VIS 2018

Correlation Judgment and Visualization Features

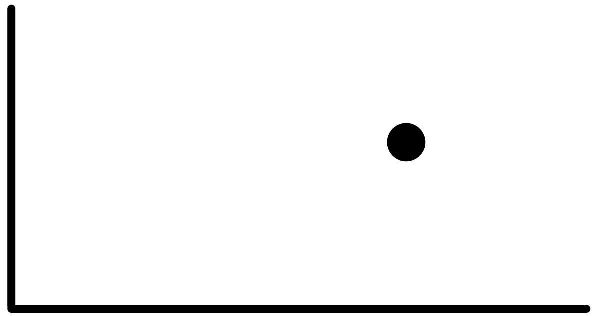
A COMPARATIVE STUDY

Fumeng Yang	Brown University
Lane Harrison	Worcester Polytechnic Institute
Ronald Rensink	University of British Columbia
Steven Franconeri	Northwestern University
Remco Chang	Tufts University

data and code are available at
https://github.com/Fumeng-Yang/VisualFeature_TVCG

DATA MAPPING

(25, 10)

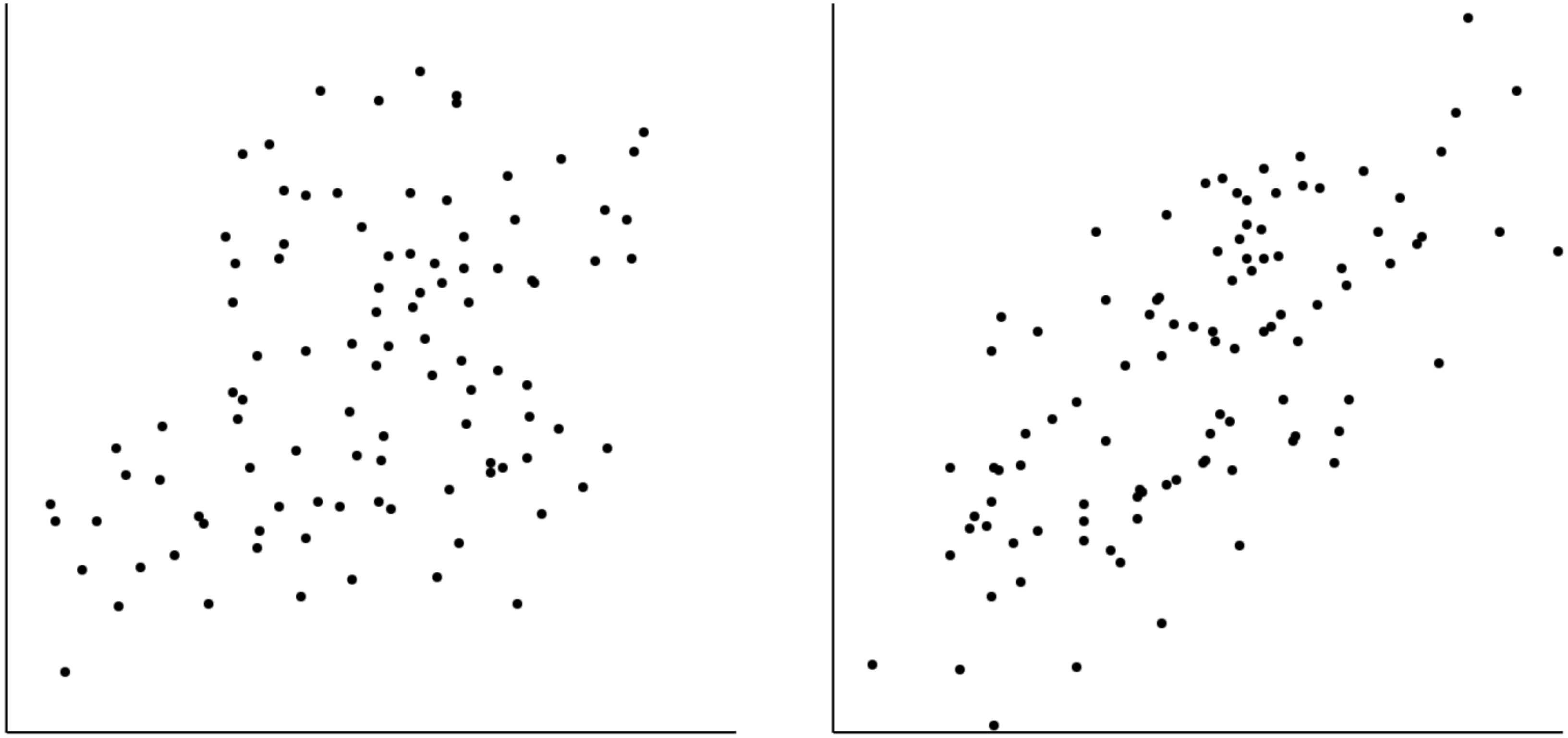


data element

visual element

CORRELATION VISUALIZATION

Which one shows the more correlated dataset?



Weber's law (linear) for correlation perception

2010

[RENSINK 2010]

Weber's law applies to other visualizations

2014

[RENSINK 2014, HARRISON 2014]

MODELING

Correlation Perception

Log-transformation for individual observations

2015

[KAY 2015]

Why can Weber's law describe correlation perception?

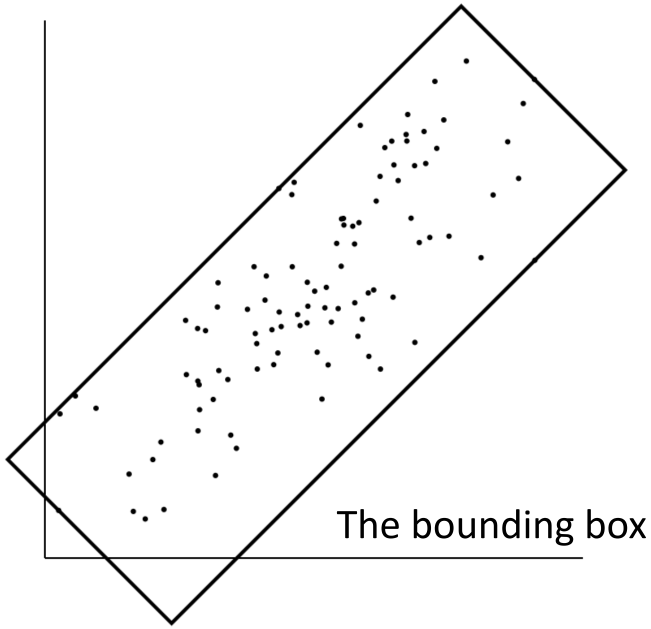
[WEBER'S LAW IS FOR LOW-LEVEL PERCEPTION, e.g., LENGTH]

