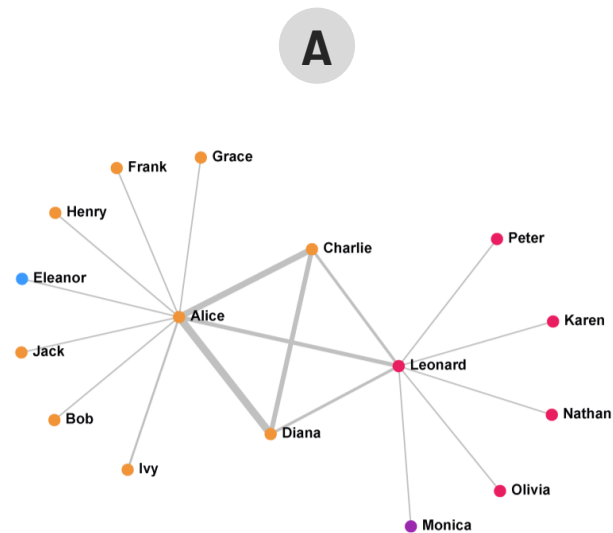
1. Inter-scales reliability
2. Discriminant validity
3. **Convergent validity:**
the results should
positively correlate
with independent
measures that also
relate to readability

Results: construct validity

1. Inter-scales reliability
2. Discriminant validity
3. **Convergent validity**

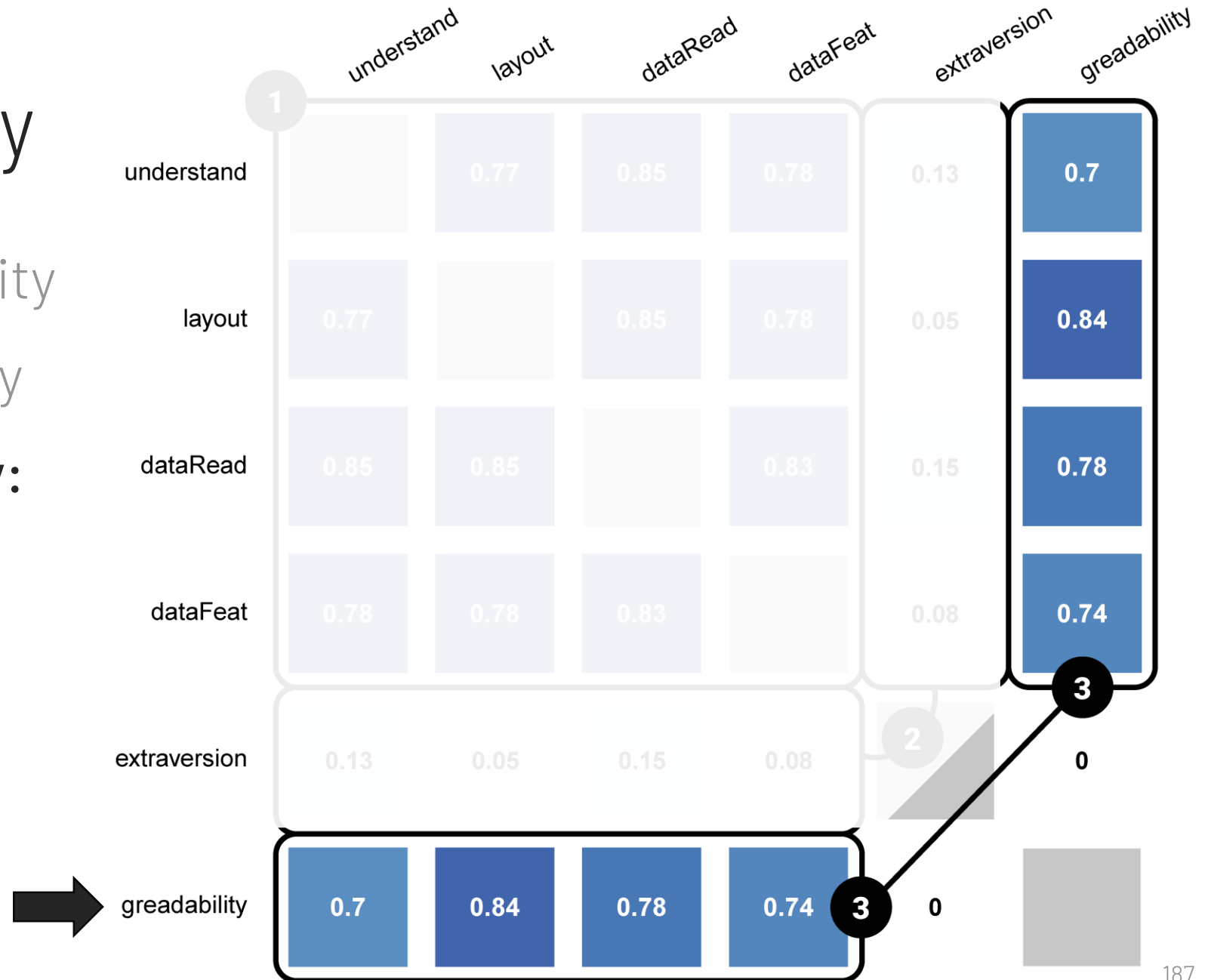
Computed reference layout measures from greadability.js (Gove, 2018)

Stimulus A	0.872
Stimulus B	0.786
Stimulus C	0.679



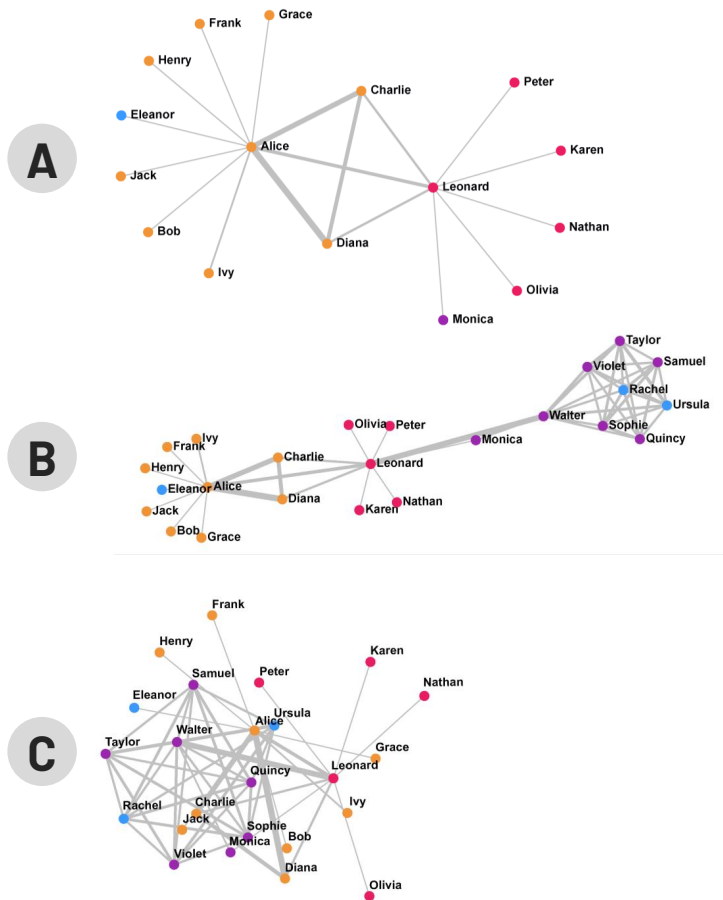
Results: construct validity

1. Inter-scales reliability
2. Discriminant validity
3. **Convergent validity:**
the results should positively correlate with independent measures that also relate to readability

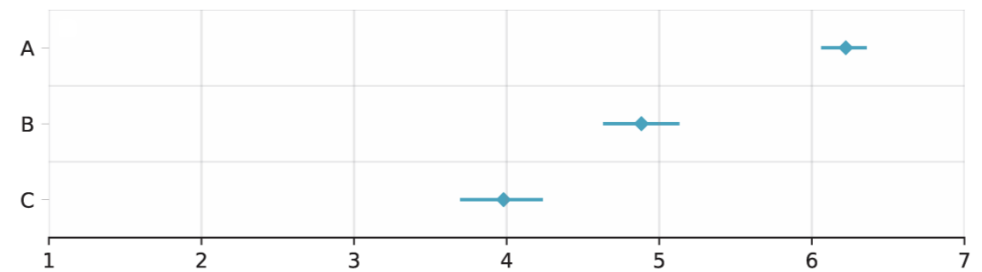


Results: ratings

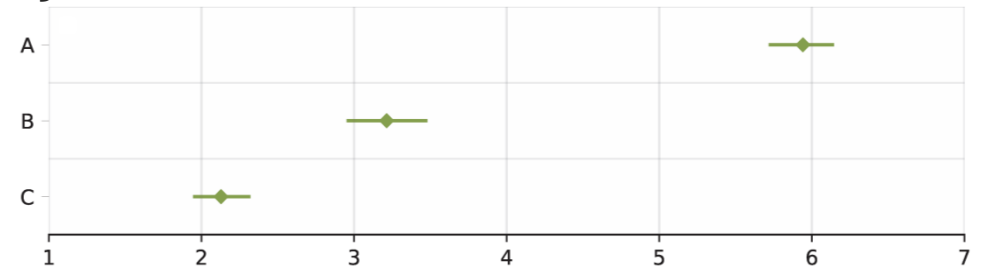
Point estimates with 95% CI



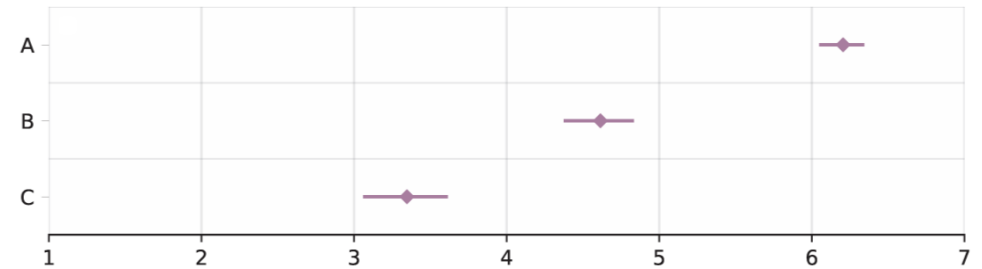
◆ Understand obvious + represent + understand



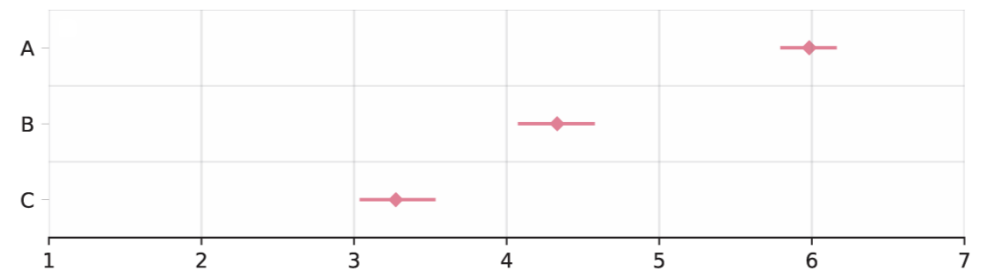
◆ Layout messy + crowded + distracting



◆ DataRead find + identify + inform



◆ DataFeat visible + clear



◆◆◆◆ PREVis: implementation guidelines

Why?

