Comparative Visualization: Interactive Designs and Algorithms Depending on Data and Tasks

Tatiana von Landesberger[1], Kathrin Ballweg[1],
Hans-Jörg Schulz[2], Natalie Kerracher[3], Margit Pohl[4]

VIS Tutorial 2018

1. TU Darmstadt, Darmstadt, Germany
2. Aarhus University, Denmark
3. Edinburgh Napier University, UK
4. TU Wien, Austria

PART I:
SIMILARITY MODELS IN COGNITIVE PSYCHOLOGY

Speaker: Margit Pohl

Perception and Cognition

Cognitive Psychology:

Research on perceived similarity

- Development of categories

Attention

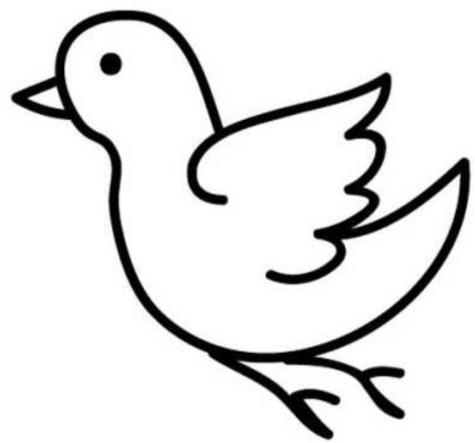
- Change blindness

Categories 1

Feature based approach

Bird

- Flies
- Has wings
- Has feathers
- Has a beak
- Lays eggs



Categories 2

Feature based approach:

Mathematical models of similarity research (multidimensional scaling, Contrast model/Tversky)

Problem: People develop theories about objects in their environment (especially about the relationships between features)

Categories 3

What makes things similar (= belong to the same category) ?

Example: Lemon (yellow, sour, round shape)

> You paint the lemon, inject sugar water, step on it – it is still seen as a lemon

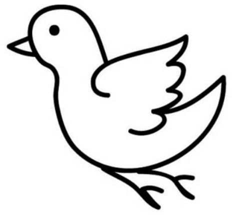
The essence of a lemon?

(grew on a lemon tree,

has lemon DNA)

Levels of Comparison

One-to-One: Prototypes

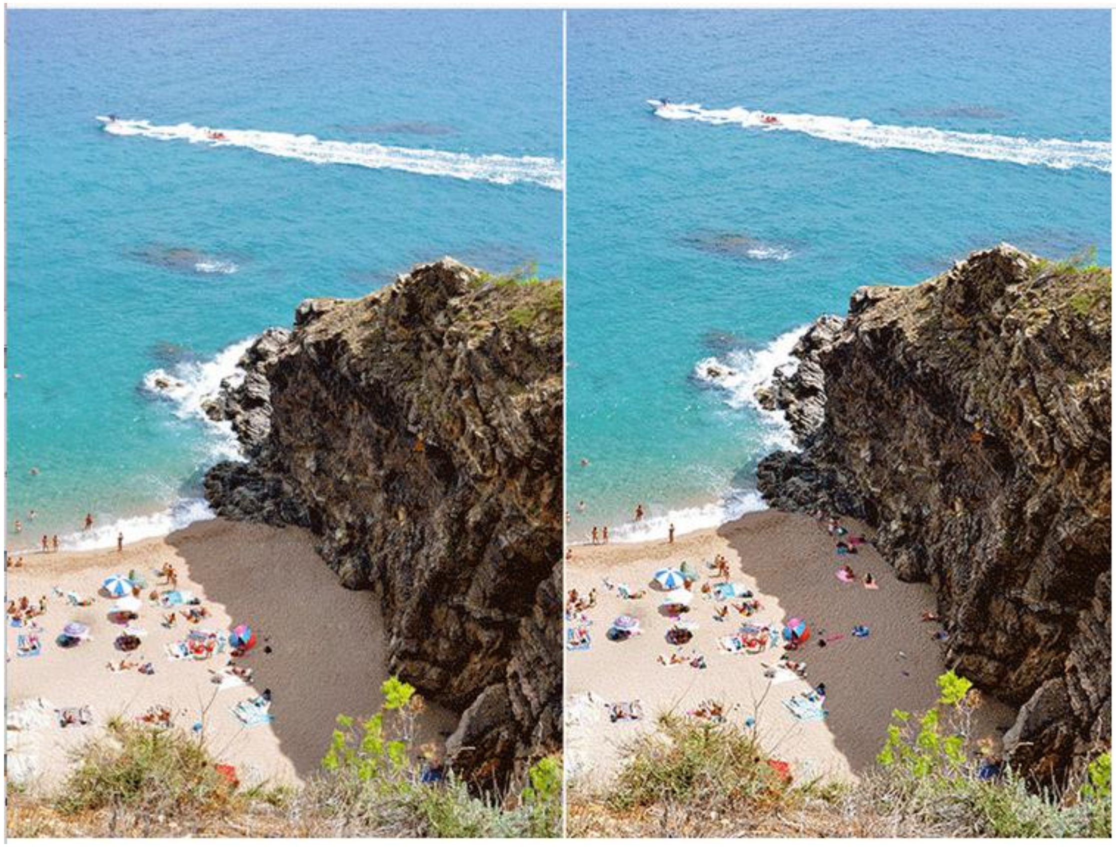


One-to-Many: Exemplars

Many-to-Many:

Attention 1

Change blindness



Attention 2

Change blindness

Guidelines (Rensink 2002)

1. Items on the screen should be easy to identify
2. Visual events on the screen should be minimized
3. Important elements should be emphasized

Holistic vs. Analytic 1

Is the percepetion of similarity holistic/global or analytic/componential?

Structuralist approach: analytic

Gestalt psychology: holistic

Perception of faces probably holistic

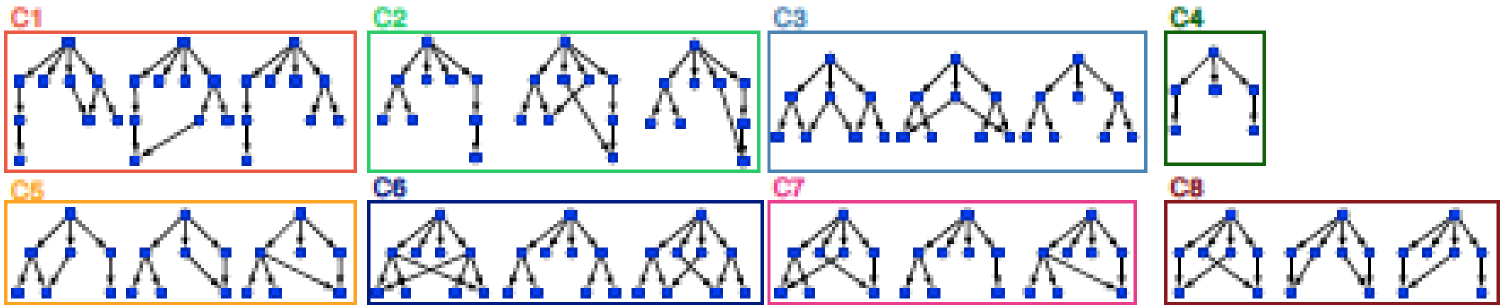
Perception of objects probably analytic

Dual route to recognition

Holistic vs. Analytic 2

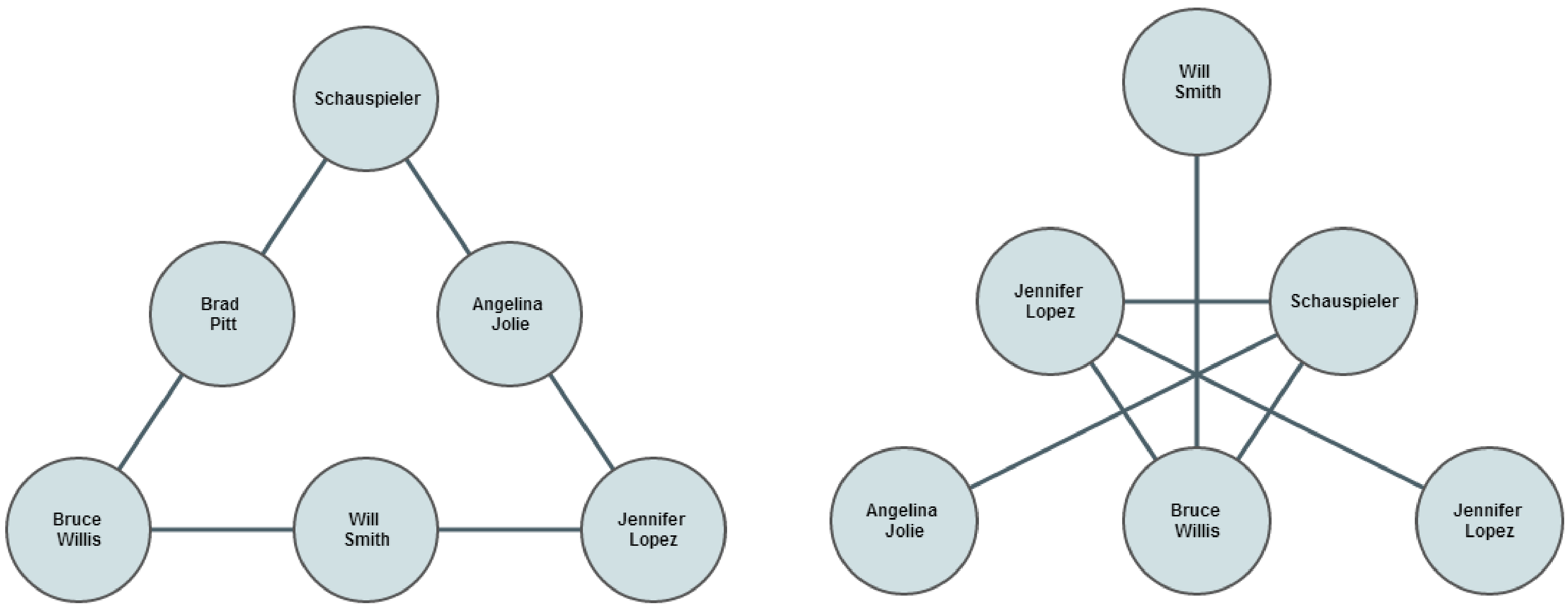
Is the percepetion of similarity holistic or analytic?

