

Supplementary Material 2

A SEARCH TERMS AND SCREENING

The table combines the term-level screening results used to construct the final corpus; the two search rounds are not split. At the later search, Vispubs contained 6,253 papers from 1986–2026. The 36 in-scope hits comprise the 22 + 10 distinct papers reported in the main text and four repeated hits across terms. Of the 32 distinct in-scope papers, 18 were already in the development corpus, leaving 14 validation papers.

Table 1: Combined term-level search and screening results. Counts are term–paper hits, not deduplicated corpus sizes.

Term group	Regular expression	Returned	In scope	In-scope share
Duplicate	\bduplicat\w*	17	0	0.0%
Re-evaluate	\bre-?evaluat\w*	2	0	0.0%
Rethink	\bre-?think\w*	7	1	14.3%
Reassess	\bre-?assess\w*	2	0	0.0%
Revisit	\brevisit\w*	35	5	14.3%
Replication	\breplica\w*	46	19	41.3%
Reproduction	\breproduc\w*	65	4	6.2%
Follow-up experiment or study	\bfollow(?:\s+ -)up\s+(?:experiment\w* stud\w*)	8	1	12.5%
Secondary analysis	\bsecondary\s+analys\w*	1	1	100.0%
Reanalysis	\bre-?analys\w*	4	0	0.0%
Previous or prior finding or result	\b(?:previous prior)\s+(?:finding\w* result\w*)	12	5	41.7%
Total hits		199	36	18.1%

\b marks a word boundary; \w* allows a word suffix; -? allows an optional hyphen; \s+ matches whitespace; and (?:a|b) groups alternatives.

B CORPUS REGISTRY

Table 2: Replication studies and their primary references.

Study	Replication paper	Primary reference
1	ExperiScope: an analysis tool for interaction data (2007)	The springboard: multiple modes in one spring-loaded control (2006)
2.1	Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design (2010)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
2.2	Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design (2010)	Alpha, contrast and the perception of visual metadata (2009)
2.3	Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design (2010)	Sizing the horizon: the effects of chart size and layering on the graphical perception of time series visualizations (2009)
3	Perceptual Guidelines for Creating Rectangular Treemaps (2010)	Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design (2010)
4	Do Mechanical Turks Dream of Square Pie Charts? (2010)	The Shaping of Information by Visual Metaphors (2008)
5	The Impact of Social Information on Visual Judgments (2011)	Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design (2010)
6	Does an Eye Tracker Tell the Truth about Visualizations?: Findings while Investigating Visualizations for Decision Making (2012)	A comparative study of three sorting techniques in performing cognitive tasks on a tabular representation (2012)
7.1	Assessing the Effect of Visualizations on Bayesian Reasoning through Crowdsourcing (2012)	Pictorial representations in statistical reasoning (2009)
7.2	Assessing the Effect of Visualizations on Bayesian Reasoning through Crowdsourcing (2012)	Pictorial representations in statistical reasoning (2009)
8	Priming Locus of Control to Affect Performance (2012)	How locus of control influences compatibility with visualization style (2011)
9.1	An Empirical Model of Slope Ratio Comparisons (2012)	Graphical perception: The visual decoding of quantitative information on graphical displays of data (1987)
9.2	An Empirical Model of Slope Ratio Comparisons (2012)	Graphical perception: The visual decoding of quantitative information on graphical displays of data (1987)
10	Influencing Visual Judgment through Affective Priming (2013)	Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design (2010)
11	Interactive Visualizations on Large and Small Displays: The Interrelation of Display Size, Information Space, and Scale (2013)	Sizing Up Visualizations: Effects of Display Size in Focus+Context, Overview+Detail, and Zooming Interfaces (2011)

Continued on next page

Table 2 (continued)

Study	Replication paper	Primary reference
12	How Visualization Layout Relates to Locus of Control and Other Personality Factors (2013)	Towards the Personal Equation of Interaction: The Impact of Personality Factors on Visual Analytics Interface Interaction (2010)
13.1	Ranking Visualizations of Correlation Using Weber's Law (2014)	The Perception of Correlation in Scatterplots (2010)
13.2	Ranking Visualizations of Correlation Using Weber's Law (2014)	The Perception of Correlation in Scatterplots (2010)
14.1	Four Experiments on the Perception of Bar Charts (2014)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
14.2	Four Experiments on the Perception of Bar Charts (2014)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
14.3	Four Experiments on the Perception of Bar Charts (2014)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
14.4	Four Experiments on the Perception of Bar Charts (2014)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
15.1	Elementary graphical perception for older adults: A comparison with the general population (2014)	An Information-Processing Analysis of Graph Perception (1987)
15.2	Elementary graphical perception for older adults: A comparison with the general population (2014)	An Information-Processing Analysis of Graph Perception (1987)
16.1	GraphUnit: Evaluating Interactive Graph Visualizations Using Crowdsourcing (2015)	A Comparison of the Readability of Graphs Using Node-Link and Matrix-Based Representations (2004)
16.2	GraphUnit: Evaluating Interactive Graph Visualizations Using Crowdsourcing (2015)	A User Study on Visualizing Directed Edges in Graphs (2009)
17	Beyond Weber's Law: A Second Look at Ranking Visualizations of Correlation (2015)	Ranking Visualizations of Correlation Using Weber's Law (2014)
18.1	Improving Bayesian Reasoning: The Effects of Phrasing, Visualization, and Spatial Ability (2015)	Assessing the Effect of Visualizations on Bayesian Reasoning through Crowdsourcing (2012)
18.2	Improving Bayesian Reasoning: The Effects of Phrasing, Visualization, and Spatial Ability (2015)	Assessing the Effect of Visualizations on Bayesian Reasoning through Crowdsourcing (2012)
19	A Psychophysical Investigation of Size as a Physical Variable (2015)	An experiment in graphical perception (1986)
20	Judgment Error in Pie Chart Variations (2016)	Arcs, Angles, or Areas: Individual Data Encodings in Pie and Donut Charts (2016)
21.1	Arcs, Angles, or Areas: Individual Data Encodings in Pie and Donut Charts (2016)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
21.2	Arcs, Angles, or Areas: Individual Data Encodings in Pie and Donut Charts (2016)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
22.1	The Attraction Effect in Information Visualization (2016)	Between a rock and a hard place: The failure of the attraction effect among unattractive alternatives (2013)
22.2	The Attraction Effect in Information Visualization (2016)	Distinguishing among models of contextually induced preference reversals (1991)
23	HindSight: Encouraging Exploration through Direct Encoding of Personal Interaction History (2016)	Storytelling in information visualizations: Does it engage users to explore data? (2015)
24.1	Blinded with Science or Informed by Charts? A Replication Study (2017)	Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy (2016)
24.2	Blinded with Science or Informed by Charts? A Replication Study (2017)	Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy (2016)
24.3	Blinded with Science or Informed by Charts? A Replication Study (2017)	Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy (2016)
24.4	Blinded with Science or Informed by Charts? A Replication Study (2017)	Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy (2016)
25.1	Modeling Color Difference for Visualization Design (2017)	An engineering model for color difference as a function of size (2014)
25.2	Modeling Color Difference for Visualization Design (2017)	An engineering model for color difference as a function of size (2014)
25.3	Modeling Color Difference for Visualization Design (2017)	An engineering model for color difference as a function of size (2014)
26.1	Open vs. Closed Shapes: New Perceptual Categories? (2017)	High perceptual load makes everybody equal: Eliminating individual differences in distractibility with load (2007)
26.2	Open vs. Closed Shapes: New Perceptual Categories? (2017)	High perceptual load makes everybody equal: Eliminating individual differences in distractibility with load (2007)
27.1	Graphical Perception of Continuous Quantitative Maps: the Effects of Spatial Frequency and Colormap Design (2018)	Color sequences for univariate maps: Theory, experiments and principles (1988)
28.1	Evaluating 'Graphical Perception' with CNNs (2018)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)