When?

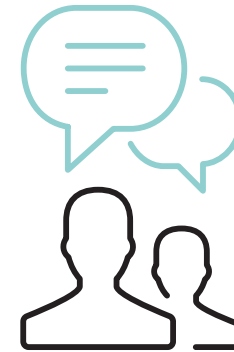
How?

◆◆◆◆ PREVis: implementation guidelines

Why? To complement other measures



Task performance



Qualitative interviews

◆◆◆◆ PREVis: implementation guidelines

Why? To complement other measures

When? After people used the visualization(s)

◆◆◆◆ PREVis: implementation guidelines

Why? To complement other measures

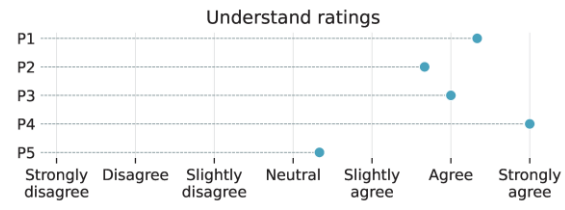
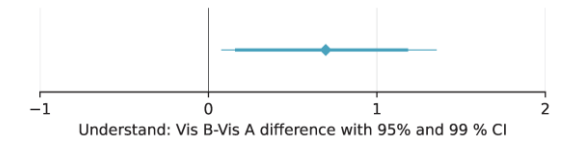
When? After people used the visualization(s)

How? Ensure reliability of your results: read the *** manual!

The screenshot displays the PREVis manual page, titled "PREVis Perceived Readability Evaluation in Visualization". It includes sections for "Understand subscale", "Layout subscale", "DataFeat subscale", and "DataRead subscale". Each section lists items with their corresponding subscale and a brief description. For example, under "Understand subscale", items include "obvious", "represent", and "understand". The "Layout subscale" items are "messy", "crowd", and "extract". The "DataFeat subscale" items are "visible" and "see". The "DataRead subscale" items are "find", "identify", and "information". Below the list, there is a Likert scale for the statement "I can easily understand this visualization" with options from "Strongly disagree" to "Strongly agree". A text box for "I don't know / Not applicable" is also present. The page footer includes a license notice and a page number "1".

- ▶ Question order and randomization
- ▶ Questions wording
- ▶ Answer options and labels
- ▶ Study cases and data analysis choices

Available materials



This document is provided on www.nature.com/articles/s41598-024-58111-1 under the CC-BY license, as a supplemental material for our IEEE VIS 2024 paper on the PREVis scale.

PREVis

Perceived Readability Evaluation in Visualization

Questionnaire items and implementation recommendations

- Understand subscale**
 - understand** I can easily understand how the data is represented in this visualization
 - understand** I can easily understand this visualization
- Layout subscale**
 - layout** I don't find this visualization messy
 - layout** I don't find this visualization crowded
 - layout** I don't find distracting parts in this visualization
- DataFeat subscale**
 - dataFeat** I find data features (for example, a minimum, or an outlier, or a trend) visible in this visualization
 - dataFeat** I can clearly see data features (for example, a minimum, or an outlier, or a trend) in this visualization
- DataRead subscale**
 - dataRead** I can easily find specific elements in this visualization
 - dataRead** I can easily identify relevant information in this visualization
 - dataRead** I can easily retrieve information from this visualization

Regarding the visualization items, I can agree or disagree with the following statement

I can easily understand this visualization

Strongly disagree Disagree Slightly disagree Neutral Slightly agree Agree Strongly agree

Agree Not applicable if you wish, feel free to substitute below

(An example of answer options for one item and how to present them.)

HOW TO USE PREVis

In this document, we provide practical considerations for effective use, implementation and analysis of PREVis measurements. We gathered the following content to guide researchers.

FAQ

- How many participants do I need?
- Is the scale appropriate for my specific participants?
- When do I analyze PREVis?
- Is it worth using the PREVis?
- Are the 4 subscales independent in PREVis?
- How do I calculate the mean?

Study protocols examples

- Study: visualization as a visualization
- Study: visualization as a visualization
- Study: visualization as a visualization

PDF forms for PREVis

- Simple questionnaire (1 page)
- Detailed questionnaire (2 pages)

PREVis: Perceived Readability Evaluation for Visualizations



Fig. 1 The six visualizations we presented in our survey to develop the PREVis instrument. How readable do you find them?

Abstract—We developed and validated an instrument to measure the perceived readability in data visualization: PREVis. Researchers and practitioners can easily use this instrument as part of their evaluations to compare the perceived readability of different visual data representations. Our instrument can complement results from controlled experiments on user task performance or provide additional data during in-depth qualitative work such as design iterations when developing a new technique. Although readability is recognized as an essential quality of data visualizations, so far there has not been a unified definition of the construct in the context of visual representations. As a result, researchers often lack guidance for determining how to ask people to rate their perceived readability of a visualization. To address this issue, we engaged a rigorous process to develop the first validated instrument targeted at the subjective readability of visual data representations. Our final instrument consists of 11 items across 4 dimensions: understandability, layout clarity, readability of data patterns, and readability of data features. We provide the implementation as a document with implementation guidelines on our GitHub. Beyond the instrument, we conducted a discussion of how researchers have previously assessed visualization readability, and an analysis of the factors underlying perceived readability in visual data representations.

Index Terms—Readability, validated instrument, perception, user experiments, empirical methods, methodology.

1 INTRODUCTION

When looking at examples of data visualizations, it is intuitively clear that some are easier to read than others. For many data analysis use cases, good readability will directly affect the amount of a visual representation of data for the viewer. As such, readability is a basic quality criterion in data visualization [1]. One of the fundamental challenges in studying the readability of data visualizations, however, is that the concept of reading and readability is held to such knowledge. The terms are often used in scientific writing without clear definitions of what they specifically mean in the context of data visualization—recalling Kieras’s “empire built on sand” [2].

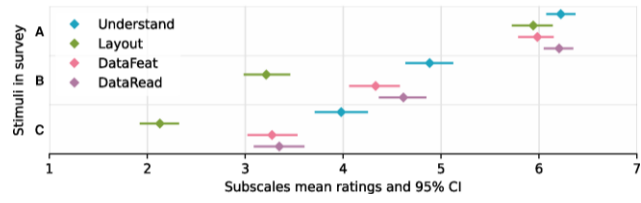
Readability is best to be broadly defined as “the quality of being easy and enjoyable to read” [18]. It applies to letters and words as well as entire books. Lingens later developed hundreds of formulas to analyze the readability of text [4], but this approach fails to take into account characteristics of readers. Hence, readability is better explained as a function of the interaction between the properties of text and the characteristics of readers [4]. Moreover, new work is analyzing readability based on cognitive theories. Such an approach may also be suitable to explore the readability of visual representations of data.

As we discuss in more detail below, a few definitions of “readability” exist in the visualization domain, yet they do not fully overlap. Our PREVis is in line with what is most common in the field, but we also address what is not. In particular, we address the lack of a definition of what “reading” a data visualization is as a cognitive activity. Cognitive processes in visualization range from

low-level visual perception [8] to high-level activities such as data exploration [10], insight and knowledge generation [9, 30], sense-making of external visualizations [26], or decision-making [22]. Current cognitive models of visualization comprehension [25, 49, 72] provide important theoretical grounding to explain how people process information from visual data representations, the models, however, do not specify the boundaries of “reading” within the cognitive continuum.

Our work is based on the fundamental premise that readability is a quality dependent on quality elements in data visualizations. As such, it requires formal definition and empirically verified methods to study it. In this paper, we present the development and validation of PREVis questionnaire. PREVis is a reliable instrument that allows respondents to rate how readable they find a static data visualization across 4 dimensions: layout clarity, ease of understanding, ease of reading data features, and ease of reading data values. During the development process, we also had to take first steps in clarifying what readability means in data visualizations. This clarification is important because discrepancies in the use of terminology pose issues of comparability and reliability of empirical findings. In particular, we observed that researchers who asked participants to rate the readability of visualization often do not use a wide variety of terms and answer options. Our PREVis tool addresses this problem because we followed well-established methodologies in scale development [18, 24]. Developing a valid scale requires researchers to review existing literature and collect experts’ opinions on how to phrase the scale’s items. In this way, it has an

	understand	layout	dataRead	dataFeat	extraversion	readability
understand	1	0.77	0.85	0.78	0.13	0.7
layout	0.77		0.85	0.78	0.05	0.84
dataRead	0.85	0.85		0.83	0.15	0.78
dataFeat	0.78	0.78	0.83		0.08	0.74
extraversion	0.13	0.05	0.15	0.08	2	0
readability	0.7	0.84	0.78	0.74	0	3



- ▶ Full paper with appendices
- ▶ Pre-registrations
- ▶ Research log

- ▶ Code and data to reproduce our work

- ▶ Implementation companion guide
- ▶ Code and example data to help you produce different figures



ITEM DEVELOPMENT

Diversity of Likert answer options

▶ Agreement (7 points)

The visualization looked cluttered.

Strongly Disagree Strongly Agree

Diversity of Likert answer options

- ▶ Agreement (7 points)
- ▶ Bipolar numbered agreement (5 points)

It was easy to understand the information displayed.

-2	-1	0	+1	+2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Disagree				Agree

Diversity of Likert answer options

- ▶ Agreement (7 points)
- ▶ Bipolar numbered agreement (5 points)
- ▶ Bipolar quality (7 points)

Is it easy or hard to interpret?