This is probably task dependent.



Semantic meaning

Semantic meaning may overrule design of a visualization



Influence of type of representation

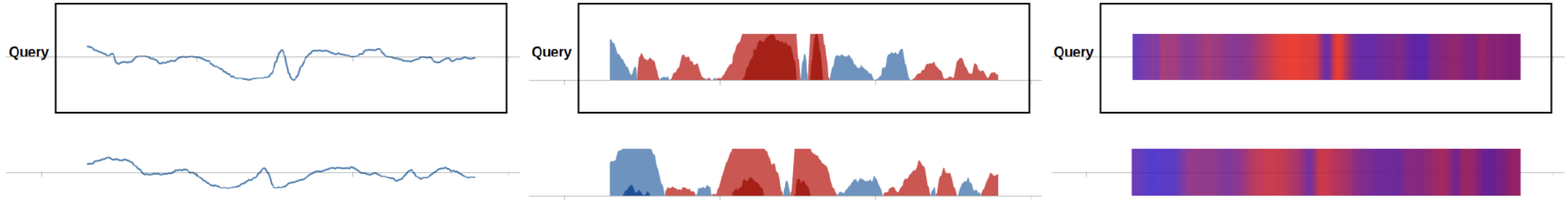
Gogolou et al (2018)

Line chart – Horizon Graph – Colorfield

Horizon Graph better for time warping

Colorfields not appropriate for time warping

- The type of representation influences what kind of similarities are perceived.



PART II:
COGNITIVE USER STUDIES FOR VISUAL COMPARISON

Speaker: Kathrin Ballweg

OVERVIEW

Visual Comparison Studies & Comparison Type

1-to-1 Comparison	1-to-Many Comparison	Many-to-Many Comparison
○ von Landesberger et al. „Investigating Graph Similarity Perception: A Preliminary Study and Methodological Challenges “ ○ Bernstein et al.. „How similar is it? towards personalized similarity measures in ontologies. ○ …	○ Klippel et al. "Color Enhanced Star Plot Glyphs – Can Salient Shape Characteristics be Overcome?" ○ Fuchs et al. „The Influence of Contour on Similarity Perception of Star Glyphs“ ○ …	○ Pandey et al. "Towards understanding human similarity perception in the analysis of large sets of scatter plots." ○ Ballweg et al. „Visual Similarity Perception of Directed Acyclic Graphs: A Study on Influencing Factors and Similarity Judgment Strategies“ ○ …

=> All comparison types are object of perception and cognition research

Commonalities of Cognitive Studies for Visual Comparison

- Feature-driven similarity model
- Explanation of “What is considered to be similar?“ usally not given to the participants
  - Reason:
    - Studies are interested in the humans‘ mental model of similarity for e.g., a specific data type
    - However, the answer to the above question would be the mathematical model
- Mostly, the studies ask for similarity and rather not for commonalities, differences or dissimilarity

$$Similarity = \Sigma\{Commonalities\} + \Sigma\{Differences\}$$

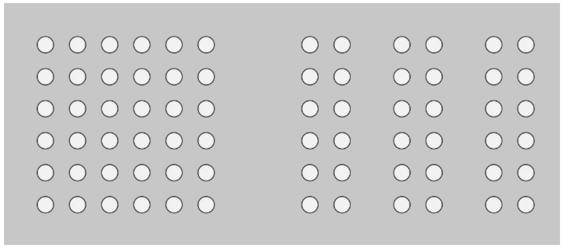
$$Dissimilarity \approx -1 * \underbrace{(\Sigma\{Commonalities\} + \Sigma\{Differences\})}_{\text{Similarity Function}}$$

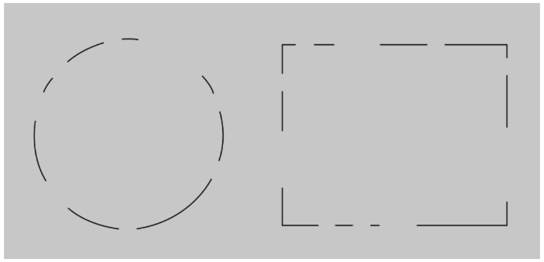
Tversky, A. (1977). Features of similarity. Psychological review, 84(4), 327.