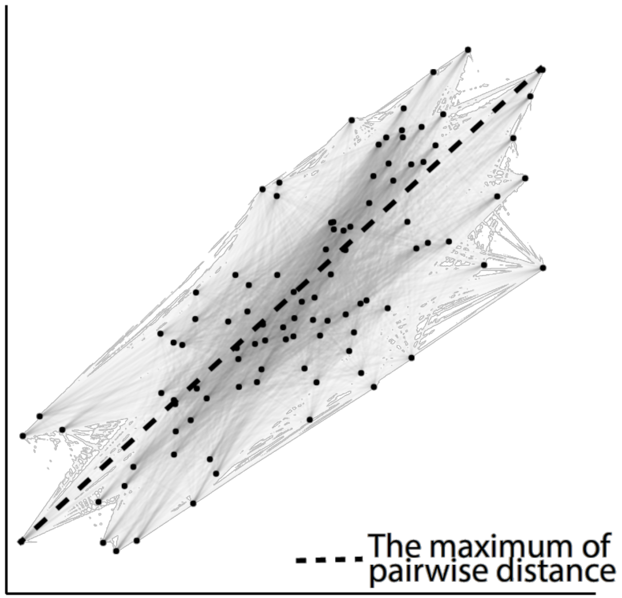
RESEARCH QUESTIONS

Do we have a ground theory to describe the data falls outside Weber's law?

[INDIVIDUAL OBVERSATIONS, LOG-TRANSFORMATION, etc.]

Visual Feature





[DEFINITION]

The perceivable and distinguishable properties (e.g., shape, dispersion, and orientation) in a 2D image or a part of an image.

OUR HYPOTHESIS

VISUAL FEATURE IS A PROXY OF CORRELATION

[WHEN COMPARING CORRELATION IN TWO SCATTERPLOTS]

WHICH VISUAL FEATURE(S) AND HOW?

Overview

Part 01 **Collecting Visual Features**

The background of visual features

Part 02 **Identifying Visual Features Used in Correlation Judgments**

The one that best aligns with the judgments

Part 03 **Using Visual Features to Describe Correlation Perception**

Connecting visual features to the modeling of correlation perception in scatterplots

Part 04 **Implications**

What can we learn from this study?

01

PART

Collecting Visual Features

Collecting Visual Features

LITERATURE

- Visualization
- Perceptual psychology
- Statistics
- Computational geometry
- ...

→ 49 visual features in scatterplots →

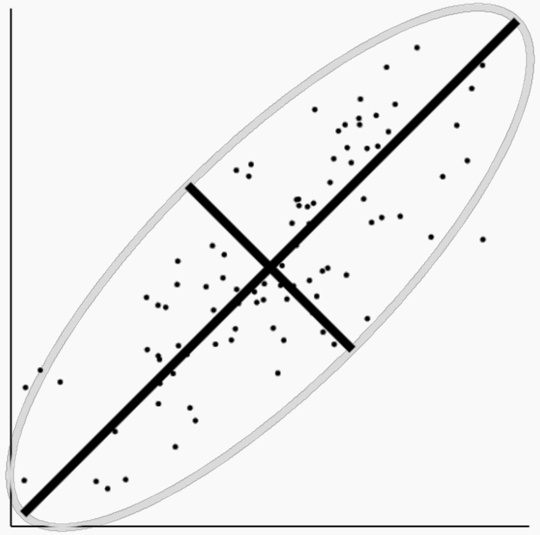
CATEGORIES

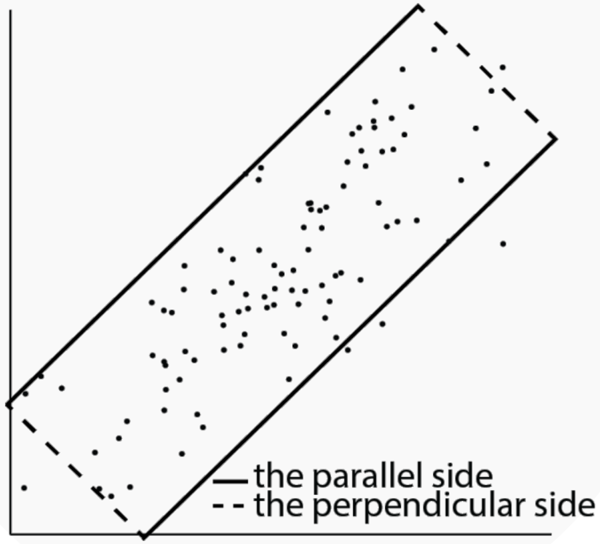
- Length
- Shape
- Area
- Density

Length

1. The minor axis of prediction ellipse
2. The major axis of prediction ellipse
3. The parallel side of bounding box
4. The perpendicular side of bounding box
5. The parallel side of confidence box
6. The perpendicular side of confidence box
7. …

[JENNINGS 1982, CLEVELAND 1982, MEYER 1992, etc.]