Very easy Very hard

The visualization was...

Well-organized Poorly-organized

Diversity of Likert answer options

- ▶ Agreement (7 points)
- ▶ Bipolar numbered agreement (5 points)
- ▶ Bipolar quality (7 points)
- ▶ Bipolar statements (5 points)

Is it clear what information is being presented and why?

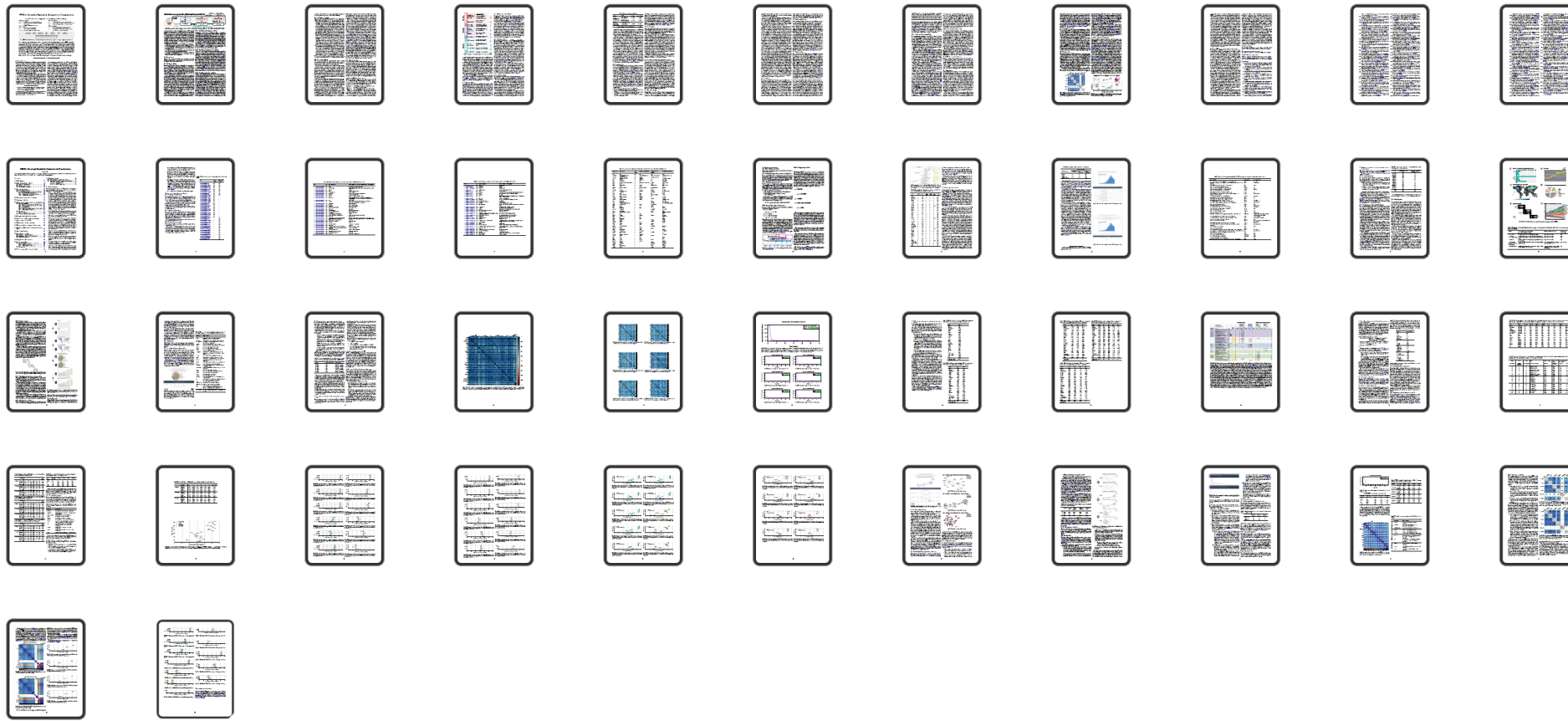
I am utterly confused It makes perfect sense

EXPLORATORY STUDY

PREVis: a standardized questionnaire to evaluate perceived readability



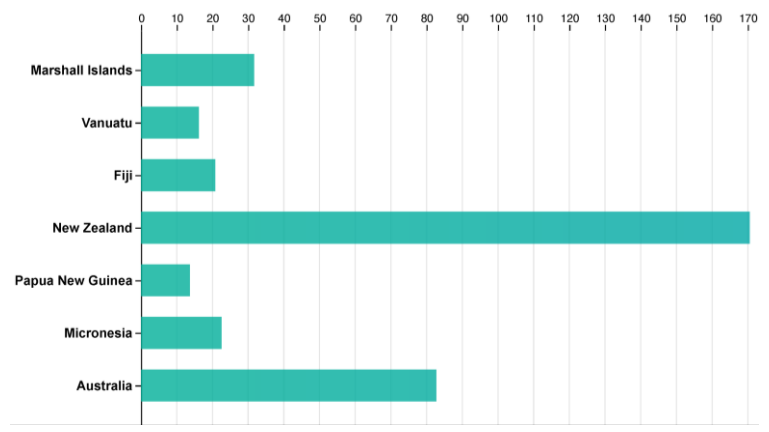
Development and validation methodology



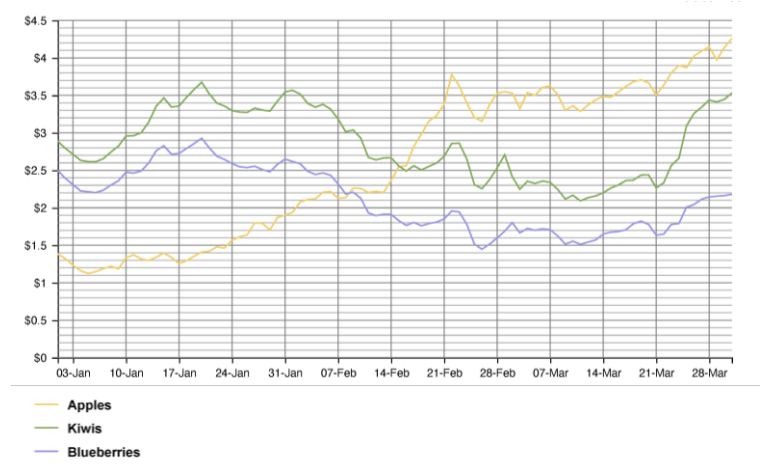
Paper

Appendices
(35 pages)

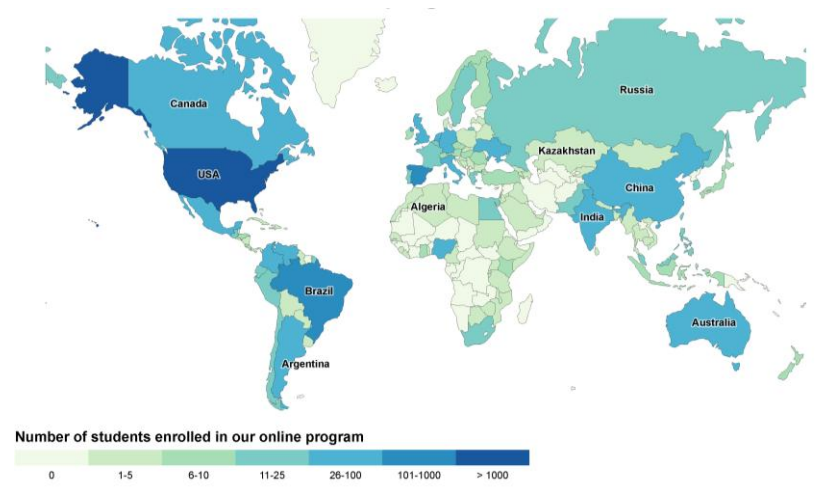
A



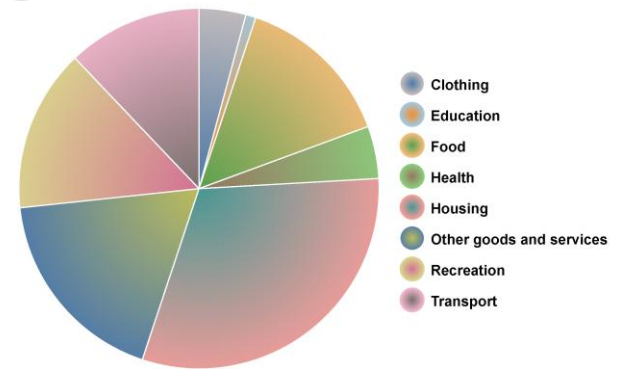
B



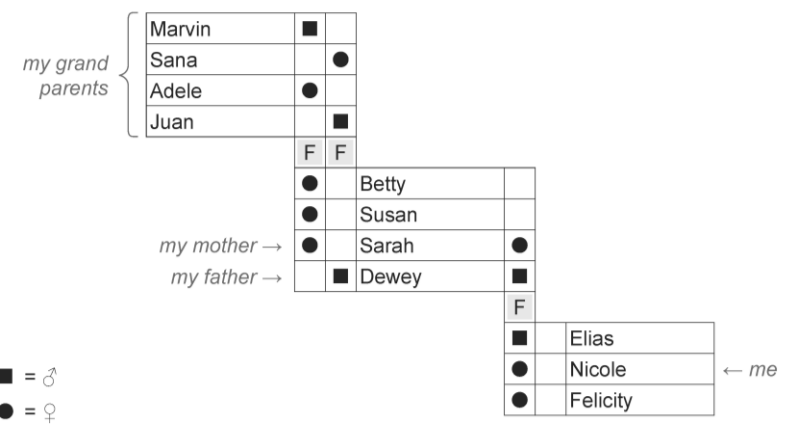
C



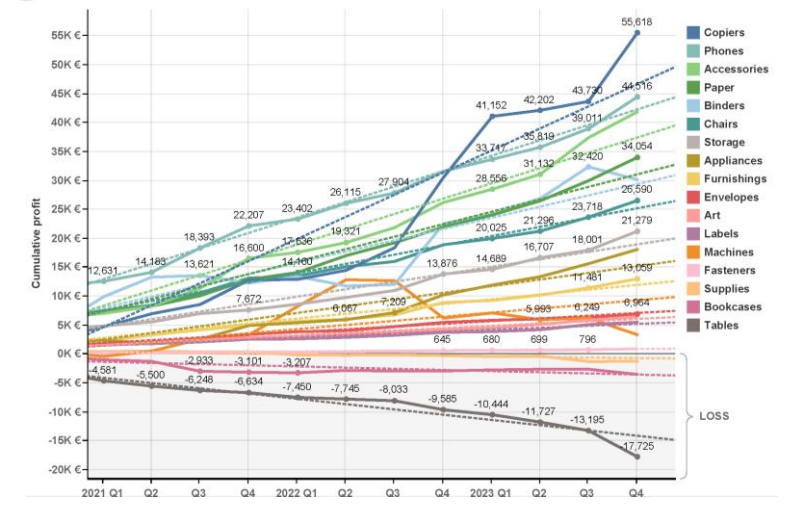
D



E



F



Exploratory survey (N=280 to 300 participants, independent samples)