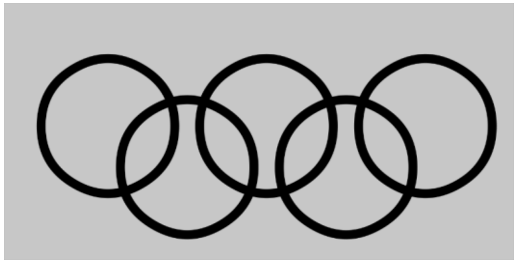
Factors Gernerally Influencing Human Similarity Perception

- Shape
  - Star Glyphs
    - Fuchs et al.., The Influence of Contour on Similarity Perception of Star Glyphs
  - Directed acyclic graphs (DAGs)
    - Ballweg et al. Visual Similarity Perception of Directed Acyclic Graphs: A Study on Influencing Factors and Similarity Judgment Strategies
  - Scatterplots
    - Pandey, Anshul Vikram, et al. Towards understanding human similarity perception in the analysis of large sets of scatter plots.

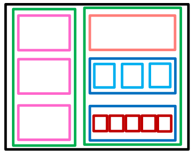
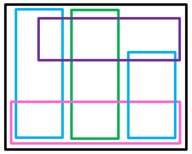
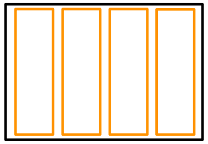
Reason: Gestalt Laws https://www.verywellmind.com/gestalt-laws-of-perceptual-organization-2795835







Human Strategies for the Task of Visual Comparison I

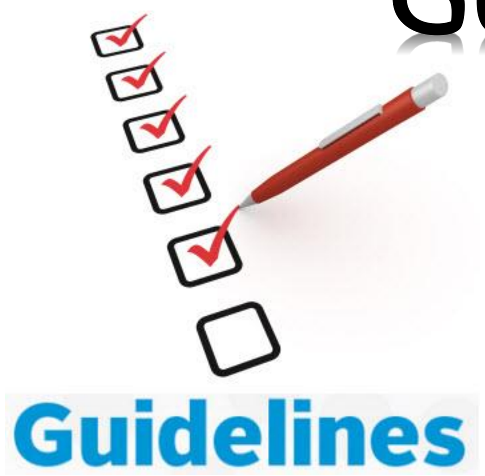
- **Many-to-Many Comparison**
  - Divide and conquer 
  - Sequential factor consideration with always the entire dataset 
  - Consideration of a single factor 

Ballweg et al. Visual Similarity Perception of Directed Acyclic Graphs: A Study on Influencing Factors and Similarity Judgment Strategies

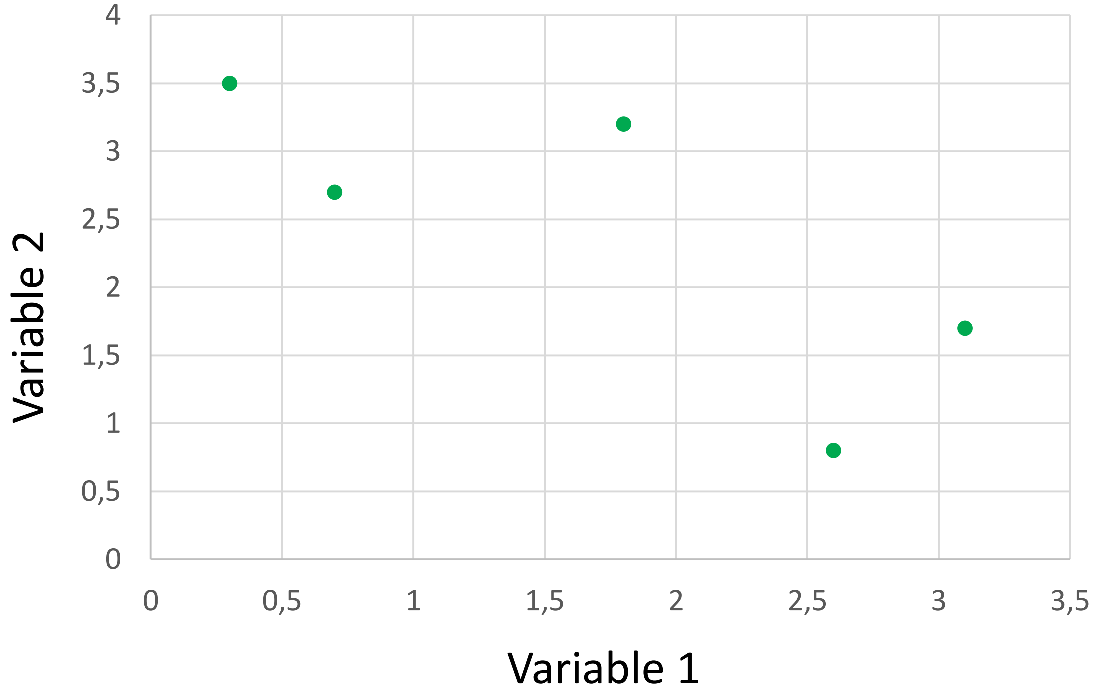
Human Strategies for the Task of Visual Comparison II