Continued on next page

Table 2 (continued)

Study	Replication paper	Primary reference
28.2	Evaluating 'Graphical Perception' with CNNs (2018)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
28.3	Evaluating 'Graphical Perception' with CNNs (2018)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
28.4	Evaluating 'Graphical Perception' with CNNs (2018)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
29	VizNet: Towards A Large-Scale Visualization Learning and Benchmarking Repository (2019)	Assessing Effects of Task and Data Distribution on the Effectiveness of Visual Encodings (2018)
30	Correlation Judgment and Visualization Features: A Comparative Study (2019)	Ranking Visualizations of Correlation Using Weber's Law (2014)
31	Revisited: Comparison of Empirical Methods to Evaluate Visualizations Supporting Crafting and Assembly Purposes (2020)	Online, VR, AR, Lab, and In-Situ: Comparison of Research Methods to Evaluate Smart Artifacts (2019)
32	Pushing the (Visual) Narrative: the Effects of Prior Knowledge Elicitation in Provocative Topics (2020)	The persuasive power of data visualization (2014)
33	Understanding Data Accessibility for People with Intellectual and Developmental Disabilities (2021)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
34	Effect of uncertainty visualizations on myopic loss aversion and the equity premium puzzle in retirement investment decisions (2021)	Risk aversion or myopia? Choices in repeated gambles and retirement investments (1999)
35	Rethinking the Ranks of Visual Channels (2021)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
36.1	Evaluating the Effects of Visual Embellishments and Context on Comprehension, Recall and Preference of Charts (2022)	Useful junk? The effects of visual embellishment on comprehension and memorability of charts (2010)
36.2	Evaluating the Effects of Visual Embellishments and Context on Comprehension, Recall and Preference of Charts (2022)	Useful junk? The effects of visual embellishment on comprehension and memorability of charts (2010)
37	Juvenile Graphical Perception: A Comparison between Children and Adults (2022)	Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design (2010)
38	VoxLens: Making Online Data Visualizations Accessible with an Interactive JavaScript Plug-In (2022)	Understanding Screen-Reader Users' Experiences with Online Data Visualizations (Study 2) (2021)
39.1	More Useful Junk? Replicating the Effects of Visual Embellishment on Description and Recall of Charts (2023)	Useful junk? The effects of visual embellishment on comprehension and memorability of charts (2010)
39.2	More Useful Junk? Replicating the Effects of Visual Embellishment on Description and Recall of Charts (2023)	Useful junk? The effects of visual embellishment on comprehension and memorability of charts (2010)
39.3	More Useful Junk? Replicating the Effects of Visual Embellishment on Description and Recall of Charts (2023)	Useful junk? The effects of visual embellishment on comprehension and memorability of charts (2010)
39.4	More Useful Junk? Replicating the Effects of Visual Embellishment on Description and Recall of Charts (2023)	Useful junk? The effects of visual embellishment on comprehension and memorability of charts (2010)
40	The Risks of Ranking: Revisiting Graphical Perception to Model Individual Differences in Visualization Performance (2024)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)
41	Glanceable Data Visualizations for Older Adults: Establishing Thresholds and Examining Disparities Between Age Groups (2024)	Glanceable Visualization: Studies of Data Comparison Performance on Smartwatches (2018)
42	Transparent Risks: The Impact of the Specificity and Visual Encoding of Uncertainty on Decision Making (2024)	Evaluating the impact of visualization of wildfire hazard upon decision-making under uncertainty (2016)
43.1	Doorways Do Not Always Cause Forgetting: Studying the Effect of Locomotion Technique and Doorway Visualization in Virtual Reality (2024)	Walking through doorways causes forgetting: Situation models and experienced space (2006)
43.2	Doorways Do Not Always Cause Forgetting: Studying the Effect of Locomotion Technique and Doorway Visualization in Virtual Reality (2024)	Walking through doorways causes forgetting: Situation models and experienced space (2006)
44	Validation through Replication of Augmented Reality as a Visualization Technique for Scholarly Publications in Astronomy (2025)	Augmented Reality as a Visualization Technique for Scholarly Publications in Astronomy: An Empirical Evaluation (2023)
45.1	Seeing Eye to AI? Applying Deep-Feature-Based Similarity Metrics to Information Visualization (2025)	Discriminability Tests for Visualization Effectiveness and Scalability (2019)
45.2	Seeing Eye to AI? Applying Deep-Feature-Based Similarity Metrics to Information Visualization (2025)	Learning Perceptual Kernels for Visualization Design (2014)
46.1	ReVISit 2: A Full Experiment Life Cycle User Study Framework (2025)	Ranking Visualizations of Correlation Using Weber's Law (2014)
46.2	ReVISit 2: A Full Experiment Life Cycle User Study Framework (2025)	Design Characterization for Black-and-White Textures in Visualization (2023)
46.3	ReVISit 2: A Full Experiment Life Cycle User Study Framework (2025)	The Effects of Adding Search Functionality to Interactive Visualizations on the Web (2018)
47	They Aren't Built For Me: An Exploratory Study of Strategies for Measurement of Graphical Primitives in Tactile Graphics (2025)	Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984)

Continued on next page

Table 2 (continued)

Study	Replication paper	Primary reference
48.1	Characterizing Visualization Perception with Psychological Phenomena: Uncovering the Role of Subitizing in Data Visualization (2025)	Measuring categorical perception in color-coded scatterplots (2023)
48.2	Characterizing Visualization Perception with Psychological Phenomena: Uncovering the Role of Subitizing in Data Visualization (2025)	Measuring categorical perception in color-coded scatterplots (2023)
48.3	Characterizing Visualization Perception with Psychological Phenomena: Uncovering the Role of Subitizing in Data Visualization (2025)	Improving the robustness of scagnostics (2019)
49.1	Do Graph Drawing Aesthetics Matter for AI? A Replication of Foundational Studies in Graph Readability (2026)	Which aesthetic has the greatest effect on human understanding? (1997)
49.2	Do Graph Drawing Aesthetics Matter for AI? A Replication of Foundational Studies in Graph Readability (2026)	Performance of layout algorithms: Comprehension, not computation (1998)
50.1	Can LLMs Simulate Target Users in Visualization Case Studies? (2026)	Parallel Edge Splatting for Scalable Dynamic Graph Visualization (2011)
50.2	Can LLMs Simulate Target Users in Visualization Case Studies? (2026)	Visual Analysis of Document Editing Patterns (2025)
51.1	Taking a Walk on the Wild Side: Effects of Walking in Synchrony with Pitch-Altered Footstep Sounds on Body Perception in Outside the Lab Contexts (2026)	As light as your footsteps: Altering walking sounds to change perceived body weight, emotional state and gait (2015)
51.2	Taking a Walk on the Wild Side: Effects of Walking in Synchrony with Pitch-Altered Footstep Sounds on Body Perception in Outside the Lab Contexts (2026)	As light as your footsteps: Altering walking sounds to change perceived body weight, emotional state and gait (2015)

C CODING RESULTS

The 86 replication studies are ordered by paper year and numbered accordingly. Decimal suffixes distinguish studies within the same paper. In the open-code columns, — denotes no assigned unified phrase. *N/A* denotes an inactive comparison dimension. HM denotes an honourable mention. BP denotes a best paper award.