Stimuli	Display: encodings, data attributes <i>(on a canvas of max 800 x 500px)</i>	Individual: idiom familiarity	Task difficulty
A - Bar chart	2 visual variables, few entities (7 bars), clear layout	Very familiar/intuitive	Easy
B - Line chart	3 visual variables, few entities (3 lines), cluttered layout (grid)	Very familiar/intuitive	Easy
C - World choropleth map	4 visual variables, many entities (all countries), simple encodings	Very familiar/intuitive	Easy
D - Pie chart	2 visual variables, few entities (8 categories), messy encodings (colors, labels,)	Very familiar/intuitive	Easy
E - GeneaQuilts (family tree in a matrix style)	3 visual variables (4 visual encodings), reasonable amount of entities (11 people), simple encodings	Not familiar, mapping inference possible from labels with effort	- (not in VLAT)
F - Many lines chart	4 visual variables, many entities (18 categories x 2 types of lines), overplotting (too much information)	Somewhat familiar/intuitive (general trend dashed lines)	Easy

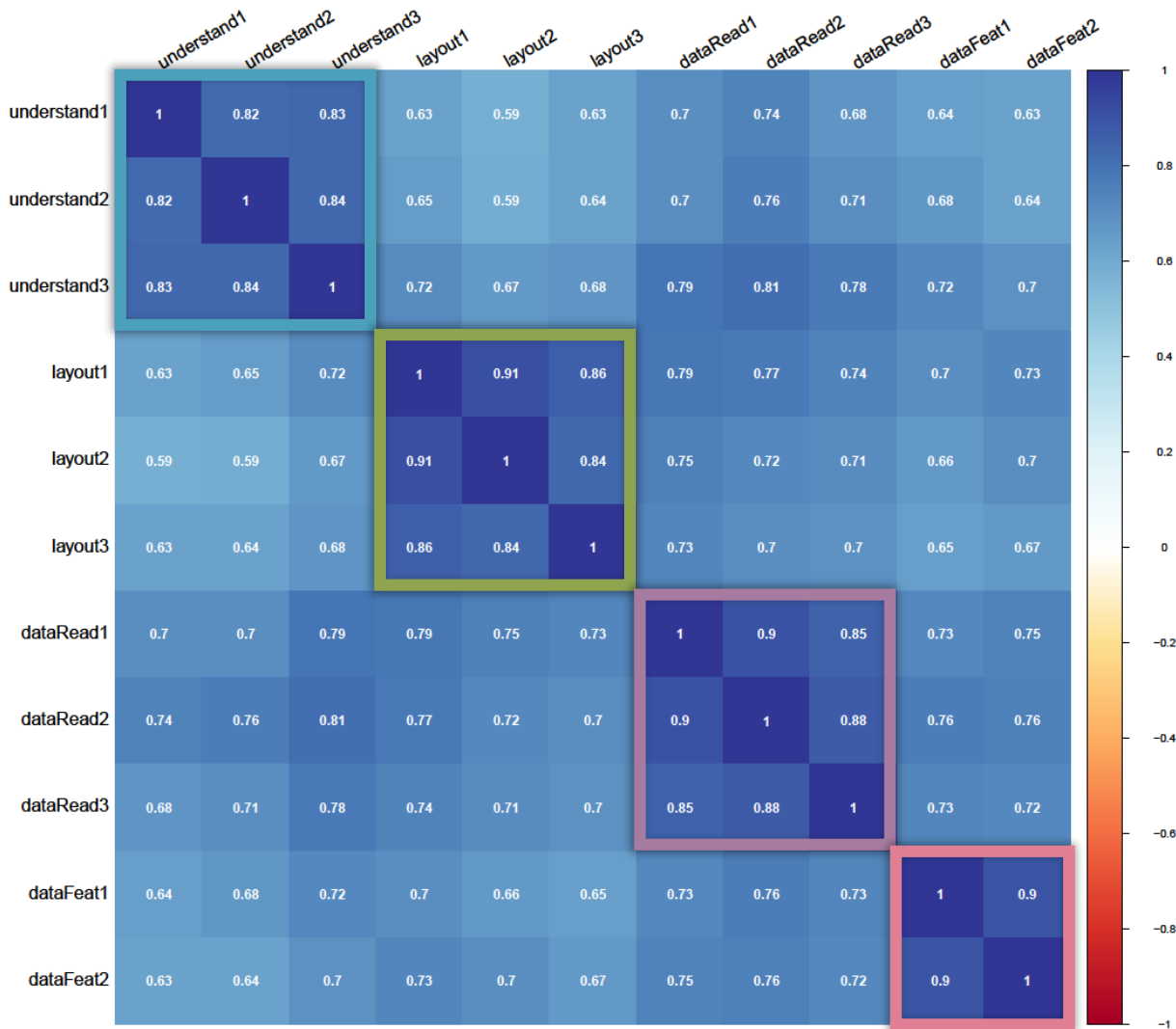
Note: there is a large number of terms in the literature for encodings taxonomies: visual variable and visual channel refer more to perceptual channels. Visual encodings is a broader term that also encompasses marks and possibly even idioms. « *Encoding Channels: position (X/latitude, Y/longitude), length, angle, area, texture, shape, color saturation, color hue, orientation, column, row. Mark Types: point, line, area-circle, area-rect, area-arc, area-other, text, geoshape, box-plot.* » (Zeng et Battle, 2023)

Reducing the number of questions

Exclusion criteria:

- ▶ Weak factor loading
- ▶ Cross-factors loadings
- ▶ Final selection on reliability metrics

	conceptual meaning >				understanding, meanings		visual clarity (layout quality)		reading the data		reading between the data
	Factor variable name in R >				understand		layout		dataRead		dataFeat
	Variance proportion explained >				0.43		0.25		0.25		0.07
variable name in R in full model	statement	item	Cross-loadings in Agg	factor rank in PA1 Agg	PA1 Agg	factor rank in PA2 Agg	PA2 Agg	factor rank in PA4 Agg	PA4 Agg	factor rank in PA3 Agg	PA3 Agg
understand1	It is obvious for me how to read this visualization	obvious	0	1	1.062337821						
understand2	I can easily understand the overall meaning of this data visualization	meanOverall	0	2	0.900123582						
understand3	I am confident in my understanding of this visualization	confid	0	3	0.88850186						
understand4	I can easily understand how the data is represented in this visualization	represent	0	4	0.882837581						
understand5	I can easily understand this visualization	understandEas	0	5	0.845610731						
understand6	I can quickly understand this visualization	understandQui	0	6	0.768301622						
understand7	I can easily understand what the different elements of the visualization mean	meanElem	0	7	0.584213463						
	I can easily answer some questions about the represented data with this visualization	answer	1	8	0.559170953				0.388855235		
	I don't feel lost trying to read this visualization	lost	1	9	0.53136155		0.362853201				
	I don't find this visualization confusing	confus	1	10	0.514550271		0.412462743				
	I don't find this visualization complex to read	complex	1	11	0.503407051		0.395441902				
	I don't find the presentation of the data difficult to decipher in this visualization	deciph	1	12	0.472499253		0.331068973				
	The representation of the data makes the information clear to me in this visualization	clearRepresent	0	13	0.401080111						
layout1	I don't find this visualization crowded	crowd	0			1	1.002355239				
layout2	I don't find this visualization messy	messi	0			2	0.97629751				
layout3	I don't find distracting parts in this visualization	distract	0			3	0.716152421				
layout4	I find this visualization well organized	organiz	0			4	0.601772217				
	I find this visualization simple to read	simpl	1		0.380395624	5	0.393556178				
dataFeat1	I find data features (for example, a minimum, or an outlier, or a trend) visible in this visualization	visibl	0							1	0.801035197
dataFeat2	I can clearly see data features (for example, a minimum, or an outlier, or a trend) in this visualization	see	0							2	0.745991576
dataRead1	I can easily find specific elements in this visualization	find	0					1	0.813425997		
	I can easily distinguish individual elements of the represented data (for example individual lines, or dots, or areas, or colors...)	distinguish	1				0.355062313	2	0.655425343		
dataRead2	I can easily identify relevant information in this visualization	identifi	0					3	0.603689212		
dataRead3	I can read data values from this visualization	valu	0					4	0.600135661		
dataRead4	I can easily retrieve information from this visualization	inform	0					5	0.599757226		
dataRead5	I find this visualization readable	readabl	0					6	0.505646165		
	I can easily read this visualization	read	1		0.333193092			7	0.484235378		
	This visualization effectively shows the data to me	effect	1		0.35791681			8	0.399574603		
	This visualization shows the data in a clear way for me	clearData	1				0.347569591	9	0.36121256		



◆ Understand subscale

understand1 It is **obvious** for me how to read this visualization

understand2 I can easily understand **how the data is represented** in this visualization

understand3 I can easily understand this visualization

◆ Layout subscale

layout1 I **don't** find this visualization **messy**

layout2 I **don't** find this visualization **crowded**

layout3 I **don't** find **distracting parts** in this visualization

◆ DataRead subscale

dataRead1 I can easily **find specific elements** in this visualization

dataRead2 I can easily **identify relevant information** in this visualization

dataRead3 I can easily **retrieve information** from this visualization

◆ DataFeat subscale

dataFeat1 I find data features (for example, a minimum, or an outlier, or a trend) **visible** in this visualization

dataFeat2 I can **clearly see** data features (for example, a minimum, or an outlier, or a trend) in this visualization