- **1-to-1 Comparison**
  - Tipp: consult psychology research
- **1-to-Many Comparison**
  - Tipp: consult psychology research

GUIDELINES



http://www.cctca.com/uploads/gallery/Guidelines2.jpg



SCATTERPLOTS

Many-to-Many Comparison of Scatterplots

- Scagnostics not perception-aware
- 6 concepts used by humans to describe the similarity of scatterplots:
  1) density:
  Density refers to the concentration of data points in certain region of the plot and can vary from high-density to low-density. Regardless of how the shapes of the plots vary, as long as there exists a high density pattern, the plots are often grouped together

Pandey et al. Towards Understanding Human Similarity Perception in the Analysis of Large Sets of Scatter Plots

Many-to-Many Comparison of Scatterplots

- 6 concepts used by humans to describe the similarity of scatterplots (continued):
  4) regularity:
  „Regularity refers to the consistency with which certain concepts, like shape or density, appear throughout the plot. In other words, regularity can refer to the repetition of certain patterns in a plot."

Pandey et al. Towards Understanding Human Similarity Perception in the Analysis of Large Sets of Scatter Plots

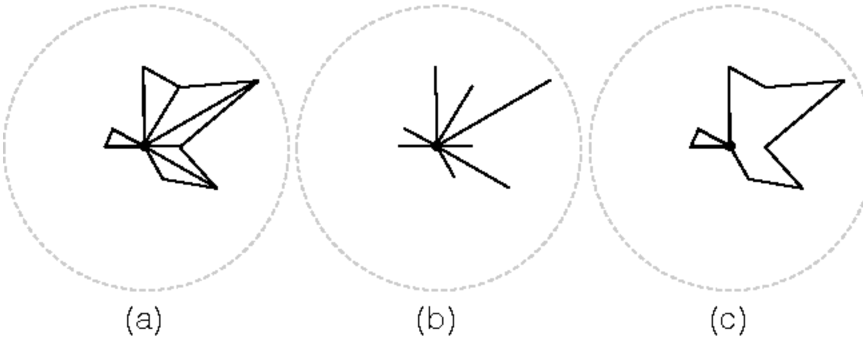


Figure 2. *The three considered star glyph variations [42]: (a) The common star glyph uses data lines radiating from the center; maximum values of the*

Jäckle et al. Star Glyph Insets for Overview Preservation of Multivariate Data

STAR GLYPHS

1-to-Many Comparison of Star Glyphs

Design Choices:

- Contours make the participants perceive the visualizations as shapes enforcing a comparison of geometrical shapes rather than data properties/values the visualizations show
  - --> even experts
  - --> novices fall back to shape comparison due to their lack of experience
  - Reason: factors enforcing perceptual unity of shape – e.g., contours – lead humans to naturally make shape judgments of similarity rather than data
- Starglyphs **without** contours promote data similarity comparison rather than shape

Fuchs et al.., The Influence of Contour on Similarity Perception of Star Glyphs

1-to-Many Comparison of Star Glyphs

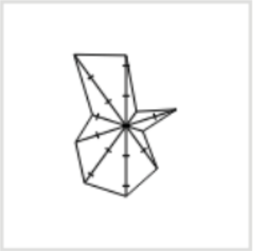
- The issue with perceptual unity enforcing factors generalizes also to filling

Fuchs et al.., The Influence of Contour on Similarity Perception of Star Glyphs

1-to-Many Comparison of Star Glyphs

- Adding reference structures to the star glyph did not have the effect on accuracy of the data similarity judgment task the authors were expecting
  - Statistical trend shows that overall reference in the background is to preferer as compared to tick marks