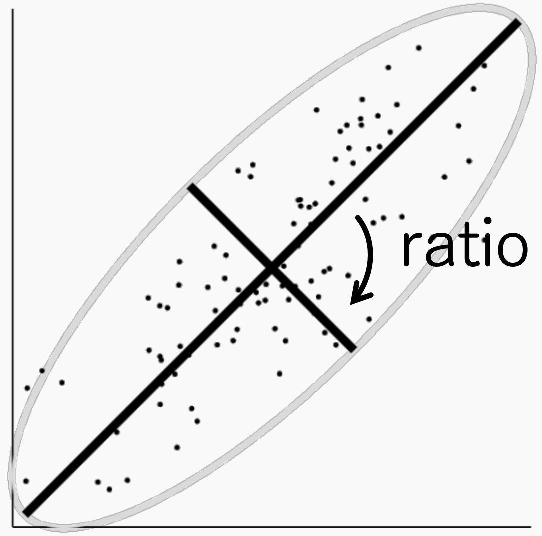


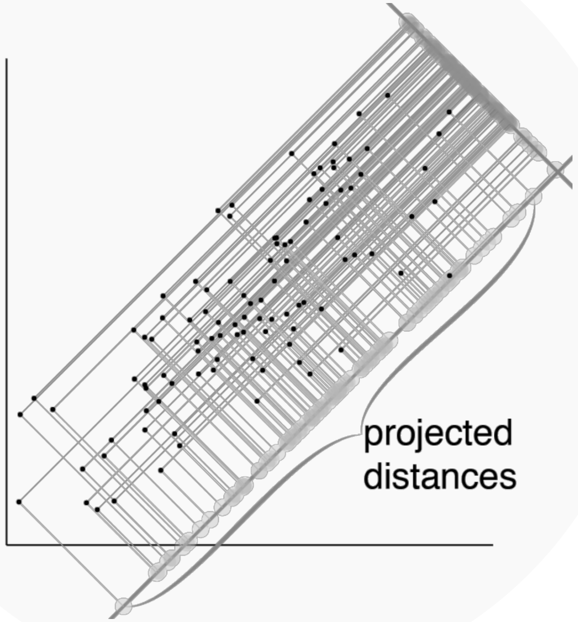


Shape

1. The ratio of a minor axis to a major axis
2. The ratio of a major axis to a minor axis
3. The skewness of the distances
4. The SD of the projections
5. …

[JENNINGS 1982, WILKINSON 2005, etc.]

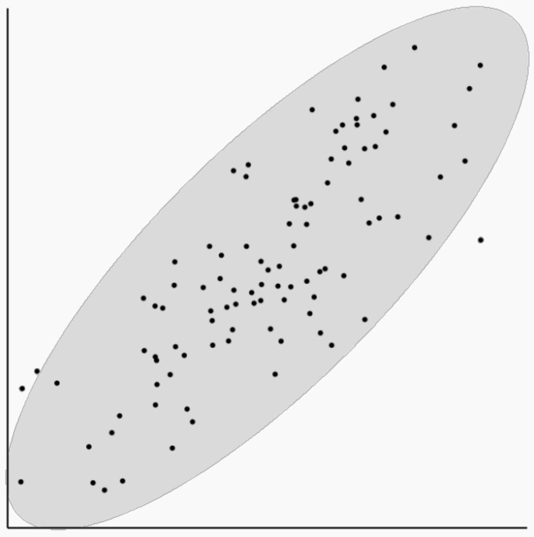


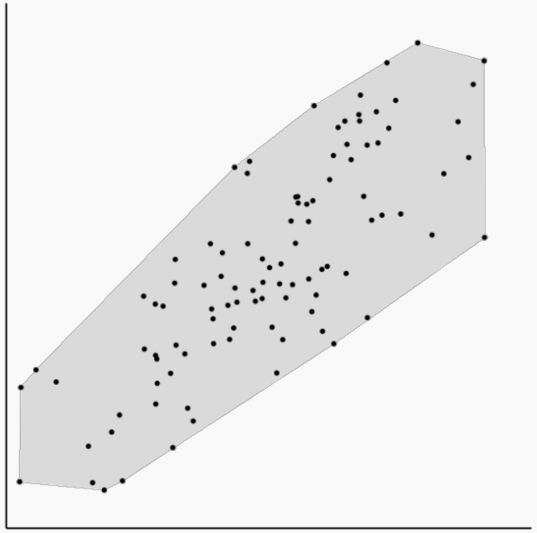


Area

1. The area of prediction ellipse
2. The area of bounding box
3. The area of confidence box
4. The area of convex hull
5. …

[CLEVELAND 1982, MEYER 1992, etc.]

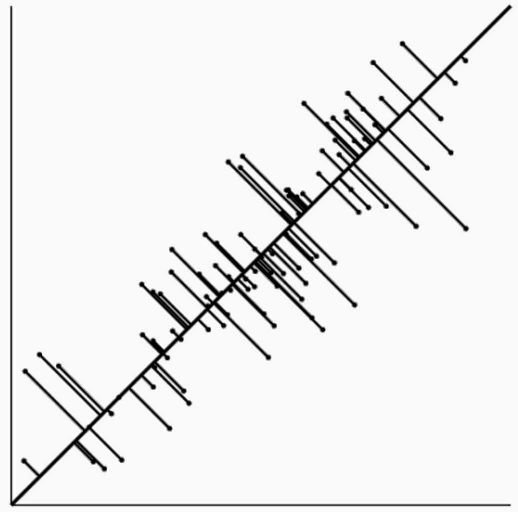


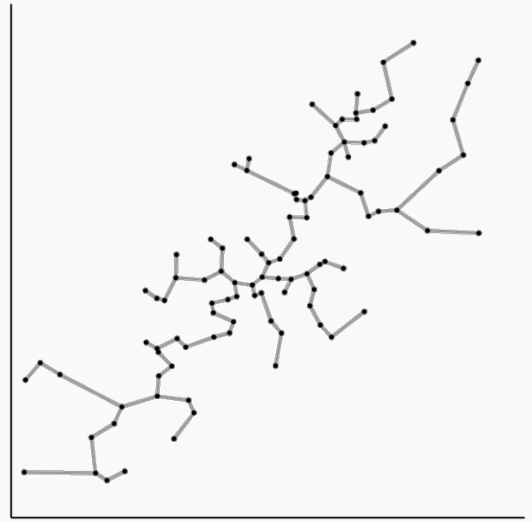


Density

1. The SD of the edges on MST
2. The average of all the inverted distances
3. The average of all the distances to the line
4. The SD of all the distances to the line
5. …

[LAUER 1989, WILKINSON 2005, etc.]