Ratings with PREVis scales on exploratory stimuli

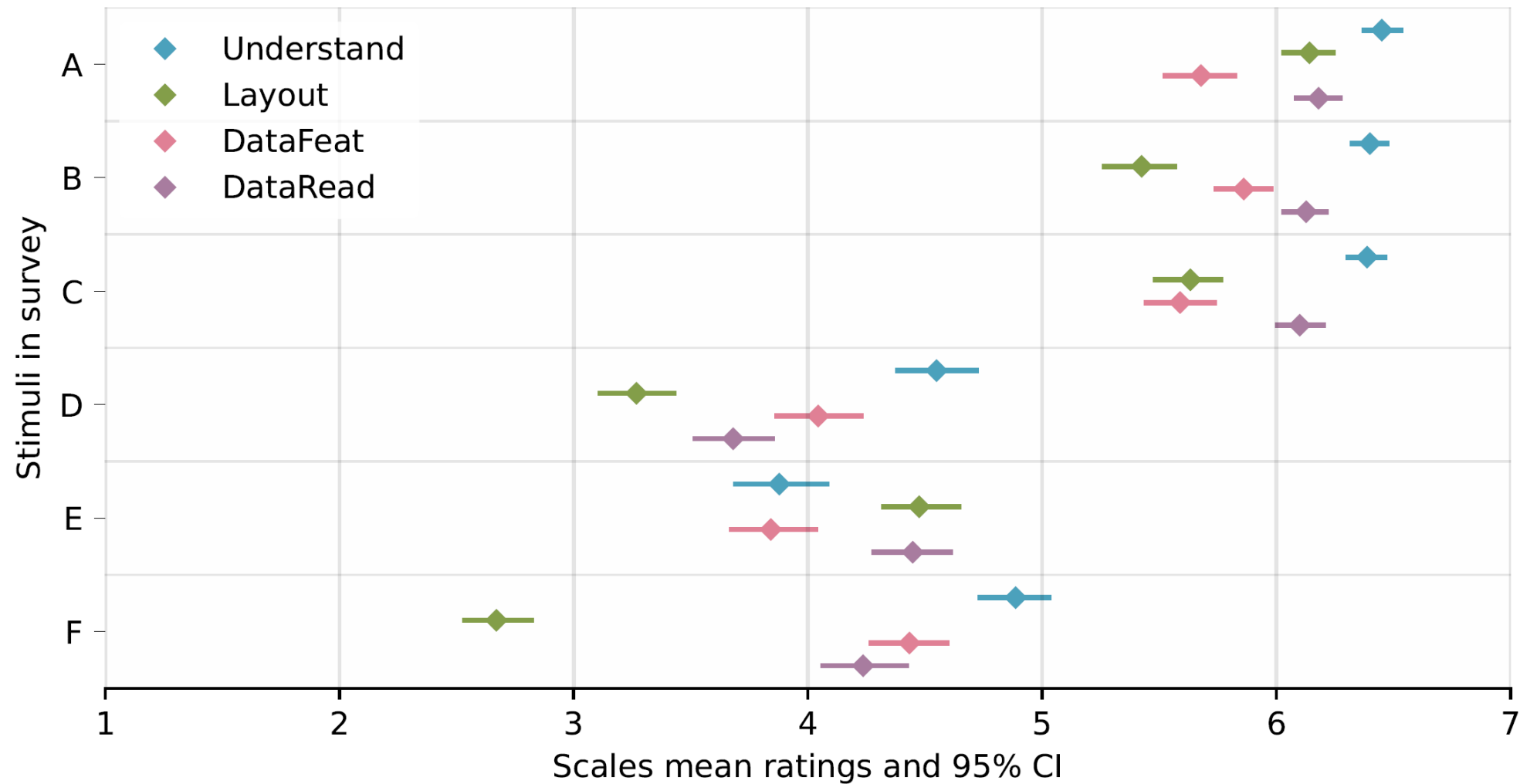
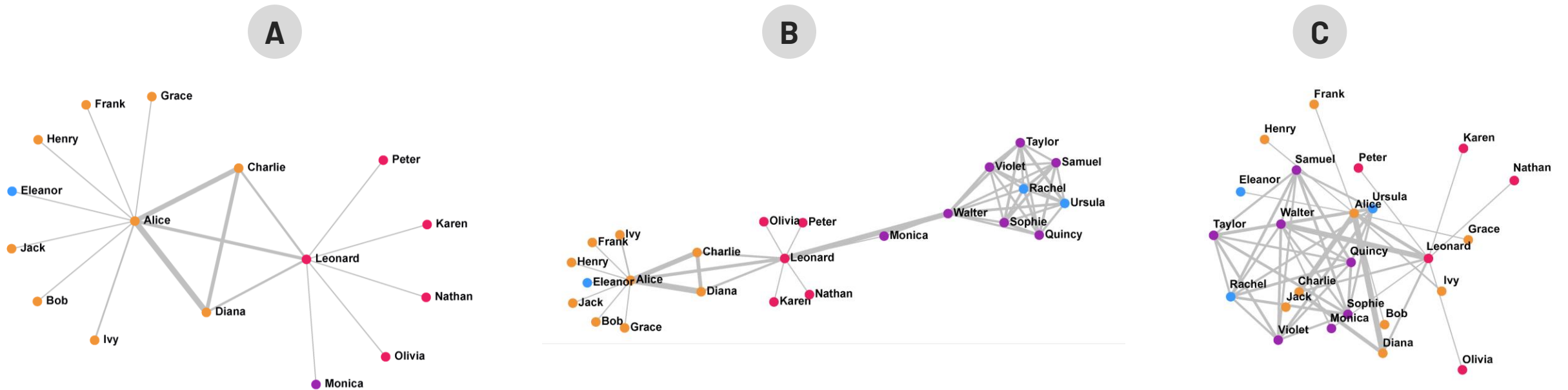


Fig. 29: Average ratings (from 1 = “Strongly disagree” to 7 = “Strongly agree”) using the four PREVis subscales \blacklozenge \blacklozenge \blacklozenge \blacklozenge on 6 visualizations of different readability: (A / B / C) > (D / E / F). For a given subscale, ratings and 95% CI do not overlap and fit the readability ranking.

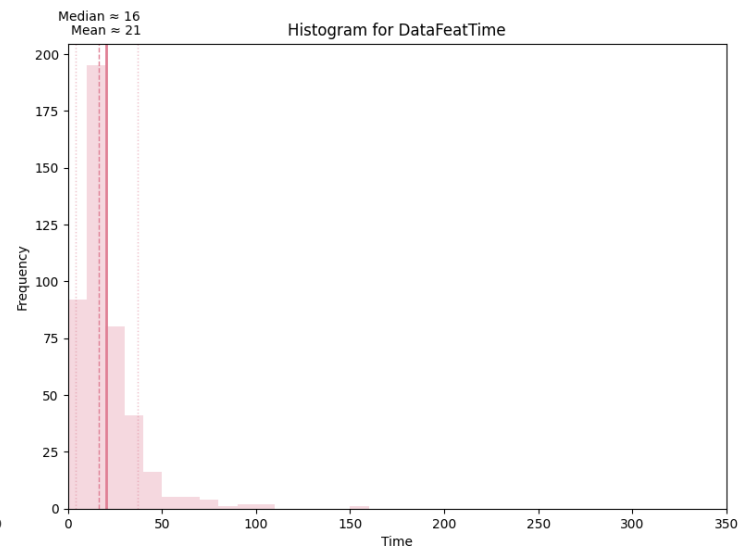
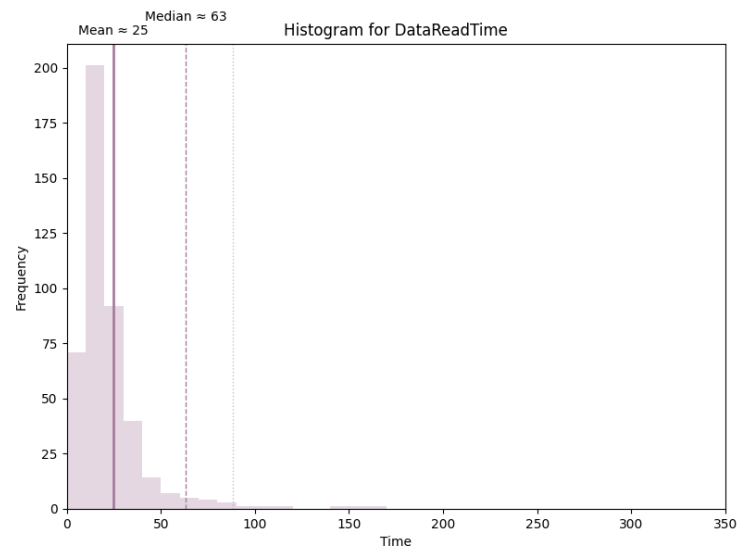
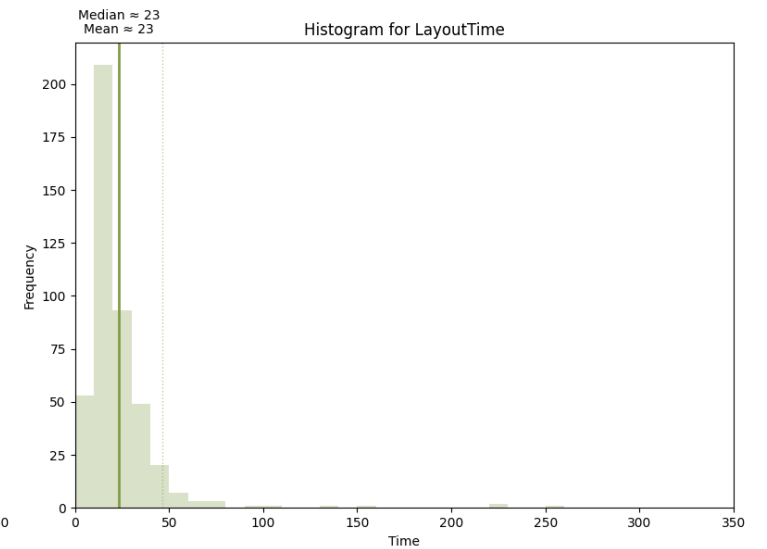
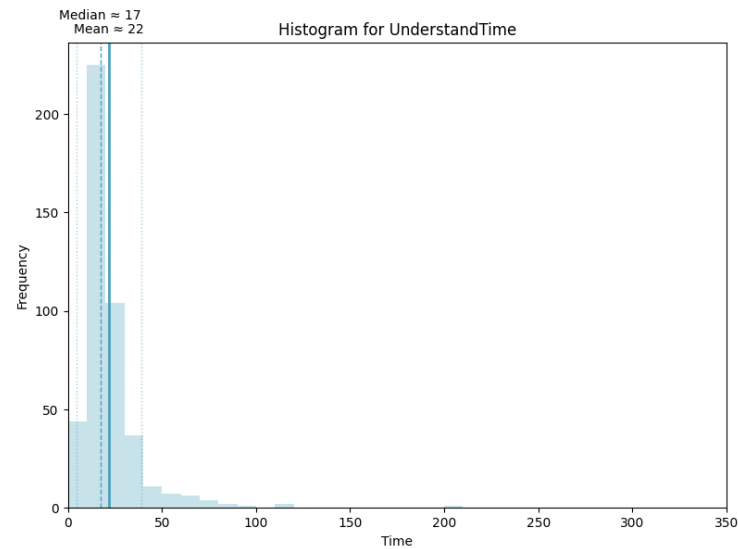
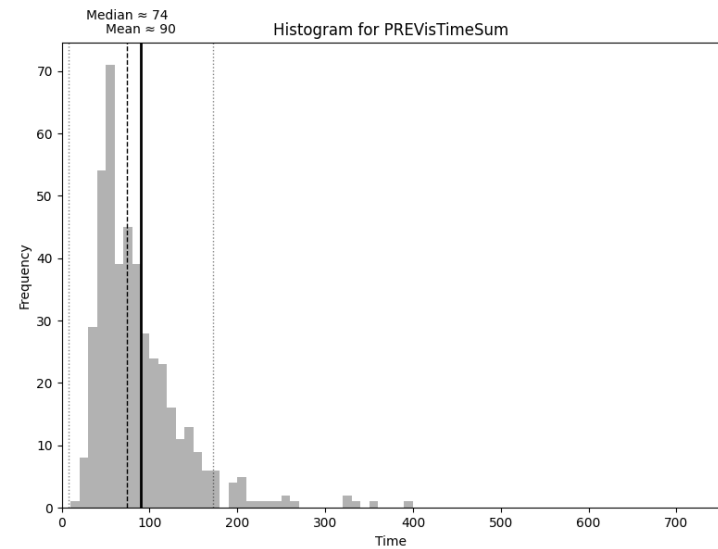
VALIDATION STUDY