Table 3: Replication-study coding results.

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
1	ExperiScope: an analysis tool for interaction data 2007 Award: — Primary reference: The springboard: multiple modes in one spring-loaded control (2006).	Research team: Overlap Explicit claim: No Result: Consistent Main motivation: Evaluate a research method or tool	STIMULI: <i>N/A</i> TASK: <i>N/A</i> PROCEDURE: <i>N/A</i> INTERFACE: <i>N/A</i> ENVIRONMENT: <i>N/A</i> DATA: <i>identical</i> PARTICIPANT: <i>N/A</i> ANALYSIS: <i>different</i>	EXPERIMENT: — DATA: Reuse existing evidence PARTICIPANT: — ANALYSIS: Change qualitative or visual analysis	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
2.1	Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design 2010 Award: HM (CHI 2010) Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Evaluate a research method or tool	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>identical</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: —	EXPERIMENT: Add a condition or benchmark comparison DATA: Collect additional performance evidence; Record additional operational or contextual metadata PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons; Assess operational feasibility or study process
2.2	Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design 2010 Award: HM (CHI 2010) Primary reference: Alpha, contrast and the perception of visual metadata (2009).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Evaluate a research method or tool	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: Change the scope of comparisons	EXPERIMENT: — DATA: Collect additional performance evidence; Record additional operational or contextual metadata PARTICIPANT: — ANALYSIS: Analyse background or contextual associations
2.3	Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design 2010 Award: HM (CHI 2010) Primary reference: Sizing the horizon: the effects of chart size and layering on the graphical perception of time series visualizations (2009).	Research team: Overlap Explicit claim: Yes Result: Partly consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Expand stimulus or condition coverage; Change stimulus representation or modality; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Change the scope of comparisons	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
3	Perceptual Guidelines for Creating Rectangular Treemaps 2010 Award: HM (InfoVis 2010) Primary reference: Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design (2010).	Research team: Overlap Explicit claim: Yes Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>identical</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>identical</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Expand stimulus or condition coverage; Restrict stimulus or condition coverage; Adjust the study administration protocol DATA: Collect new observations PARTICIPANT: — ANALYSIS: Change statistical tests or models; Change the scope of comparisons	EXPERIMENT: Add a condition or benchmark comparison; Add a formative or pilot study DATA: Collect additional performance evidence PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons
4	Do Mechanical Turks Dream of Square Pie Charts? 2010 Award: — Primary reference: The Shaping of Information by Visual Metaphors (2008).	Research team: Overlap Explicit claim: Yes Result: Partly consistent Main motivation: Evaluate a research method or tool	STIMULI: <i>identical</i> TASK: <i>identical</i> PROCEDURE: <i>identical</i> INTERFACE: <i>identical</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>identical</i>	EXPERIMENT: Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: —	EXPERIMENT: Add a new task or question DATA: Collect additional performance evidence; Collect subjective ratings or preferences; Collect participant-background or ability measures; Record additional operational or contextual metadata PARTICIPANT: Recruit a separate extension sample ANALYSIS: Analyse additional outcomes or comparisons; Analyse background or contextual associations; Assess operational feasibility or study process
5	The Impact of Social Information on Visual Judgments 2011 Award: — Primary reference: Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design (2010).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Evaluate an intervention using a replicated baseline	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>identical</i> INTERFACE: <i>identical</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>identical</i> ANALYSIS: <i>identical</i>	EXPERIMENT: Restrict stimulus or condition coverage DATA: Collect new observations PARTICIPANT: — ANALYSIS: —	EXPERIMENT: Add a new task or question; Add a condition or benchmark comparison DATA: Collect additional performance evidence PARTICIPANT: Recruit a separate extension sample ANALYSIS: Analyse additional outcomes or comparisons
6	Does an Eye Tracker Tell the Truth about Visualizations?: Findings while Investigating Visualizations for Decision Making 2012 Award: — Primary reference: A comparative study of three sorting techniques in performing cognitive tasks on a tabular representation (2012).	Research team: Overlap Explicit claim: Yes Result: Consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>identical</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>different</i>	EXPERIMENT: Adjust the study administration protocol DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation	EXPERIMENT: Add a condition or benchmark comparison DATA: Collect additional performance evidence PARTICIPANT: Recruit a separate extension sample ANALYSIS: Analyse additional outcomes or comparisons

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
7.1	Assessing the Effect of Visualizations on Bayesian Reasoning through Crowdsourcing 2012 Award: HM (InfoVis 2012) Primary reference: Pictorial representations in statistical reasoning (2009).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Examine generalisability across human populations	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Expand stimulus or condition coverage; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation	EXPERIMENT: Add a condition or benchmark comparison; Add a follow-up assessment or questionnaire DATA: Collect additional performance evidence; Collect subjective ratings or preferences; Collect qualitative accounts or recordings; Collect participant-background or ability measures; Record additional operational or contextual metadata PARTICIPANT: Recontact or select existing participants ANALYSIS: Analyse additional outcomes or comparisons; Analyse background or contextual associations; Analyse additional qualitative accounts
7.2	Assessing the Effect of Visualizations on Bayesian Reasoning through Crowdsourcing 2012 Award: HM (InfoVis 2012) Primary reference: Pictorial representations in statistical reasoning (2009).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Reassess the robustness of a reported finding	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Restrict stimulus or condition coverage; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change outcome scoring or aggregation; Revise uncertainty or robustness assessment	EXPERIMENT: Add a condition or benchmark comparison DATA: Collect additional performance evidence; Collect subjective ratings or preferences; Collect qualitative accounts or recordings PARTICIPANT: Recruit a separate extension sample ANALYSIS: Analyse additional outcomes or comparisons; Analyse background or contextual associations
8	Priming Locus of Control to Affect Performance 2012 Award: — Primary reference: How locus of control influences compatibility with visualization style (2011).	Research team: Overlap Explicit claim: Yes Result: Consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>identical</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Restrict stimulus or condition coverage; Adjust the study administration protocol; Change the experimental structure or paradigm DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Analyse individual differences or moderators	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
9.1	An Empirical Model of Slope Ratio Comparisons 2012 Award: — Primary reference: Graphical perception: The visual decoding of quantitative information on graphical displays of data (1987).	Research team: Independent Explicit claim: No Result: Partly consistent Main motivation: Develop or refine a model of performance	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>similar</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Expand stimulus or condition coverage; Adjust the study administration protocol; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators; Revise uncertainty or robustness assessment	EXPERIMENT: Add a formative or pilot study DATA: Collect additional performance evidence PARTICIPANT: Recruit a separate extension sample ANALYSIS: —
9.2	An Empirical Model of Slope Ratio Comparisons 2012 Award: — Primary reference: Graphical perception: The visual decoding of quantitative information on graphical displays of data (1987).	Research team: Independent Explicit claim: No Result: Partly consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>similar</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Expand stimulus or condition coverage; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
10	Influencing Visual Judgment through Affective Priming 2013 Award: — Primary reference: Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design (2010).	Research team: Independent Explicit claim: No Result: Consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Restrict stimulus or condition coverage; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Analyse individual differences or moderators; Revise uncertainty or robustness assessment	EXPERIMENT: Add a formative or pilot study DATA: Collect subjective ratings or preferences; Collect qualitative accounts or recordings PARTICIPANT: Recruit a separate extension sample ANALYSIS: Analyse additional outcomes or comparisons