02

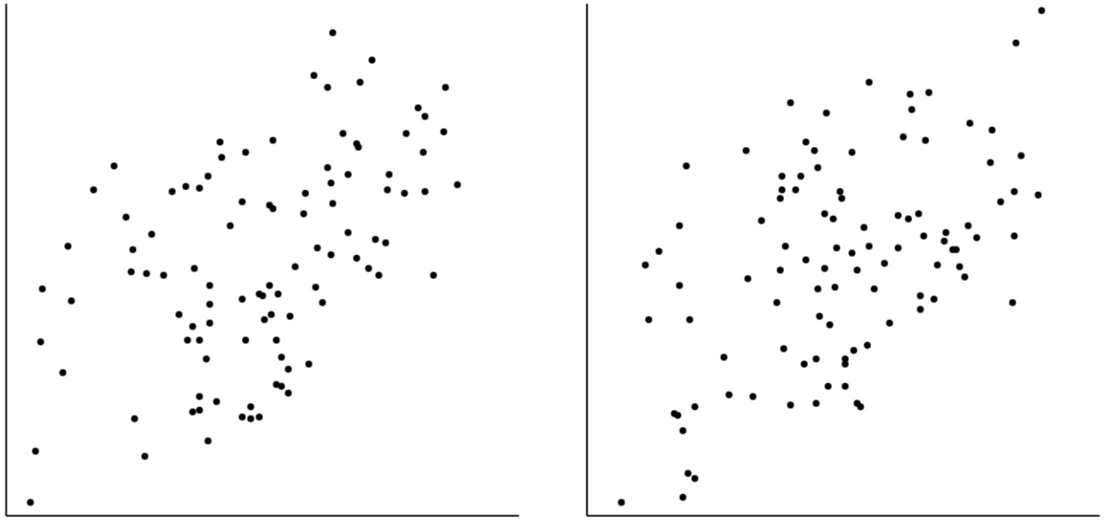
PART

Identifying Visual Features Used in Correlation Judgments

Collecting Judgments

Replication Experiment

REPLICATING HARRISON 2014

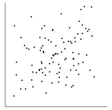
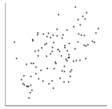


WHICH ONE IS MORE CORRELATED?

$r = \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}$

Collecting Judgments

EACH PARTICIPANT

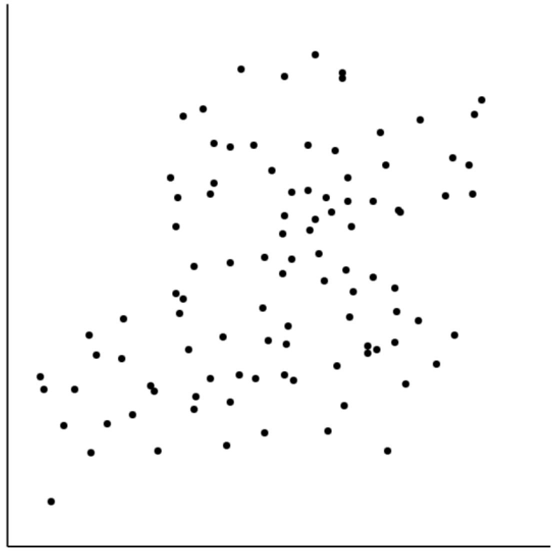
TWO r		DIRECTION	{r, DIRECTION}
0.3		above (0.5 vs. 0.55)	50 judgments
0.5		below (0.5 vs. 0.45)	[STAIRCASE METHOD]

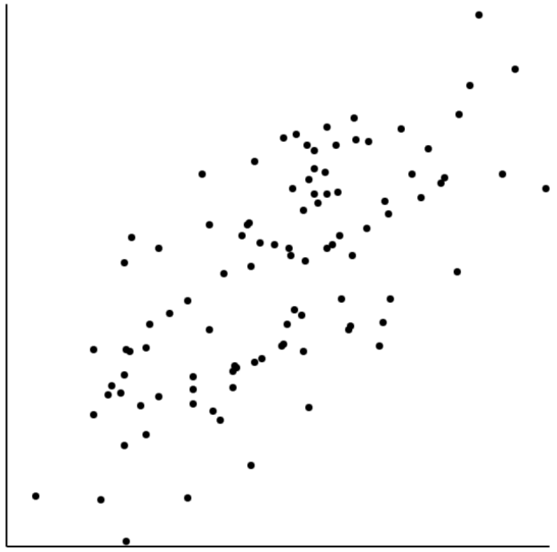
2 r values x 2 directions x 95 participants (AMT) = 19,000 judgments

[SEE HARRISON 2014 AND OUR PAPER]