Validation study



- ▶ N=147 on Prolific - All participants saw A, B, and C in random order
- ▶ People answered 2 reading questions then 11 PREVis questions
- ▶ Exclusion criteria: 1 comprehension check and 1 attention check

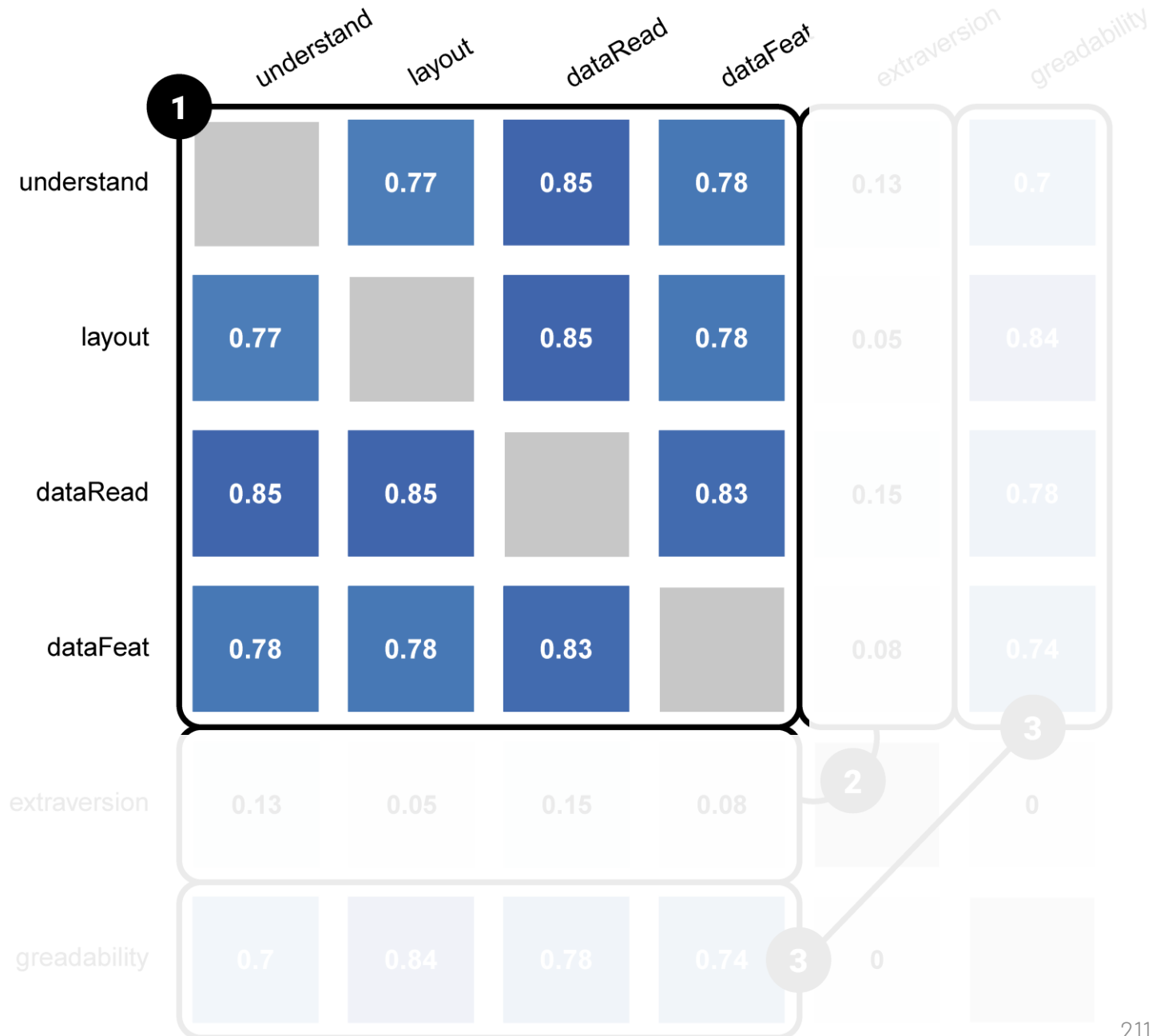
Time spent on PREVis ratings



Results: construct validity

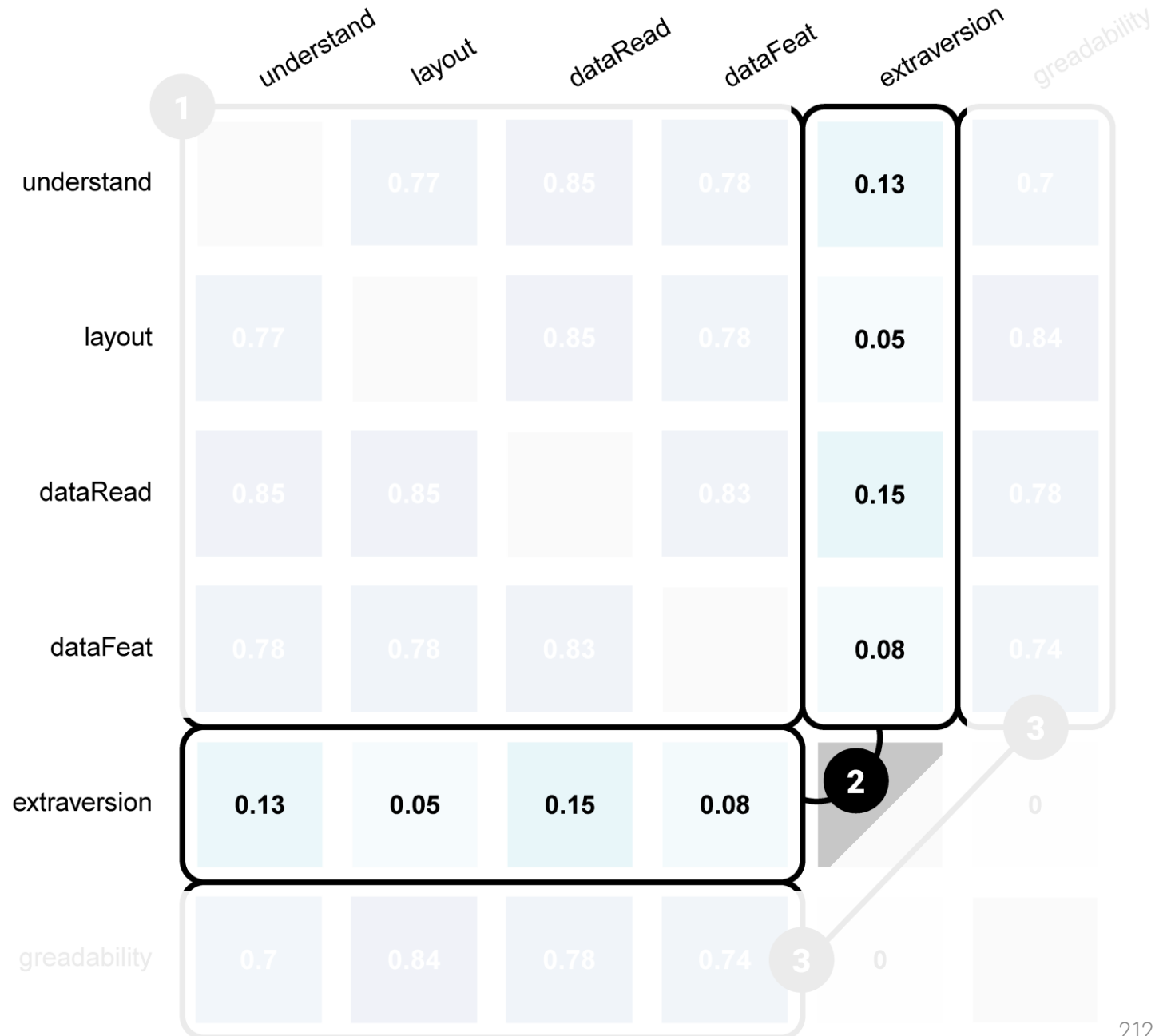
1. Inter-scales reliability:

each scale should highly correlate with the others, because they are all related to the same underlying construct: *readability*