Overall reference



Tick marks

- Reference structures are not able to mitigate the influence of shape

Fuchs et al.., The Influence of Contour on Similarity Perception of Star Glyphs

Star Glyph Comparison – Guidelines as a Summary

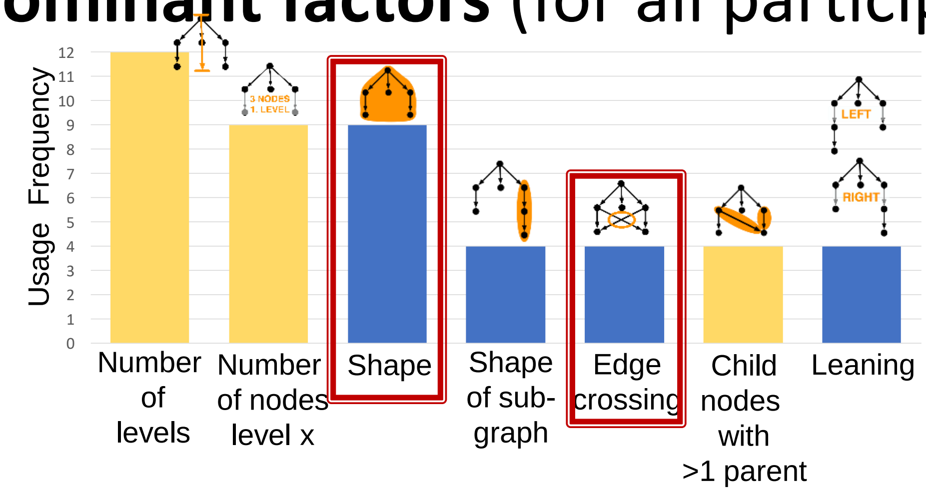
When judging data similarity avoid contours in glyph designs.
For low number of dimensions (around 4) any glyph variation can safely be used for data similarity judgments.
When there is a need for contours, add data lines to the design to strengthen data similarity judgments.
When there is a need for contours, the designer can decide whether or not to use fill color.
When clutter is an issue avoid reference structures in non-contour star glyphs for similarity search tasks.
If references are required use grids rather than tickmarks.

Fuchs et al.., The Influence of Contour on Similarity Perception of Star Glyphs

GRAPHS

Many-to-Many Comparison of Small Graphs

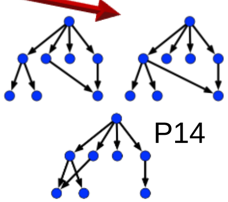
- **Perception influencing factors** (for all participants)
  - 27 distinct factors
    (15 visual, 10 graph theoretical)
    - **=> clear tendency of „visual factors over graph theoretical factors"**
- **Dominant factors** (for all participants)



Edges related factors were nearly irrelevant

Further edge related factors:
- Number of edges to level ID+1
- Long edges
- Visual symmetry - edges

Used by <20% of participants



Ballweg et al. Visual Similarity Perception of Directed Acyclic Graphs: A Study on Influencing Factors

Many-to-Many Comparison of Small Graphs

Value Ranges of the Used Factors

Value Range Types:

- **boolean** – (Non-)Existence of a factor *e.g.: "There are nodes which only have one predecessor"*
- **1 value** – focus on a specific value *e.g. "number of layers = 4"*
- **> 1 value** – > 1 value attributed to the factor & these values are not Boolean *e.g. "number of nodes on level 2 = 3, 4"*

In total usage frequency

- boolean: 36
- 1 value: 4
- >1 value: 38

27 distinct factors used by the participants	boolean	1 value	>1 value
Factors used by at least 20% of the participants			
number of levels	0	2	10
number of nodes on a specific level	0	1	8
shape	2	0	7
arm/branch (DAG sub-shape)	4	0	0
edge crossing	4	0	0
child node(s) with > 1 parent node	3	0	1
leaning	0	0	4
Factors used by less than 20% of the participants			
one parent node	3	0	0
visual symmetry (entire DAG)	3	0	0
number of nodes in the entire DAG	0	0	2
node position/layout	1	0	2
level style	0	0	2
graph type	2	0	0
DAG appears nearly full	2	0	0
leave nodes on higher level than the lowest level	2	0	0
number of edges to level $ID + 1$	0	1	1
visually approximated number of nodes in branch	1	0	0
one root-like node	1	0	0
cycle	1	0	0
long edges	1	0	0
outlier subgraph (self-defined)	1	0	0
visual symmetry (edges)	1	0	0
hierarchy violations (self-defined)	0	0	1
isomorphic	1	0	0
balance	1	0	0
"Other"	1	0	0
"Not classifiable"	1	0	0