FROM FEATURES TO JUDGMENTS

One judgment $\longrightarrow$ {Correct, Incorrect}



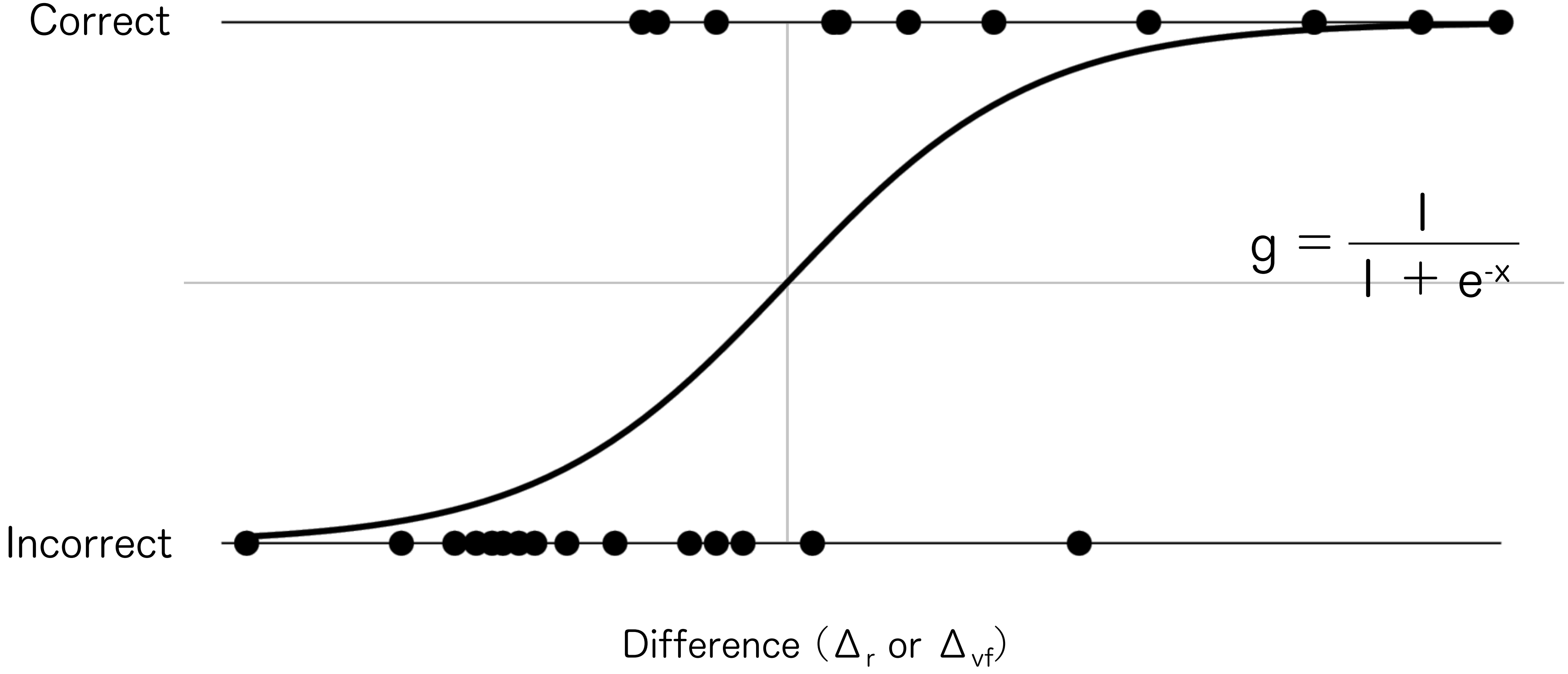


$| \ r \ - \ r \ | = \text{difference } (\Delta_r)$

$| \ \text{feature} \ - \ \text{feature} \ | = \text{difference } (\Delta_{vf})$

LOGISTIC REGRESSION

[HOSMER 1957]



$$g = \frac{1}{1 + e^{-x}}$$

LOGISTIC REGRESSION

$$\Delta r \;—\; g_r = \beta_0 + \beta_1 a_i + \beta_2 r_i + \beta_3 \Delta_i$$

$$\Delta vf \left\{ \begin{array}{l} g_1 = \beta_0 + \beta_1 a_i + \beta_2 r_i + \beta_3 \Delta_i \\ \cdots \\ g_{49} = \beta_0 + \beta_1 a_i + \beta_2 r_i + \beta_3 \Delta_i \end{array} \right.$$

[DIRECTION]

The **visual feature** should **better** explain error in people's judgments than **correlation.**

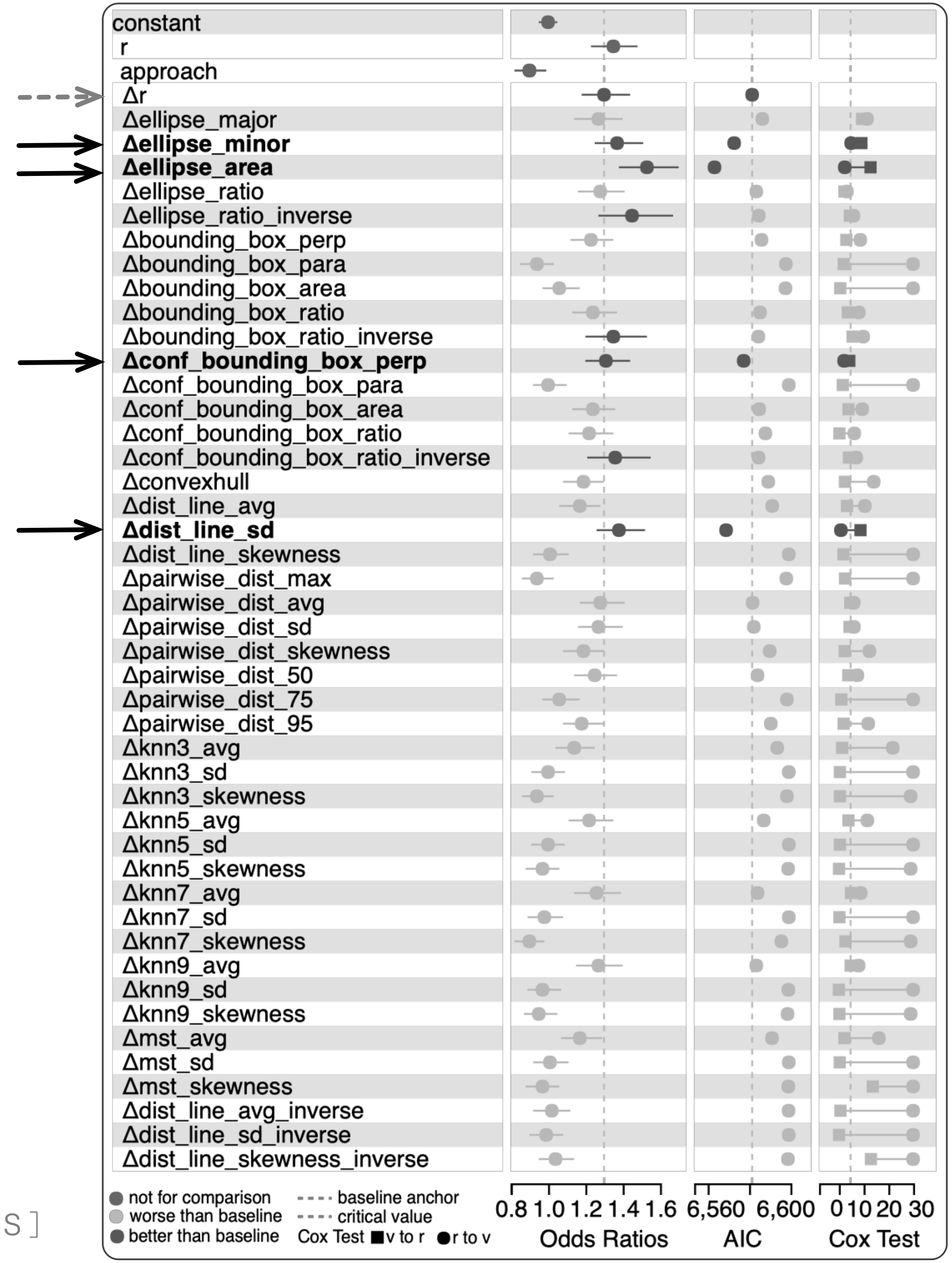
[WEIGHTED ON CORRECT/INCORRECT]

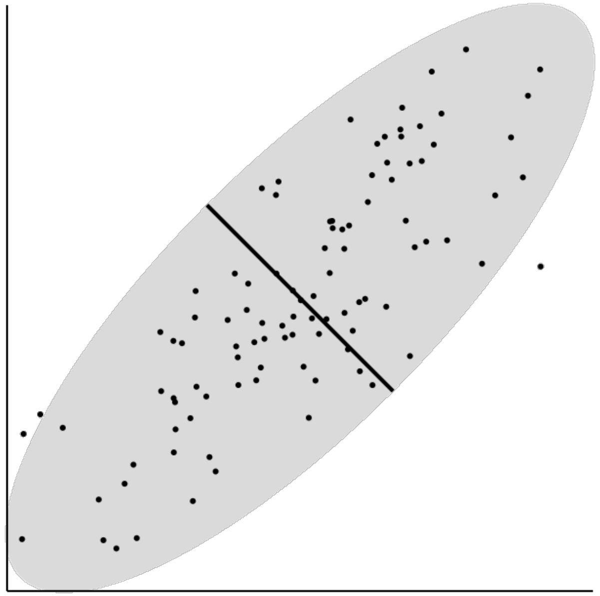
[STANDARDIZED MODELS]