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
11	Interactive Visualizations on Large and Small Displays: The Interrelation of Display Size, Information Space, and Scale 2013 Award: — Primary reference: Sizing Up Visualizations: Effects of Display Size in Focus+Context, Overview+Detail, and Zooming Interfaces (2011).	Research team: Overlap Explicit claim: Yes Result: Partly consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>identical</i> INTERFACE: <i>identical</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Adjust the study administration protocol DATA: Collect new observations PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Change the scope of comparisons	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
12	How Visualization Layout Relates to Locus of Control and Other Personality Factors 2013 Award: — Primary reference: Towards the Personal Equation of Interaction: The Impact of Personality Factors on Visual Analytics Interface Interaction (2010).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Restrict stimulus or condition coverage; Change stimulus representation or modality; Change the task or response requirement; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Analyse individual differences or moderators; Change the scope of comparisons	EXPERIMENT: — DATA: Collect subjective ratings or preferences PARTICIPANT: — ANALYSIS: Analyse background or contextual associations
13.1	Ranking Visualizations of Correlation Using Weber's Law 2014 Award: — Primary reference: The Perception of Correlation in Scatterplots (2010).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Evaluate a research method or tool	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Restrict stimulus or condition coverage; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population ANALYSIS: Change outcome scoring or aggregation; Change the scope of comparisons	EXPERIMENT: — DATA: Collect qualitative accounts or recordings; Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: —

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
13.2	Ranking Visualizations of Correlation Using Weber's Law 2014 Award: — Primary reference: The Perception of Correlation in Scatterplots (2010).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Expand stimulus or condition coverage; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: Collect qualitative accounts or recordings; Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: —
14.1	Four Experiments on the Perception of Bar Charts 2014 Award: — Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: No Result: Consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Expand stimulus or condition coverage; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change outcome scoring or aggregation; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: Collect additional performance evidence PARTICIPANT: — ANALYSIS: —
14.2	Four Experiments on the Perception of Bar Charts 2014 Award: — Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: No Result: Consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Expand stimulus or condition coverage; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change outcome scoring or aggregation; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: Collect additional performance evidence PARTICIPANT: — ANALYSIS: —

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
14.3	Four Experiments on the Perception of Bar Charts 2014 Award: — Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: No Result: Consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Expand stimulus or condition coverage; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change outcome scoring or aggregation; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: Collect additional performance evidence PARTICIPANT: — ANALYSIS: —
14.4	Four Experiments on the Perception of Bar Charts 2014 Award: — Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: No Result: Not consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Expand stimulus or condition coverage; Restrict stimulus or condition coverage; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change outcome scoring or aggregation; Analyse individual differences or moderators; Change the scope of comparisons	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
15.1	Elementary graphical perception for older adults: A comparison with the general population 2014 Award: — Primary reference: An Information-Processing Analysis of Graph Perception (1987).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Examine generalisability across human populations	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Analyse individual differences or moderators	EXPERIMENT: — DATA: Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: —

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
15.2	Elementary graphical perception for older adults: A comparison with the general population 2014 Award: — Primary reference: An Information-Processing Analysis of Graph Perception (1987).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Examine generalisability across human populations	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Analyse individual differences or moderators	EXPERIMENT: — DATA: Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: —
16.1	GraphUnit: Evaluating Interactive Graph Visualizations Using Crowdsourcing 2015 Award: — Primary reference: A Comparison of the Readability of Graphs Using Node-Link and Matrix-Based Representations (2004).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Evaluate a research method or tool	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Restrict stimulus or condition coverage; Change the task or response requirement; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment	EXPERIMENT: Add a method or model evaluation DATA: Record additional operational or contextual metadata PARTICIPANT: Recruit a separate extension sample ANALYSIS: Assess operational feasibility or study process
16.2	GraphUnit: Evaluating Interactive Graph Visualizations Using Crowdsourcing 2015 Award: — Primary reference: A User Study on Visualizing Directed Edges in Graphs (2009).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Evaluate a research method or tool	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Restrict stimulus or condition coverage; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: Add a method or model evaluation DATA: Record additional operational or contextual metadata PARTICIPANT: Recruit a separate extension sample ANALYSIS: Assess operational feasibility or study process

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
17	Beyond Weber's Law: A Second Look at Ranking Visualizations of Correlation 2015 Award: HM (InfoVis 2015) Primary reference: Ranking Visualizations of Correlation Using Weber's Law (2014).	Research team: Independent Explicit claim: No Result: Partly consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>N/A</i> TASK: <i>N/A</i> PROCEDURE: <i>N/A</i> INTERFACE: <i>N/A</i> ENVIRONMENT: <i>N/A</i> DATA: <i>similar</i> PARTICIPANT: <i>N/A</i> ANALYSIS: <i>different</i>	EXPERIMENT: — DATA: Reuse existing evidence PARTICIPANT: — ANALYSIS: Change statistical tests or models; Analyse individual differences or moderators; Revise uncertainty or robustness assessment	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
18.1	Improving Bayesian Reasoning: The Effects of Phrasing, Visualization, and Spatial Ability 2015 Award: — Primary reference: Assessing the Effect of Visualizations on Bayesian Reasoning through Crowdsourcing (2012).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>identical</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Expand stimulus or condition coverage; Change the task or response requirement; Adjust the study administration protocol; Change the study interface DATA: Collect new observations PARTICIPANT: — ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Change the scope of comparisons	EXPERIMENT: — DATA: Record additional process or interaction evidence; Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: —
18.2	Improving Bayesian Reasoning: The Effects of Phrasing, Visualization, and Spatial Ability 2015 Award: — Primary reference: Assessing the Effect of Visualizations on Bayesian Reasoning through Crowdsourcing (2012).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change stimulus representation or modality; Change the task or response requirement; Adjust the study administration protocol; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators	EXPERIMENT: — DATA: Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: —