Results: construct validity

1. Inter-scales reliability
2. **Discriminant validity:**
the results should not correlate with measures of an unrelated construct (extraversion scale from the Big Five Inventory)



Results:

construct validity

1. Inter-scales reliability
2. Discriminant validity
3. **Convergent validity:**
the results should positively correlate with independent measures that also relate to readability

Results: construct validity

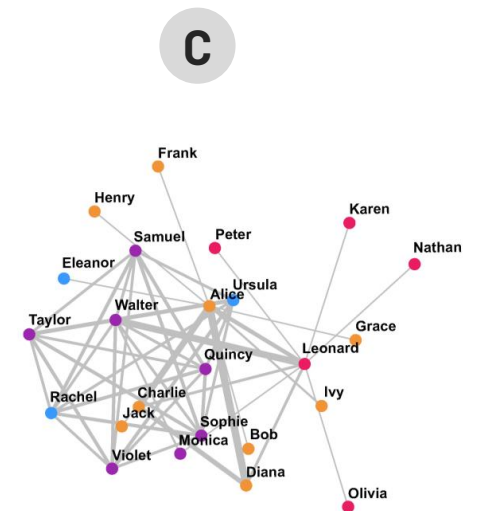
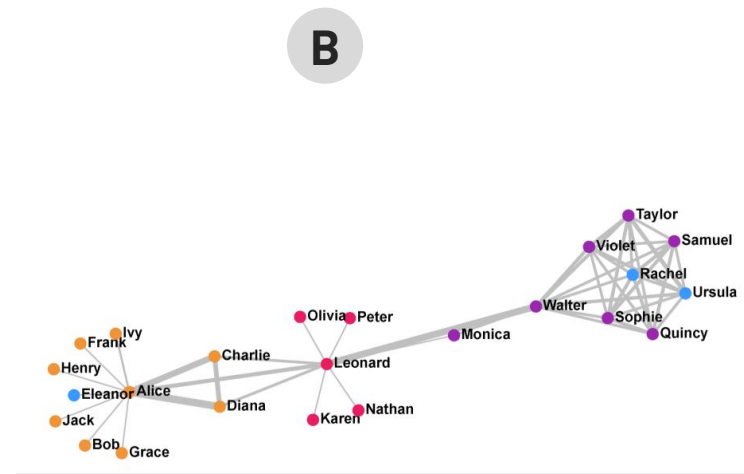
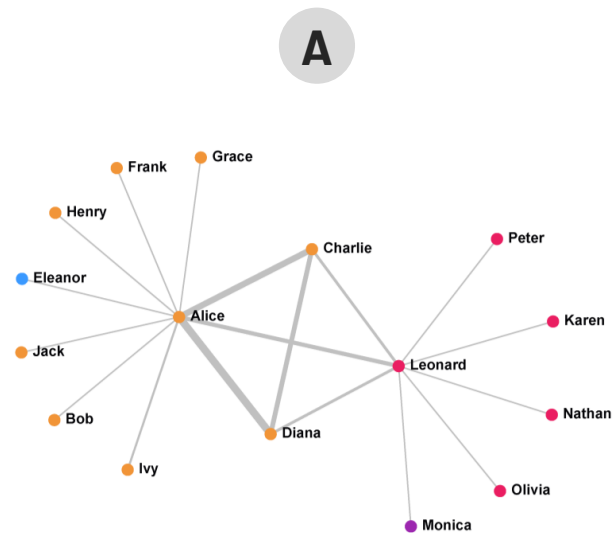
1. Inter-scales reliability
2. Discriminant validity
3. **Convergent validity**

Computed reference layout measures from greadability.js (Gove, 2018)

Stimulus A 0.872

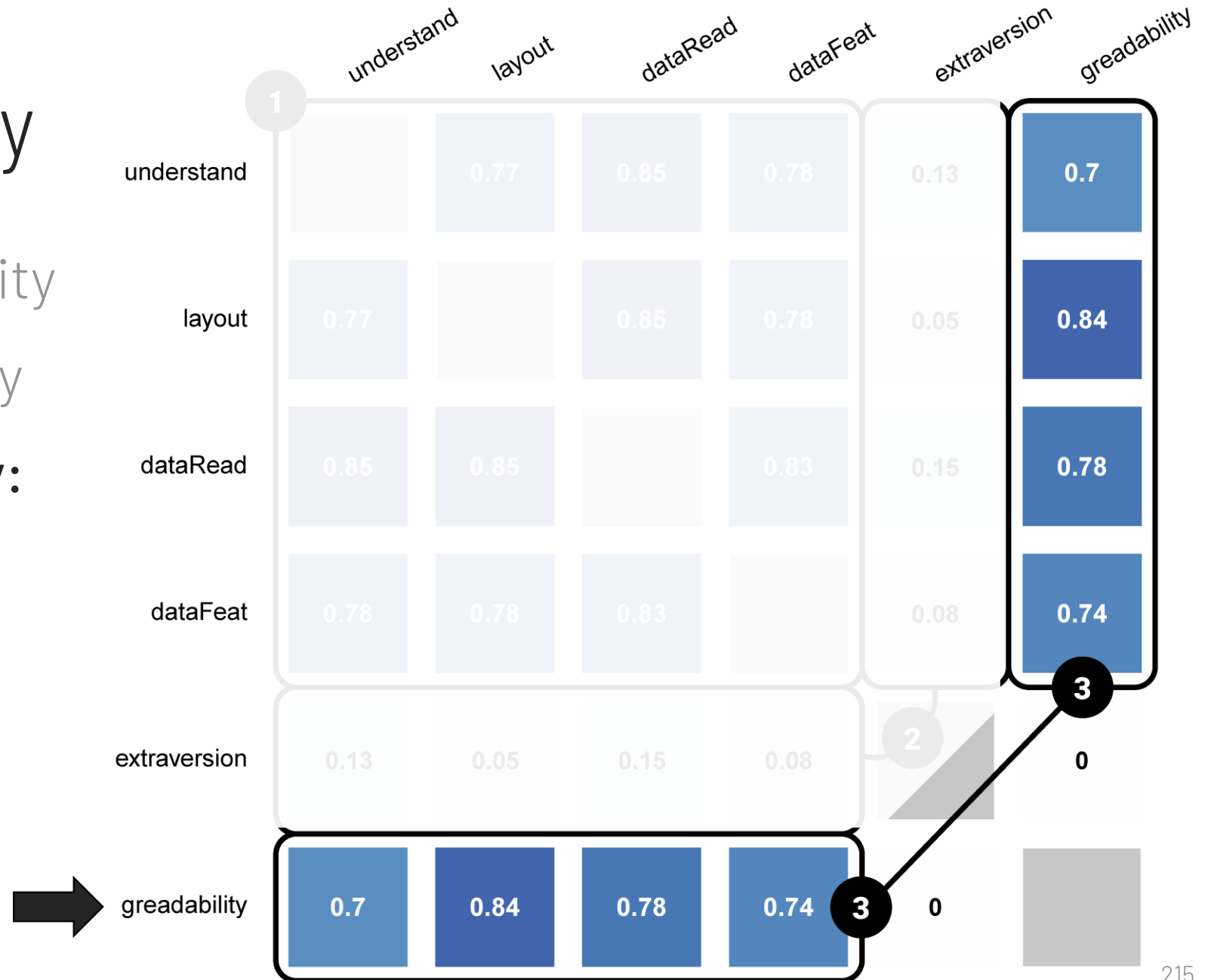
Stimulus B 0.786

Stimulus C 0.679



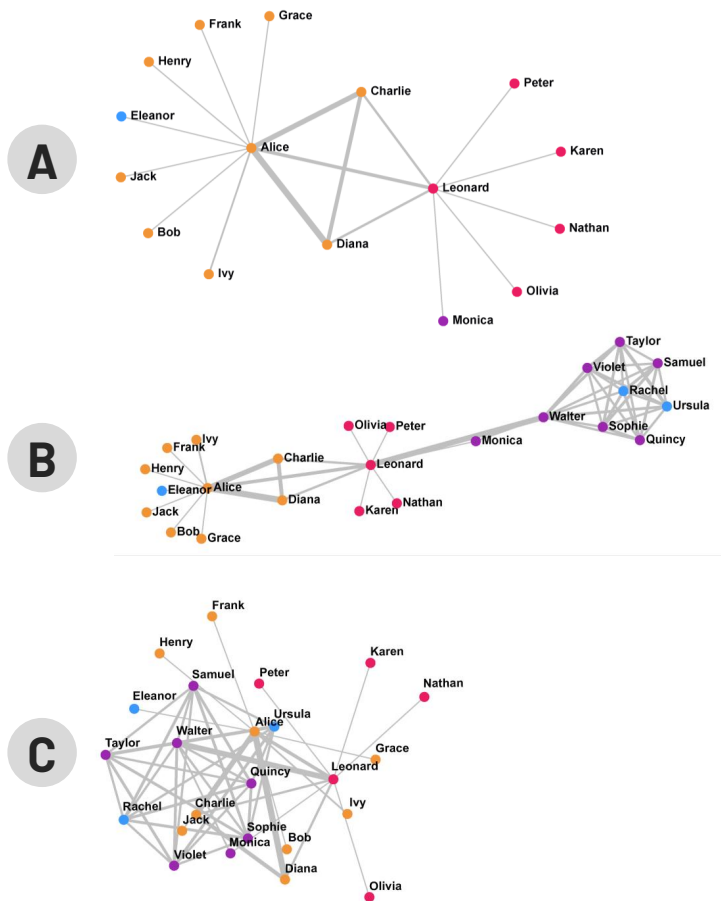
Results: construct validity

1. Inter-scales reliability
2. Discriminant validity
3. **Convergent validity:**
the results should positively correlate with independent measures that also relate to readability