METRICS

- Odds ratios
- AIC
- Cox test
- …

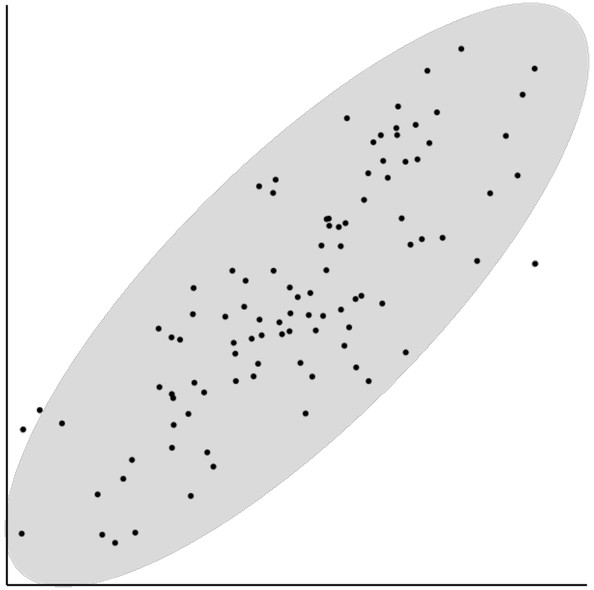
[SEE THE SUPPLEMENTARY MATERIALS FOR OTHER METRICS]

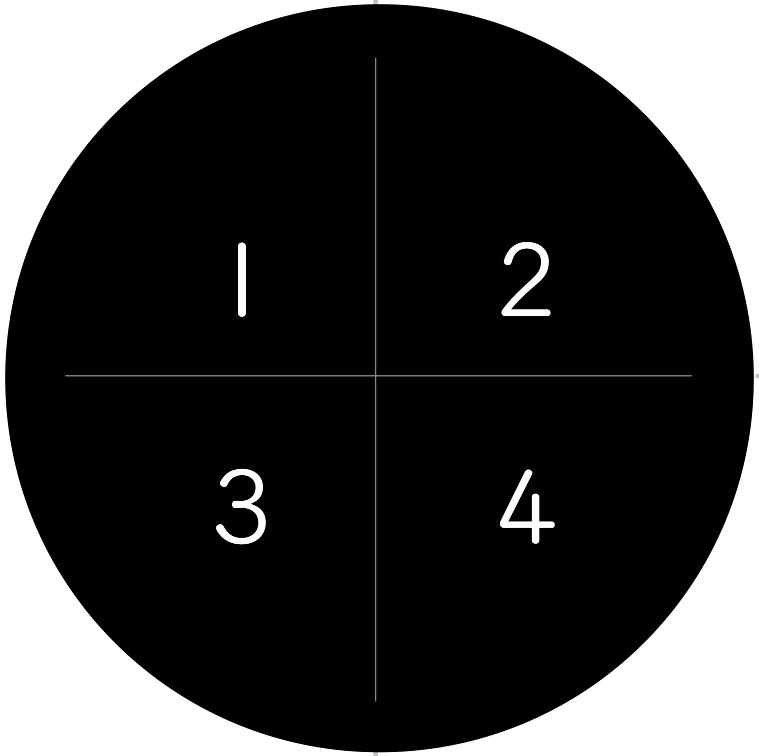


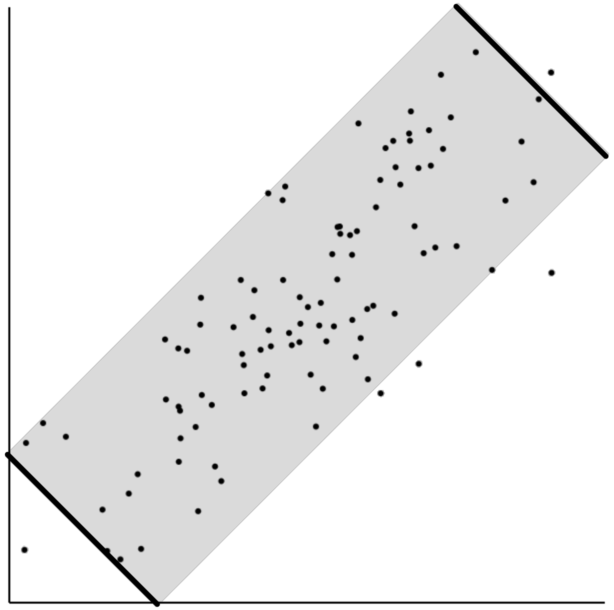


The minor axis of prediction ellipse

The area of prediction ellipse

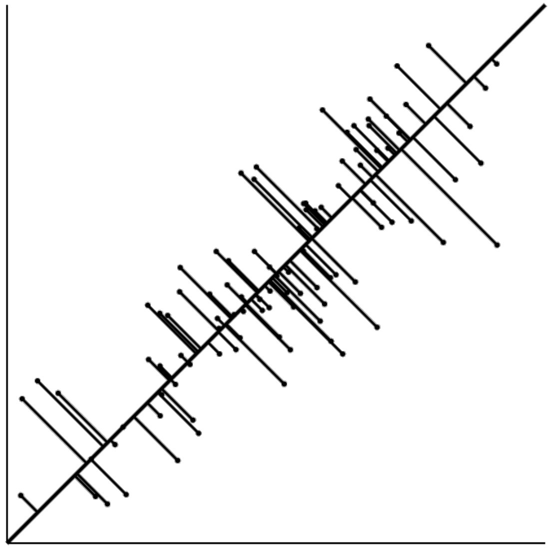






The perpendicular side of confidence bounding box

The SD of all the distances to the regression line



03

PART

Using Visual Features to Describe Correlation Perception

Weber's law (linear) for correlation perception

$JND = \beta_0 + \beta_1 r$

2010

[RENSINK 2010]

Weber's law (linear) to rank visualizations

$JND = \beta_0 + \beta_1 r$

2014

[HARRISON 2014]