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
19	A Psychophysical Investigation of Size as a Physical Variable 2015 Award: — Primary reference: An experiment in graphical perception (1986).	Research team: Independent Explicit claim: No Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>different</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Restrict stimulus or condition coverage; Change stimulus representation or modality; Change the task or response requirement; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface DATA: Collect new observations PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators; Revise uncertainty or robustness assessment	EXPERIMENT: — DATA: Collect additional performance evidence; Collect subjective ratings or preferences; Collect qualitative accounts or recordings; Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons; Analyse background or contextual associations; Analyse additional qualitative accounts
20	Judgment Error in Pie Chart Variations 2016 Award: — Primary reference: Arcs, Angles, or Areas: Individual Data Encodings in Pie and Donut Charts (2016).	Research team: Overlap Explicit claim: No Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>identical</i> ANALYSIS: <i>identical</i>	EXPERIMENT: Expand stimulus or condition coverage; Restrict stimulus or condition coverage; Adjust the study administration protocol DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: — ANALYSIS: —	EXPERIMENT: — DATA: Collect additional performance evidence PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons; Develop or evaluate additional models
21.1	Arcs, Angles, or Areas: Individual Data Encodings in Pie and Donut Charts 2016 Award: — Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: No Result: Consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Restrict stimulus or condition coverage; Change stimulus representation or modality; Change the task or response requirement; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Revise uncertainty or robustness assessment	EXPERIMENT: — DATA: Collect subjective ratings or preferences; Collect qualitative accounts or recordings; Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons; Analyse background or contextual associations

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
21.2	Arcs, Angles, or Areas: Individual Data Encodings in Pie and Donut Charts 2016 Award: — Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: No Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Expand stimulus or condition coverage; Restrict stimulus or condition coverage; Change the task or response requirement; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Revise uncertainty or robustness assessment	EXPERIMENT: — DATA: Collect qualitative accounts or recordings; Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: Analyse background or contextual associations
22.1	The Attraction Effect in Information Visualization 2016 Award: HM (InfoVis 2016) Primary reference: Between a rock and a hard place: The failure of the attraction effect among unattractive alternatives (2013).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Expand stimulus or condition coverage; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change outcome scoring or aggregation; Revise uncertainty or robustness assessment	EXPERIMENT: — DATA: Collect subjective ratings or preferences; Collect qualitative accounts or recordings; Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: Analyse background or contextual associations; Analyse additional qualitative accounts
22.2	The Attraction Effect in Information Visualization 2016 Award: HM (InfoVis 2016) Primary reference: Distinguishing among models of contextually induced preference reversals (1991).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Expand stimulus or condition coverage; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment	EXPERIMENT: — DATA: Collect subjective ratings or preferences; Collect qualitative accounts or recordings; Collect participant-background or ability measures; Record additional operational or contextual metadata PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons; Analyse additional qualitative accounts

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
23	HindSight: Encouraging Exploration through Direct Encoding of Personal Interaction History 2016 Award: — Primary reference: Storytelling in information visualizations: Does it engage users to explore data? (2015).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Evaluate an intervention using a replicated baseline	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change the task or response requirement; Adjust the study administration protocol; Change the experimental structure or paradigm DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: Change outcome scoring or aggregation; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: Add a condition or benchmark comparison DATA: Record additional process or interaction evidence; Collect qualitative accounts or recordings PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons; Analyse additional qualitative accounts
24.1	Blinded with Science or Informed by Charts? A Replication Study 2017 Award: — Primary reference: Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy (2016).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>identical</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Adjust the study administration protocol; Change the study interface DATA: Collect new observations PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Revise uncertainty or robustness assessment	EXPERIMENT: Add a new task or question DATA: Collect additional performance evidence; Collect qualitative accounts or recordings; Collect participant-background or ability measures; Record additional operational or contextual metadata PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons; Synthesise findings across studies
24.2	Blinded with Science or Informed by Charts? A Replication Study 2017 Award: — Primary reference: Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy (2016).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Reassess the robustness of a reported finding	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Restrict stimulus or condition coverage; Change the task or response requirement; Adjust the study administration protocol; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: Collect additional performance evidence; Collect participant-background or ability measures; Record additional operational or contextual metadata PARTICIPANT: — ANALYSIS: Synthesise findings across studies

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
24.3	Blinded with Science or Informed by Charts? A Replication Study 2017 Award: — Primary reference: Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy (2016).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Reassess the robustness of a reported finding	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Expand stimulus or condition coverage; Change the task or response requirement; Adjust the study administration protocol; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: Collect additional performance evidence; Record additional process or interaction evidence; Collect participant-background or ability measures; Record additional operational or contextual metadata PARTICIPANT: — ANALYSIS: Synthesise findings across studies
24.4	Blinded with Science or Informed by Charts? A Replication Study 2017 Award: — Primary reference: Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy (2016).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Reassess the robustness of a reported finding	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change the task or response requirement; Adjust the study administration protocol; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators; Revise uncertainty or robustness assessment	EXPERIMENT: Add a follow-up assessment or questionnaire DATA: Collect subjective ratings or preferences; Collect qualitative accounts or recordings; Record additional operational or contextual metadata PARTICIPANT: Recontact or select existing participants ANALYSIS: Analyse background or contextual associations; Analyse additional qualitative accounts; Synthesise findings across studies

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
25.1	Modeling Color Difference for Visualization Design 2017 Award: BP (InfoVis 2017) Primary reference: An engineering model for color difference as a function of size (2014).	Research team: Overlap Explicit claim: Yes Result: Consistent Main motivation: Reassess the robustness of a reported finding	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>identical</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Restrict stimulus or condition coverage; Change stimulus representation or modality; Adjust the study administration protocol DATA: Collect new observations PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators	EXPERIMENT: — DATA: Collect participant-background or ability measures; Record additional operational or contextual metadata PARTICIPANT: — ANALYSIS: Analyse background or contextual associations
25.2	Modeling Color Difference for Visualization Design 2017 Award: BP (InfoVis 2017) Primary reference: An engineering model for color difference as a function of size (2014).	Research team: Overlap Explicit claim: Yes Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>identical</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Expand stimulus or condition coverage; Change stimulus representation or modality; Adjust the study administration protocol; Change the experimental structure or paradigm DATA: Collect new observations PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators; Change the scope of comparisons	EXPERIMENT: — DATA: Collect participant-background or ability measures; Record additional operational or contextual metadata PARTICIPANT: — ANALYSIS: Analyse background or contextual associations
25.3	Modeling Color Difference for Visualization Design 2017 Award: BP (InfoVis 2017) Primary reference: An engineering model for color difference as a function of size (2014).	Research team: Overlap Explicit claim: Yes Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>identical</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change stimulus representation or modality; Adjust the study administration protocol DATA: Collect new observations PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators; Change the scope of comparisons	EXPERIMENT: — DATA: Collect participant-background or ability measures; Record additional operational or contextual metadata PARTICIPANT: — ANALYSIS: Analyse background or contextual associations

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
26.1	Open vs. Closed Shapes: New Perceptual Categories? 2017 Award: — Primary reference: High perceptual load makes everybody equal: Eliminating individual differences in distractibility with load (2007).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Change stimulus representation or modality; Adjust the study administration protocol; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
26.2	Open vs. Closed Shapes: New Perceptual Categories? 2017 Award: — Primary reference: High perceptual load makes everybody equal: Eliminating individual differences in distractibility with load (2007).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>different</i> PROCEDURE: <i>different</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Restrict stimulus or condition coverage; Change stimulus representation or modality; Change the task or response requirement; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: Add a new task or question DATA: Collect additional performance evidence PARTICIPANT: Recruit a separate extension sample ANALYSIS: Analyse additional outcomes or comparisons