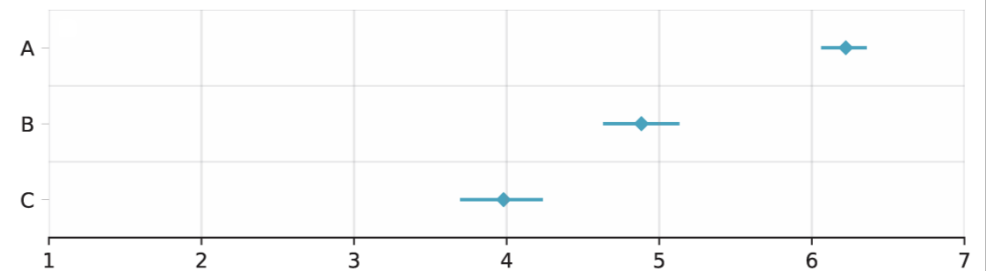


Results: ratings

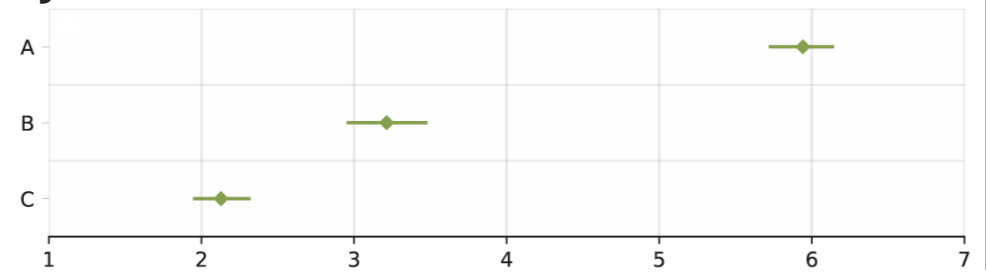
Point estimates with 95% CI



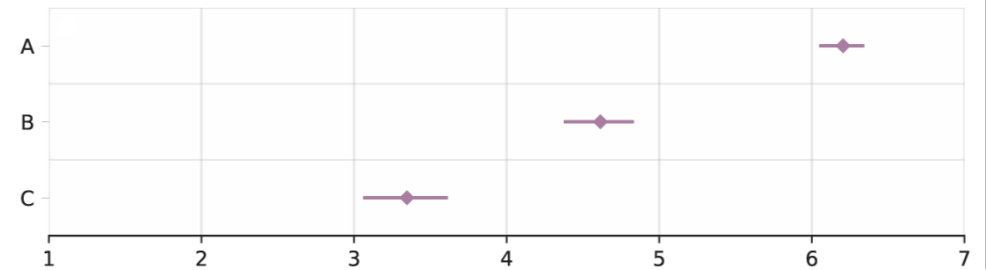
◆ Understand obvious + represent + understand



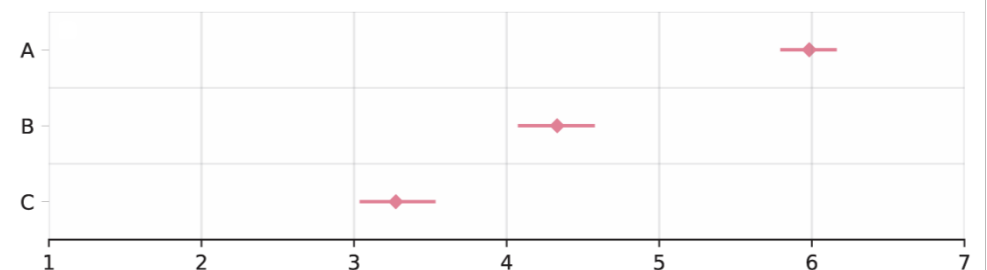
◆ Layout messy + crowded + distracting



◆ DataRead find + identify + inform



◆ DataFeat visible + clear



◆◆◆◆ PREVis: implementation guidelines

Why?

When?

How?

◆◆◆◆ PREVis: implementation guidelines

Why? To complement other measures



Task performance



Qualitative interviews

◆◆◆◆ PREVis: implementation guidelines

Why? To complement other measures

When? After people used the visualization(s)

◆◆◆◆ PREVis: implementation guidelines

Why? To complement other measures

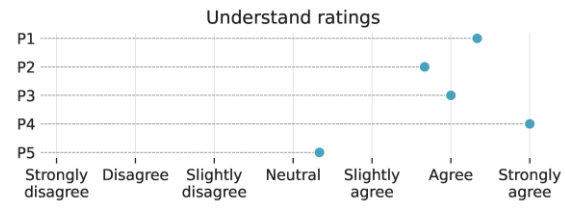
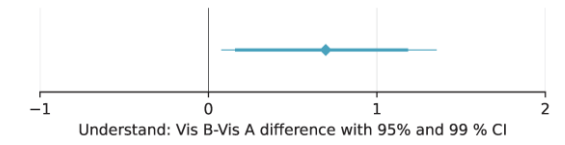
When? After people used the visualization(s)

How? Ensure reliability of your results: read the *** manual!

The screenshot displays the PREVis manual page, titled "PREVis Perceived Readability Evaluation in Visualization". It includes sections for "Understand subscale", "Layout subscale", "DataFeat subscale", and "DataRead subscale". Each section lists items with corresponding Likert scale anchors. For example, the "Understand subscale" includes items like "It is obvious for me how to read this visualization" and "I can easily understand how the data is represented in this visualization". The "Layout subscale" includes items like "I don't find this visualization messy" and "I don't find this visualization crowded". The "DataFeat subscale" includes items like "I find data features (for example, a minimum, or an outlier, or a trend) visible in this visualization" and "I can clearly see data features (for example, a minimum, or an outlier, or a trend) in this visualization". The "DataRead subscale" includes items like "I can easily find specific elements in this visualization" and "I can easily identify relevant information in this visualization". Below the items, there is a Likert scale ranging from "Strongly disagree" to "Strongly agree" with a "Neutral" option in the middle. A text box for "I don't know / Not applicable (if you wish, feel free to elaborate below)" is also present. The page footer indicates "An example of answer options for one item and how to present them".

- ▶ Question order and randomization
- ▶ Questions wording
- ▶ Answer options and labels
- ▶ Study cases and data analysis choices

Available materials



This document is provided on www.nature.com under the CC-BY license, as a supplemental material for our IEEE VIS 2024 paper on the PREVis scales.

PREVis

Perceived Readability Evaluation in Visualization

Questionnaire items and implementation recommendations

- Understand subscale**
 - understand** I can easily understand how the data is represented in this visualization
 - understand** I can easily understand this visualization
- Layout subscale**
 - messy** I don't find this visualization messy
 - clean** I don't find this visualization crowded
 - clearcut** I don't find distracting parts in this visualization
- DataFeat subscale**
 - visible** I find data features (for example, a minimum, or an outlier, or a trend) visible in this visualization
 - see** I can clearly see data features (for example, a minimum, or an outlier, or a trend) in this visualization
- DataRead subscale**
 - find** I can easily find specific elements in this visualization
 - identify** I can easily identify relevant information in this visualization
 - retrieve** I can easily retrieve information from this visualization

Regarding the visualization items, I can agree or disagree with the following statement

I can easily understand this visualization

Strongly disagree Disagree Slightly disagree Neutral Slightly agree Agree Strongly agree

Agree Not applicable if you wish, feel free to substitute below

An example of answer options for one item and how to present them:

PREVis: Perceived Readability Evaluation for Visualizations



Abstract—We developed and validated an instrument to measure the perceived readability in data visualization: PREVis. Researchers and practitioners can easily use this instrument as part of their evaluations to compare the perceived readability of different visual data representations. Our instrument can complement results from controlled experiments on user task performance or provide additional data during in-depth qualitative work such as design iterations when developing a new technique. Although readability is recognized as an essential quality of data visualizations, so far there has not been a unified definition of the construct in the context of visual representations. As a result, researchers often lack guidance for determining how to ask people to rate their perceived readability of a visualization. To address this issue, we engaged a rigorous process to develop the first validated instrument targeted at the subjective readability of visual data representations. Our final instrument consists of 11 items across 4 dimensions: understandability, layout clarity, readability of data patterns, and readability of data features. We provide the questionnaire as a document with implementation guidelines on our GitHub. Beyond the questionnaire, we contribute a discussion of how researchers have previously assessed visualization readability, and an analysis of the factors underlying perceived readability in visual data representations.

Index Terms—Readability, validated instrument, perception, user experiments, empirical methods, methodology.

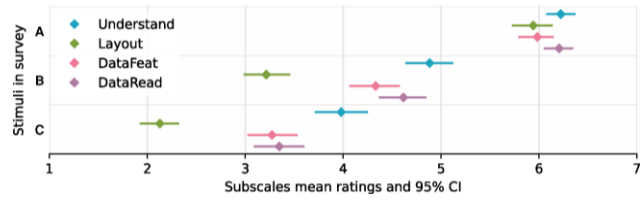
1 INTRODUCTION

When looking at examples of data visualizations, it is intuitive that some are easier to read than others. For many data analysis use cases, good readability will directly affect the amount of a visual representation of data for the viewer. As such, readability is a basic quality criterion in data visualization [1]. One of the fundamental challenges in studying the readability of data visualizations, however, is that the concept of reading and readability is held to different conceptions. The terms are often used in scientific writing without clear definitions of what they specifically mean in the context of data visualization—recalling Kieras’s “temple built on sand” [2].

Readability is best broadly defined as “the quality of being easy and enjoyable to read” [18]. It applies to letters and words as well as entire books. Lingens later developed hundreds of formulas to analyze the readability of text [4], but this approach fails to take into account characteristics of readers. Hence, readability is better explained as a function of the interaction between the properties of text and the characteristics of readers [4]. Moreover, new work is analyzing readability based on cognitive theories. Such an approach may also be suitable to explore the readability of visual representations of data.

As we discuss in more detail below, a few definitions of “readability” exist in the visualization domain, yet they do not fully overlap. Our PREVis tool addresses this problem because we followed well-established methodology in scale development [18, 24]. Developing a valid scale requires researchers to review existing literature and collect experts’ opinions on how to phrase the scale’s items. In this way it has an

	understand	layout	dataRead	dataFeat	extraversion	readability
understand	1	0.77	0.85	0.78	0.13	0.7
layout	0.77	1	0.85	0.78	0.05	0.84
dataRead	0.85	0.85	1	0.83	0.15	0.78
dataFeat	0.78	0.78	0.83	1	0.08	0.74
extraversion	0.13	0.05	0.15	0.08	2	0
readability	0.7	0.84	0.78	0.74	0	3



- ▶ Full paper with appendices
- ▶ Pre-registrations
- ▶ Research log

- ▶ Code and data to reproduce our work

- ▶ Implementation companion guide
- ▶ Code and example data to help you produce different figures