VISUAL FEATURE

extend

JND

Just-Noticeable Difference

[THE DIFFERENCE THAT PEOPLE CAN TELL 75/50% OF THE TIME]

[Example: the SD of all the distances]

Log-transformation for individual observations

$\log(JND) = \beta_0 + \beta_1 a + \beta_2 r + U$

2015

[KAY 2015]

REPLICATING

Weber's law (linear)

[RENSINK 2010, 2014, & HARRISON 2014]

Log-transformation

[KAY 2015]

CORRELATION

VISUAL FEATURE

PROPOSING

Power-transformation

Linear Model (Weber's law) for Mean Observations

[RENSINK 2010 & 2014, HARRISON 2014]

$$JND_r = \beta_0 + \beta_1 r$$

$$JND_{vf} = \beta_0 + \beta_1 v$$

Linear Model (Weber's law) for Mean Observations

[RENSINK 2010 & 2014, HARRISON 2014]

Model	Correlation Coefficients	Coefficients				R2	AIC
		$\beta 0$	p	$\beta 1$	p		
f1: JNDr ~ r	-.9778	0.1860	< .001	-0.1791	< .001	.9561	-79.7720
f2: JNDv ~ v	.9011	0.7975	.0101	0.0708	< .001	.8119	-3.0762

Both the visual feature and correlation can be described by Weber's law.
A single visual feature fairly describes the data from 95 participants.

Log-transformation for Individual Observations

[KAY 2015]

TO CORRECT SKEWED DATA

$$\log(JND_r) = \beta_0 + \beta_1 r + U_k$$

$$\log(JND_{vf}) = \beta_0 + \beta_1 v + U_k$$

Log-transformation for Individual Observations

[KAY 2015]

Method	Coefficients								R2	AIC	Normality of residuals	Skewness	Kurtosis	Homoscedasticity
	$\beta 0$	p	$\beta 1$	p	$\beta 2$	p	$\beta 3$	p						
f1: log(JNDr) ~ r	-1.4137	< .001	-2.0152	< .001	0.1365	.0021	-0.0815	.2837	.7941	-1724.3940	p = .0586	0.1170	0.5619	p = .9961
f2: log(JNDv) ~ v	0.1903	< .001	0.0297	< .001	-0.1074	.0516	0.0059	.0130	.7104	748.3163	p < .001	0.4009	0.8595	p = .2798

Both the visual feature and r can be fit into a log-transformed model.

So Far and Next

- The log transformation improves the models of correlation and visual features based on individual observations.
- Log transformations are used in HCI (e.g., completion time).
- A model supported by perceptual psychology?

STEVENS' POWER LAW

[STEVENS 1957]

A widely used perceptual law.
Replace Weber's law, though debates remain.

$$P(I) = \alpha \cdot I^a \longrightarrow JND^b = \beta \cdot I$$

The perceived stimulus is a power function of the objective stimulus.

[SEE OUR PAPER FOR THE DETAILED CONNECTION BETWEEN JNDS AND STEVENS' POWER LAW]

Power-transformation for Individual Observations

$$JND_{r}^{w1} = \beta_0 + \beta_1 r + U_k$$

$$JND_{vf}^{w2} = \beta_0 + \beta_1 v + U_k$$

Power-transformation vs. Log-transformation

[LOG]

Method	Coefficients								R2	AIC	Normality of residuals	Skewness	Kurtosis	Homosce-dasticity
	β0	p	β1	p	β2	p	β3	p						
f1: log(JNDr) ~ r	-1.4137	< .001	-2.0152	< .001	0.1365	.0021	-0.0815	.2837	.7941	-1724.3940	p = .0586	0.1170	0.5619	p = .9961
f2: log(JNDv) ~ v	0.1903	< .001	0.0297	< .001	-0.1074	.0516	0.0059	.0130	.7104	748.3163	p < .001	0.4009	0.8595	p = .2798

[POWER]

Method	BCT	Coefficients								R2	AIC	Normality of residuals	Skewness	Kurtosis	Homosce-dasticity
	ω	β0	p	β1	p	β2	p	β3	p						
f1: JNDr ~ r	0.26	0.6629	< .001	-0.2646	< .001	0.0186	< .001	-0.0155	.0615	.8099	-1742.0840	p = .2866	-0.0298	-0.2223	p = .3326
f2: JNDv ~ v	-0.23	0.9465	< .001	-0.0037	< .001	0.0187	< 001	-0.0010	< .001	.7387	703.7296	p = .0702	0.1061	-0.2850	p = .0607

All the metrics were improved if no worse than.

LINEAR MEAN MODEL (WEBER'S LAW)

→ simple and enough for comparing aggregated results
→ based on perceptual psychology

LOG-TRANSFORMATION MODEL

→ better than linear for individual observations
→ a quick correction for skewed residuals

[KAY 2015]

POWER-TRANSFORMATION MODEL

→ more precise for individual observations
→ a link to perceptual psychology
→ linear and log are a special case of power

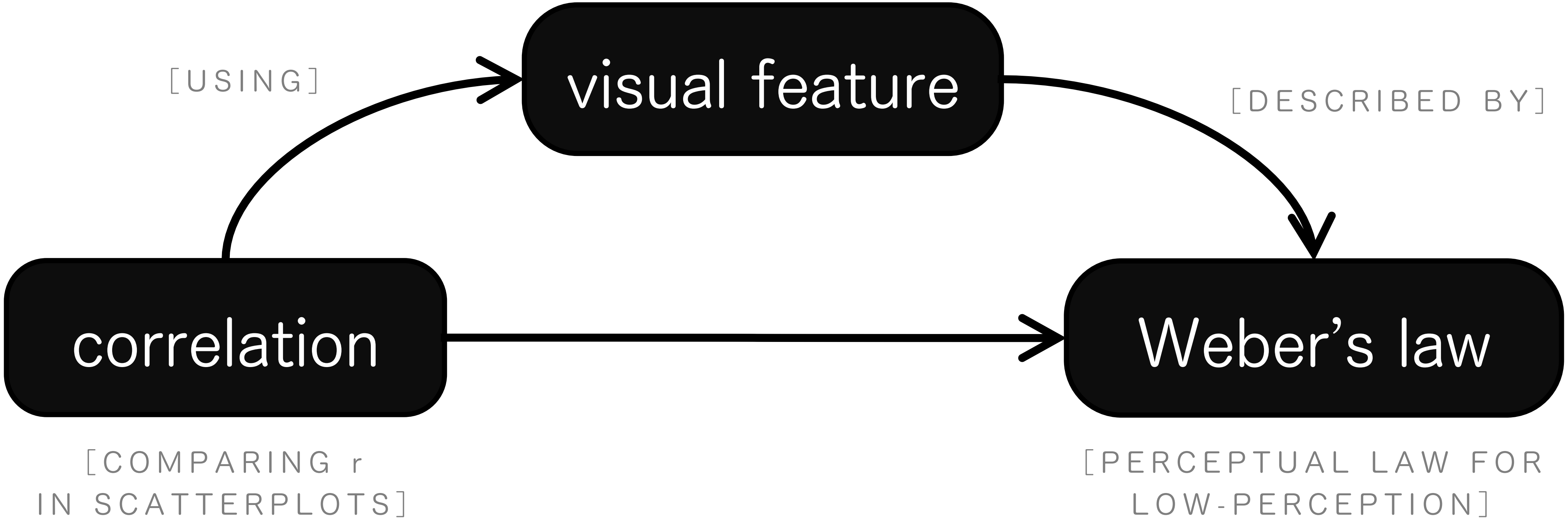
Which Model?