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
27.1	Graphical Perception of Continuous Quantitative Maps: the Effects of Spatial Frequency and Colormap Design 2018 Award: — Primary reference: Color sequences for univariate maps: Theory, experiments and principles (1988).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Expand stimulus or condition coverage; Change stimulus representation or modality; Change the task or response requirement; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators	EXPERIMENT: Add a new task or question DATA: Collect additional performance evidence PARTICIPANT: Recruit a separate extension sample ANALYSIS: Analyse additional outcomes or comparisons; Synthesise findings across studies
28.1	Evaluating 'Graphical Perception' with CNNs 2018 Award: — Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Assess computational observers against established findings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>different</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Expand stimulus or condition coverage; Restrict stimulus or condition coverage; Change the task or response requirement; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Replace the participant agent ANALYSIS: Change outcome scoring or aggregation; Revise uncertainty or robustness assessment	EXPERIMENT: Add a condition or benchmark comparison; Add a method or model evaluation DATA: Collect additional performance evidence; Record additional process or interaction evidence PARTICIPANT: Recruit a separate extension sample ANALYSIS: Develop or evaluate additional models

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
28.2	Evaluating 'Graphical Perception' with CNNs 2018 Award: — Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Assess computational observers against established findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>different</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Replace the participant agent ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment	EXPERIMENT: Add a method or model evaluation DATA: Collect additional performance evidence; Record additional process or interaction evidence PARTICIPANT: — ANALYSIS: Develop or evaluate additional models
28.3	Evaluating 'Graphical Perception' with CNNs 2018 Award: — Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Assess computational observers against established findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>different</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Replace the participant agent ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment	EXPERIMENT: Add a method or model evaluation DATA: Collect additional performance evidence PARTICIPANT: — ANALYSIS: Develop or evaluate additional models
28.4	Evaluating 'Graphical Perception' with CNNs 2018 Award: — Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Assess computational observers against established findings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>different</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>different</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>different</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change the task or response requirement; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations; Replace the primary evidence type PARTICIPANT: Replace the participant agent ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Revise uncertainty or robustness assessment	EXPERIMENT: Add a new task or question; Add a condition or benchmark comparison DATA: Collect additional performance evidence PARTICIPANT: Recruit a separate extension sample ANALYSIS: Analyse additional outcomes or comparisons

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
29	VizNet: Towards A Large-Scale Visualization Learning and Benchmarking Repository 2019 Award: — Primary reference: Assessing Effects of Task and Data Distribution on the Effectiveness of Visual Encodings (2018).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Restrict stimulus or condition coverage; Adjust the study administration protocol; Change the study interface DATA: Collect new observations PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment	EXPERIMENT: Add a new task or question; Add a method or model evaluation DATA: Collect additional performance evidence; Derive additional features or evidence representations PARTICIPANT: — ANALYSIS: Develop or evaluate additional models
30	Correlation Judgment and Visualization Features: A Comparative Study 2019 Award: — Primary reference: Ranking Visualizations of Correlation Using Weber's Law (2014).	Research team: Overlap Explicit claim: Yes Result: Partly consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>identical</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>identical</i> ANALYSIS: <i>different</i>	EXPERIMENT: Adjust the study administration protocol DATA: Collect new observations PARTICIPANT: — ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators; Revise uncertainty or robustness assessment	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
31	Revisited: Comparison of Empirical Methods to Evaluate Visualizations Supporting Crafting and Assembly Purposes 2020 Award: N/A Primary reference: Online, VR, AR, Lab, and In-Situ: Comparison of Research Methods to Evaluate Smart Artifacts (2019).	Research team: Overlap Explicit claim: Yes Result: Not consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>different</i> TASK: <i>different</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>similar</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Change stimulus representation or modality; Change the task or response requirement; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change outcome scoring or aggregation; Revise uncertainty or robustness assessment; Change qualitative or visual analysis	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
32	Pushing the (Visual) Narrative: the Effects of Prior Knowledge Elicitation in Provocative Topics 2020 Award: — Primary reference: The persuasive power of data visualization (2014).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Evaluate an intervention using a replicated baseline	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change stimulus representation or modality; Change the task or response requirement; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Revise uncertainty or robustness assessment	EXPERIMENT: Add a condition or benchmark comparison DATA: Collect qualitative accounts or recordings PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons; Analyse additional qualitative accounts
33	Understanding Data Accessibility for People with Intellectual and Developmental Disabilities 2021 Award: BP (CHI 2021) Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: No Result: Consistent Main motivation: Examine generalisability across human populations	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change the task or response requirement; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators	EXPERIMENT: Add a new task or question; Add a condition or benchmark comparison; Add a qualitative inquiry component DATA: Collect additional performance evidence; Collect qualitative accounts or recordings PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons; Analyse additional qualitative accounts

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
34	Effect of uncertainty visualizations on myopic loss aversion and the equity premium puzzle in retirement investment decisions 2021 Award: N/A Primary reference: Risk aversion or myopia? Choices in repeated gambles and retirement investments (1999).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Reassess the robustness of a reported finding	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change stimulus representation or modality; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Revise uncertainty or robustness assessment	EXPERIMENT: Add a condition or benchmark comparison DATA: Collect additional performance evidence; Collect qualitative accounts or recordings PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons; Analyse additional qualitative accounts
35	Rethinking the Ranks of Visual Channels 2021 Award: HM (VIS 2021) Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>different</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>identical</i> DATA: <i>different</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>different</i>	EXPERIMENT: Expand stimulus or condition coverage; Change the task or response requirement; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface DATA: Collect new observations; Replace the primary evidence type PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators; Revise uncertainty or robustness assessment	EXPERIMENT: — DATA: Collect additional performance evidence PARTICIPANT: — ANALYSIS: —

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
36.1	Evaluating the Effects of Visual Embellishments and Context on Comprehension, Recall and Preference of Charts 2022 Award: — Primary reference: Useful junk? The effects of visual embellishment on comprehension and memorability of charts (2010).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Reassess the robustness of a reported finding	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Restrict stimulus or condition coverage; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: Collect qualitative accounts or recordings PARTICIPANT: — ANALYSIS: Analyse additional qualitative accounts
36.2	Evaluating the Effects of Visual Embellishments and Context on Comprehension, Recall and Preference of Charts 2022 Award: — Primary reference: Useful junk? The effects of visual embellishment on comprehension and memorability of charts (2010).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Expand stimulus or condition coverage; Restrict stimulus or condition coverage; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: Collect qualitative accounts or recordings PARTICIPANT: — ANALYSIS: Analyse additional qualitative accounts