Average Factor Usage per

- **Factor Specifics:**
  - **On average 4 factors** are **considered** for a similarity judgment
    - **2** out of 4 are **visual factors**
    - **2** out of 4 are **graph theoretical (gt) factors**
      - **=> no visible tendency whether a person would rather use visual or gt factors**
        = Divergence from the results per distinct factor (15 visual, 10 gt)
  - Dominance ranking of the factors relative to each other was not identifiable (≙ Bertin's ranking of visual variables
    - **Expected contribution: to be used factors and their weighting for future perception aware similarity measures**

Ballweg et al. Visual Similarity Perception of Directed Acyclic Graphs: A Study on Influencing Factors and Similarity Judgment Strategies

1-to-1 Comparison of Small Graphs

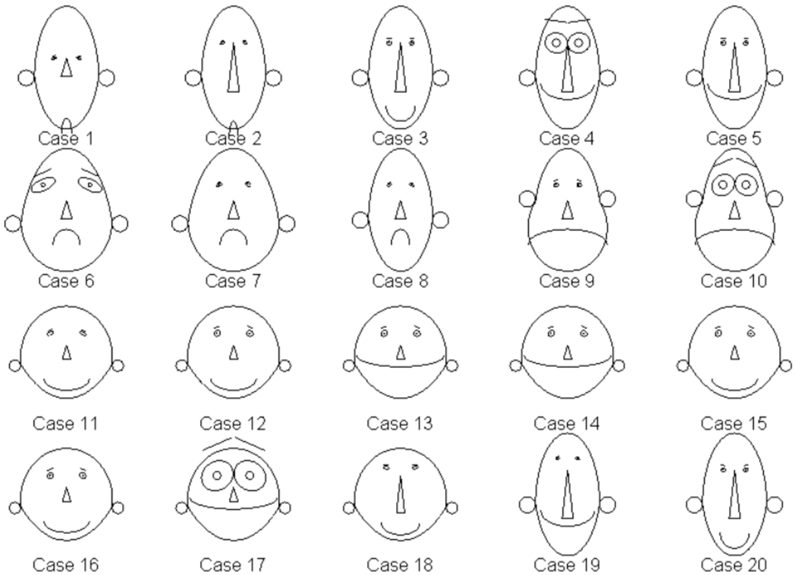
- Perception influencing factors
  - Similar to those of the many-to-many comparison case
    - **GREAT – means that the comparison type does not impact the influencing factors**
- Labels – especially categorical labels have a huge impact on the comparison result

T. Von Landesberger et al. Investigating Graph Similarity Perception: A Preliminary Study and Methodological Challenges

Data Item	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	…
Item1	7	5	67	2	0.6	56	…
Item 2	3	15	86	1	0.3	67	…
…	…	…	…	…	…	…	…

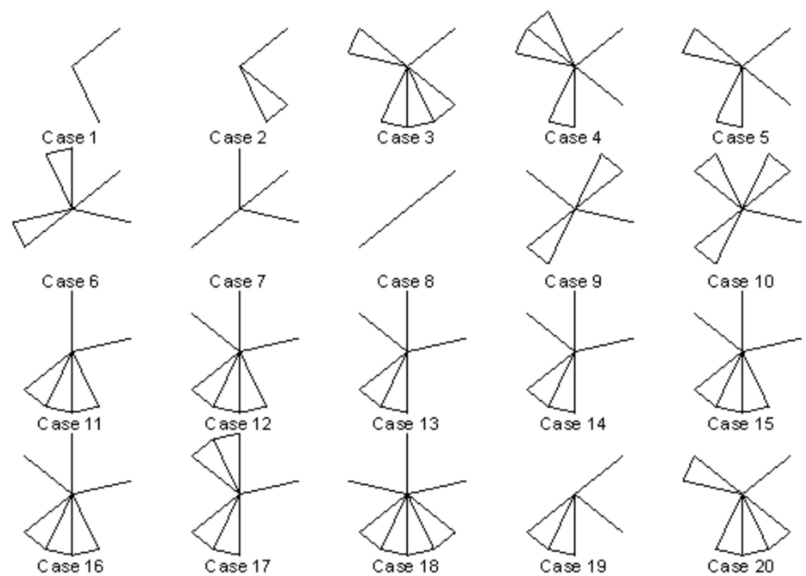
SPECIAL CASE MULTIVARIATE DATA – "WHICH VISUALIZATION TYPE SHOULD BE USED?"