04

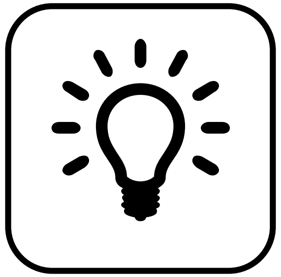
PART

Implications

WHY WEBER'S LAW



BRIDGE TWO SIDES OF RESEARCH



Perceptual Science

[FINDINGS AND THEORIES]

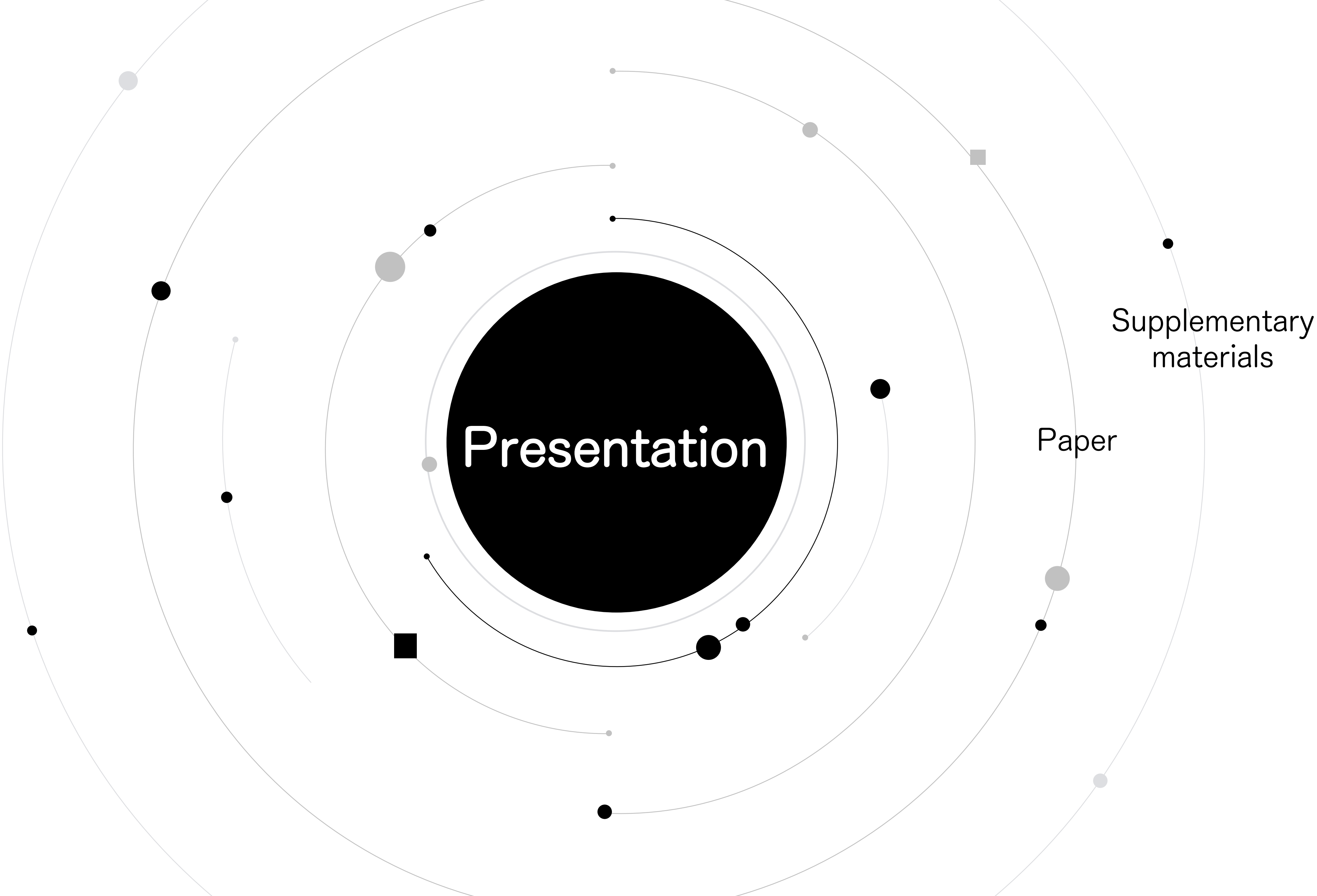
Visualization Community

[LARGE-SCALE APPROACHES FOR MODELING PERCEPTION]

CLARIFICATIONS

- **CAUSALITY?** → Evidence other than causality
- **FUTURE WORK** → Reasoning causality

- **ASSUMPTION** → People **use a single** and **the same** visual feature.
- **FUTURE WORK** → Multiple and combined visual features.


Presentation
Paper
Supplementary
materials

Thanks!

Fumeng Yang *(seek an intern position!)*
fy@brown.edu

Lane Harrison *(seek PhD students!)*
lane@cs.wpi.edu

Ronald Rensink
rensink@cs.ubc.ca

Steven Franconeri
franconeri@northwestern.edu

Remco Chang
remco@cs.tufts.edu

VIS 2018

TAKE-AWAYS

- People use visual features to make a visual judgment when comparing correlation in scatterplots.

- Power-transformation appears to better describe correlation perception and it is supported by perceptual psychology.

Correlation Judgment and Visualization Features
https://github.com/Fumeng-Yang/VisualFeature_TVCG