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
37	Juvenile Graphical Perception: A Comparison between Children and Adults 2022 Award: N/A Primary reference: Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design (2010).	Research team: Independent Explicit claim: No Result: Partly consistent Main motivation: Examine generalisability across human populations	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>similar</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>different</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Restrict stimulus or condition coverage; Change stimulus representation or modality; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators; Change the scope of comparisons	EXPERIMENT: — DATA: Derive additional features or evidence representations PARTICIPANT: — ANALYSIS: Develop or evaluate additional models
38	VoxLens: Making Online Data Visualizations Accessible with an Interactive JavaScript Plug-In 2022 Award: — Primary reference: Understanding Screen-Reader Users' Experiences with Online Data Visualizations (Study 2) (2021).	Research team: Overlap Explicit claim: No Result: Consistent Main motivation: Evaluate an intervention using a replicated baseline	STIMULI: <i>identical</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface DATA: Collect new observations; Reuse existing evidence PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change the scope of comparisons	EXPERIMENT: Add a formative or pilot study DATA: Record additional process or interaction evidence; Collect subjective ratings or preferences; Collect qualitative accounts or recordings PARTICIPANT: Recruit a separate extension sample; Recontact or select existing participants ANALYSIS: Analyse additional outcomes or comparisons; Analyse additional qualitative accounts
39.1	More Useful Junk? Replicating the Effects of Visual Embellishment on Description and Recall of Charts 2023 Award: — Primary reference: Useful junk? The effects of visual embellishment on comprehension and memorability of charts (2010).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Reassess the robustness of a reported finding	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>identical</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change the task or response requirement; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
39.2	More Useful Junk? Replicating the Effects of Visual Embellishment on Description and Recall of Charts 2023 Award: — Primary reference: Useful junk? The effects of visual embellishment on comprehension and memorability of charts (2010).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Reassess the robustness of a reported finding	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>identical</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>identical</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: — ANALYSIS: Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
39.3	More Useful Junk? Replicating the Effects of Visual Embellishment on Description and Recall of Charts 2023 Award: — Primary reference: Useful junk? The effects of visual embellishment on comprehension and memorability of charts (2010).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Reassess the robustness of a reported finding	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>identical</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Restrict stimulus or condition coverage; Change the task or response requirement; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: — ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
39.4	More Useful Junk? Replicating the Effects of Visual Embellishment on Description and Recall of Charts 2023 Award: — Primary reference: Useful junk? The effects of visual embellishment on comprehension and memorability of charts (2010).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Reassess the robustness of a reported finding	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>identical</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Restrict stimulus or condition coverage; Change the task or response requirement; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: — ANALYSIS: Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: Collect qualitative accounts or recordings PARTICIPANT: — ANALYSIS: Analyse additional qualitative accounts; Synthesise findings across studies

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
40	The Risks of Ranking: Revisiting Graphical Perception to Model Individual Differences in Visualization Performance 2024 Award: — Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Develop or refine a model of performance	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>different</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Restrict stimulus or condition coverage; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Analyse individual differences or moderators; Revise uncertainty or robustness assessment	EXPERIMENT: — DATA: Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: —
41	Glanceable Data Visualizations for Older Adults: Establishing Thresholds and Examining Disparities Between Age Groups 2024 Award: Honorable Mention (CHI 2024) Primary reference: Glanceable Visualization: Studies of Data Comparison Performance on Smartwatches (2018).	Research team: Overlap Explicit claim: Yes Result: Partly consistent Main motivation: Examine generalisability across human populations	STIMULI: <i>identical</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Adjust the study administration protocol; Change the study interface DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: Analyse individual differences or moderators; Change the scope of comparisons	EXPERIMENT: — DATA: Collect qualitative accounts or recordings; Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: Analyse background or contextual associations
42	Transparent Risks: The Impact of the Specificity and Visual Encoding of Uncertainty on Decision Making 2024 Award: — Primary reference: Evaluating the impact of visualization of wildfire hazard upon decision-making under uncertainty (2016).	Research team: Independent Explicit claim: No Result: Partly consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Expand stimulus or condition coverage; Restrict stimulus or condition coverage; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change outcome scoring or aggregation; Change the scope of comparisons	EXPERIMENT: Add a condition or benchmark comparison DATA: Collect subjective ratings or preferences; Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
43.1	Doorways Do Not Always Cause Forgetting: Studying the Effect of Locomotion Technique and Doorway Visualization in Virtual Reality 2024 Award: — Primary reference: Walking through doorways causes forgetting: Situation models and experienced space (2006).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change stimulus representation or modality; Change the task or response requirement; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
43.2	Doorways Do Not Always Cause Forgetting: Studying the Effect of Locomotion Technique and Doorway Visualization in Virtual Reality 2024 Award: — Primary reference: Walking through doorways causes forgetting: Situation models and experienced space (2006).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Expand stimulus or condition coverage; Change stimulus representation or modality; Change the task or response requirement; Adjust the study administration protocol; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
44	Validation through Replication of Augmented Reality as a Visualization Technique for Scholarly Publications in Astronomy 2025 Award: — Primary reference: Augmented Reality as a Visualization Technique for Scholarly Publications in Astronomy: An Empirical Evaluation (2023).	Research team: Overlap Explicit claim: Yes Result: Partly consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change the task or response requirement; Adjust the study administration protocol; Change the study interface DATA: Collect new observations PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: Add a follow-up assessment or questionnaire DATA: Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: Analyse background or contextual associations

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
45.1	Seeing Eye to AI? Applying Deep-Feature-Based Similarity Metrics to Information Visualization 2025 Award: — Primary reference: Discriminability Tests for Visualization Effectiveness and Scalability (2019).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Assess computational observers against established findings	STIMULI: <i>identical</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>identical</i>	EXPERIMENT: Adjust the study administration protocol; Change the study interface DATA: Collect new observations; Reuse existing evidence PARTICIPANT: Replace the participant agent ANALYSIS: Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: Add a method or model evaluation DATA: Derive additional features or evidence representations PARTICIPANT: — ANALYSIS: Develop or evaluate additional models
45.2	Seeing Eye to AI? Applying Deep-Feature-Based Similarity Metrics to Information Visualization 2025 Award: — Primary reference: Learning Perceptual Kernels for Visualization Design (2014).	Research team: Independent Explicit claim: Yes Result: Partly consistent Main motivation: Assess computational observers against established findings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>different</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>identical</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Restrict stimulus or condition coverage; Change the task or response requirement; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected; Reuse existing evidence PARTICIPANT: Replace the participant agent ANALYSIS: Change the scope of comparisons	EXPERIMENT: Add a method or model evaluation DATA: Derive additional features or evidence representations PARTICIPANT: — ANALYSIS: Develop or evaluate additional models
46.1	ReVISit 2: A Full Experiment Life Cycle User Study Framework 2025 Award: BP (VIS 2025) Primary reference: Ranking Visualizations of Correlation Using Weber's Law (2014).	Research team: Overlap Explicit claim: Yes Result: Consistent Main motivation: Evaluate a research method or tool	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Restrict stimulus or condition coverage; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface DATA: Collect new observations PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: Add a condition or benchmark comparison DATA: Collect additional performance evidence; Record additional operational or contextual metadata PARTICIPANT: — ANALYSIS: Analyse additional outcomes or comparisons