Many-to-Many Comparison of Multivariate Data



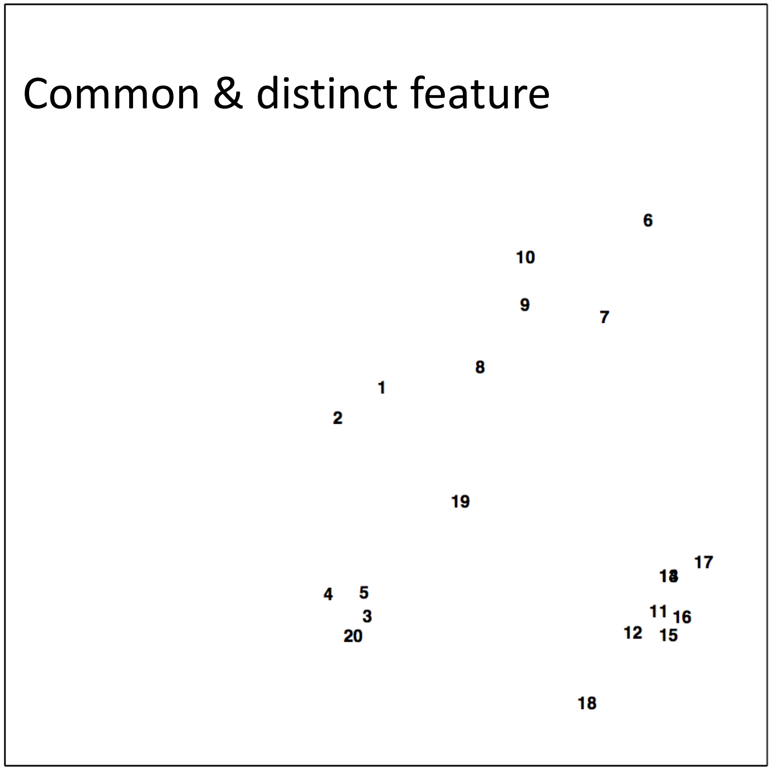
Chernoff faces

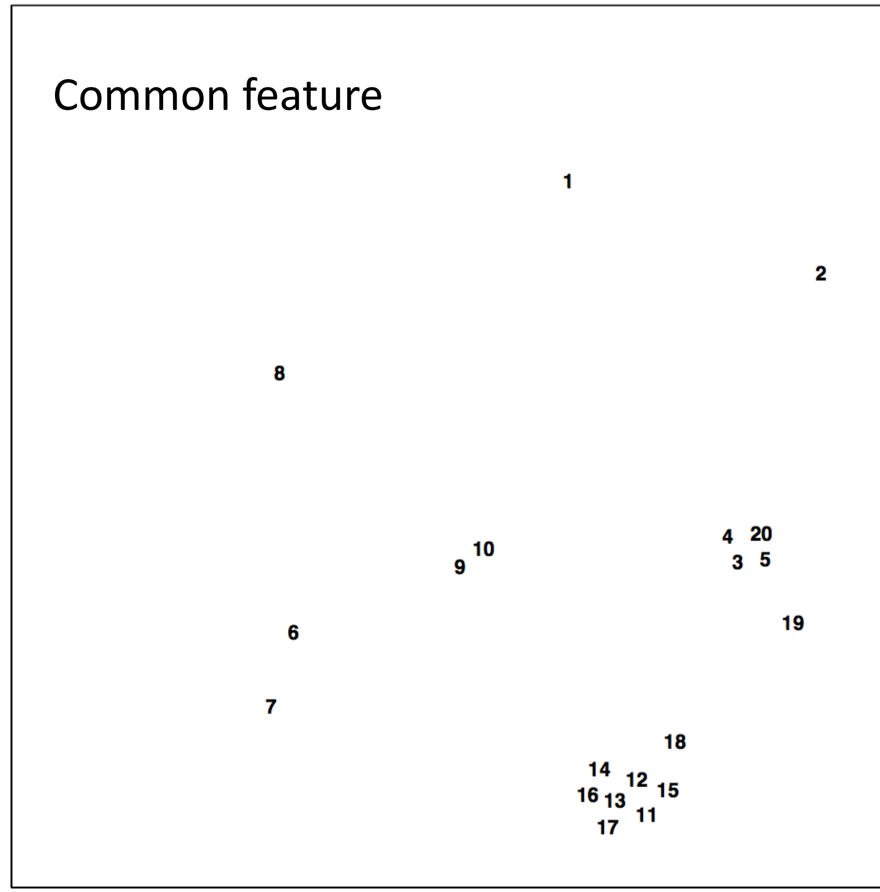
Star Glyphs



Lee et al., An Empirical Evaluation of Chernoff Faces, Star Glyphs, and Spatial Visualizations for Binary Data

Many-to-Many Comparison of Multivariate Data





Spatial visualizations

Lee et al., An Empirical Evaluation of Chernoff Faces, Star Glyphs, and Spatial Visualizations for Binary Data

Many-to-Many Comparison of Multivariate Data

Which visualization to choose?

common feature special visualization, participants were more accurate and confident* in their answers

*confidence with respect to the visualization type is dependent on the task

Lee et al., An Empirical Evaluation of Chernoff Faces, Star Glyphs, and Spatial Visualizations for Binary Data