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
46.2	ReVISit 2: A Full Experiment Life Cycle User Study Framework 2025 Award: BP (VIS 2025) Primary reference: Design Characterization for Black-and-White Textures in Visualization (2023).	Research team: Overlap Explicit claim: Yes Result: Consistent Main motivation: Evaluate a research method or tool	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Adjust the study administration protocol; Change the study interface DATA: Collect new observations PARTICIPANT: Adjust sampling and recruitment ANALYSIS: —	EXPERIMENT: Add a method or model evaluation DATA: Record additional process or interaction evidence; Collect subjective ratings or preferences PARTICIPANT: — ANALYSIS: Analyse background or contextual associations; Assess operational feasibility or study process
46.3	ReVISit 2: A Full Experiment Life Cycle User Study Framework 2025 Award: BP (VIS 2025) Primary reference: The Effects of Adding Search Functionality to Interactive Visualizations on the Web (2018).	Research team: Overlap Explicit claim: Yes Result: Consistent Main motivation: Evaluate a research method or tool	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Restrict stimulus or condition coverage; Change the study interface DATA: Collect new observations PARTICIPANT: — ANALYSIS: Change statistical tests or models	EXPERIMENT: Add a qualitative inquiry component; Add a method or model evaluation DATA: Record additional process or interaction evidence; Collect qualitative accounts or recordings PARTICIPANT: — ANALYSIS: Analyse additional qualitative accounts; Assess operational feasibility or study process
47	They Aren't Built For Me: An Exploratory Study of Strategies for Measurement of Graphical Primitives in Tactile Graphics 2025 Award: BP (VIS 2025) Primary reference: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods (1984).	Research team: Independent Explicit claim: Yes Result: Consistent Main motivation: Examine generalisability across human populations	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>different</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change stimulus representation or modality; Adjust the study administration protocol; Change the study interface DATA: Collect new observations PARTICIPANT: Change the target human population ANALYSIS: Change statistical tests or models; Analyse individual differences or moderators	EXPERIMENT: Add a qualitative inquiry component DATA: Collect additional performance evidence; Collect qualitative accounts or recordings; Collect participant-background or ability measures PARTICIPANT: — ANALYSIS: Analyse additional qualitative accounts

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
48.1	Characterizing Visualization Perception with Psychological Phenomena: Uncovering the Role of Subitizing in Data Visualization 2025 Award: HM (VIS 2025) Primary reference: Measuring categorical perception in color-coded scatterplots (2023).	Research team: Overlap Explicit claim: Yes Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>different</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Expand stimulus or condition coverage; Restrict stimulus or condition coverage; Change the task or response requirement; Adjust the study administration protocol; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change the scope of comparisons	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
48.2	Characterizing Visualization Perception with Psychological Phenomena: Uncovering the Role of Subitizing in Data Visualization 2025 Award: HM (VIS 2025) Primary reference: Measuring categorical perception in color-coded scatterplots (2023).	Research team: Overlap Explicit claim: Yes Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>identical</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Expand stimulus or condition coverage; Restrict stimulus or condition coverage; Change the task or response requirement; Adjust the study administration protocol; Change the study interface DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change the scope of comparisons	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
48.3	Characterizing Visualization Perception with Psychological Phenomena: Uncovering the Role of Subitizing in Data Visualization 2025 Award: HM (VIS 2025) Primary reference: Improving the robustness of scagnostics (2019).	Research team: Overlap Explicit claim: Yes Result: Consistent Main motivation: Test generalisability across designs or settings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>similar</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>different</i>	EXPERIMENT: Expand stimulus or condition coverage; Restrict stimulus or condition coverage; Change stimulus representation or modality; Change the task or response requirement; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Change the target human population; Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Change the scope of comparisons; Change qualitative or visual analysis	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: Synthesise findings across studies

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
49.1	Do Graph Drawing Aesthetics Matter for AI? A Replication of Foundational Studies in Graph Readability 2026 Award: — Primary reference: Which aesthetic has the greatest effect on human understanding? (1997).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Assess computational observers against established findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Replace the participant agent ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
49.2	Do Graph Drawing Aesthetics Matter for AI? A Replication of Foundational Studies in Graph Readability 2026 Award: — Primary reference: Performance of layout algorithms: Comprehension, not computation (1998).	Research team: Independent Explicit claim: Yes Result: Not consistent Main motivation: Assess computational observers against established findings	STIMULI: <i>similar</i> TASK: <i>identical</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Replace the participant agent ANALYSIS: Change statistical tests or models; Revise uncertainty or robustness assessment; Change the scope of comparisons	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
50.1	Can LLMs Simulate Target Users in Visualization Case Studies? 2026 Award: — Primary reference: Parallel Edge Splatting for Scalable Dynamic Graph Visualization (2011).	Research team: Overlap Explicit claim: Yes Result: Partly consistent Main motivation: Assess computational observers against established findings	STIMULI: <i>identical</i> TASK: <i>similar</i> PROCEDURE: <i>different</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Change the task or response requirement; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Replace the participant agent ANALYSIS: Change the scope of comparisons; Change qualitative or visual analysis	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —

Continued on next page

Table 3 (continued)

Study	Paper and reference	Context and result	DIMENSIONS	Design changes	Design additions
50.2	Can LLMs Simulate Target Users in Visualization Case Studies? 2026 Award: — Primary reference: Visual Analysis of Document Editing Patterns (2025).	Research team: Overlap Explicit claim: Yes Result: Partly consistent Main motivation: Assess computational observers against established findings	STIMULI: <i>identical</i> TASK: <i>similar</i> PROCEDURE: <i>different</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>different</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Change the task or response requirement; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations PARTICIPANT: Replace the participant agent ANALYSIS: Change the scope of comparisons; Change qualitative or visual analysis	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
51.1	Taking a Walk on the Wild Side: Effects of Walking in Synchrony with Pitch-Altered Footstep Sounds on Body Perception in Outside the Lab Contexts 2026 Award: — Primary reference: As light as your footsteps: Altering walking sounds to change perceived body weight, emotional state and gait (2015).	Research team: Overlap Explicit claim: Yes Result: Partly consistent Main motivation: Reassess the robustness of a reported finding	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>similar</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>similar</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>similar</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Change the task or response requirement; Adjust the study administration protocol; Change the study interface; Change the study setting DATA: Collect new observations; Revise the evidence measures collected PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change statistical tests or models; Analyse individual differences or moderators; Revise uncertainty or robustness assessment	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —
51.2	Taking a Walk on the Wild Side: Effects of Walking in Synchrony with Pitch-Altered Footstep Sounds on Body Perception in Outside the Lab Contexts 2026 Award: — Primary reference: As light as your footsteps: Altering walking sounds to change perceived body weight, emotional state and gait (2015).	Research team: Overlap Explicit claim: Yes Result: Partly consistent Main motivation: Explain effects or conflicting findings	STIMULI: <i>similar</i> TASK: <i>similar</i> PROCEDURE: <i>different</i> INTERFACE: <i>different</i> ENVIRONMENT: <i>different</i> DATA: <i>different</i> PARTICIPANT: <i>similar</i> ANALYSIS: <i>different</i>	EXPERIMENT: Reconstruct or adapt stimulus materials; Restrict stimulus or condition coverage; Change the task or response requirement; Adjust the study administration protocol; Change the experimental structure or paradigm; Change the study interface; Change the study setting DATA: Collect new observations; Replace the primary evidence type PARTICIPANT: Adjust sampling and recruitment ANALYSIS: Change qualitative or visual analysis	EXPERIMENT: — DATA: — PARTICIPANT: — ANALYSIS: —