Supporting information seeking behaviour with information visualization

Dr. Orland Hoeber
Department of Computer Science
University of Regina
Regina, SK, Canada

@ohoeber
orland.hoeber@uregina.ca

University of Regina

Motivation

Complex Search Tasks are Hard

- Many search tasks are complex in nature
  - ambiguous information needs
  - inability to articulate what is being sought
  - complex information structure

Complex Search Tasks are Hard

- Conducting a complex search places substantial cognitive burden on the searcher
  - craft and refine queries
  - evaluate and explore among the search results
  - sensemaking and synthesis of information
  - making use of what is found

Information Visualization Can Help

Information visualization methods can be leveraged to help searchers to perceive, interpret, and make sense of information during the various stages of the search process.

Information Visualization and Information Seeking Behaviour

- We need to understand the fundamental aspects of information visualization so that we can use them effectively in our interactive information retrieval interface designs

- We must also consider relevant information seeking behaviour and strategies as guidance for what kind of information might be useful to visualize at what time during the search process

Presentation Goals

understand the fundamental principles and theories of information visualization

appreciate how information visualization can enable and support various interactive information retrieval processes

example of a visual search interface that uses information visualization to enable information seeking

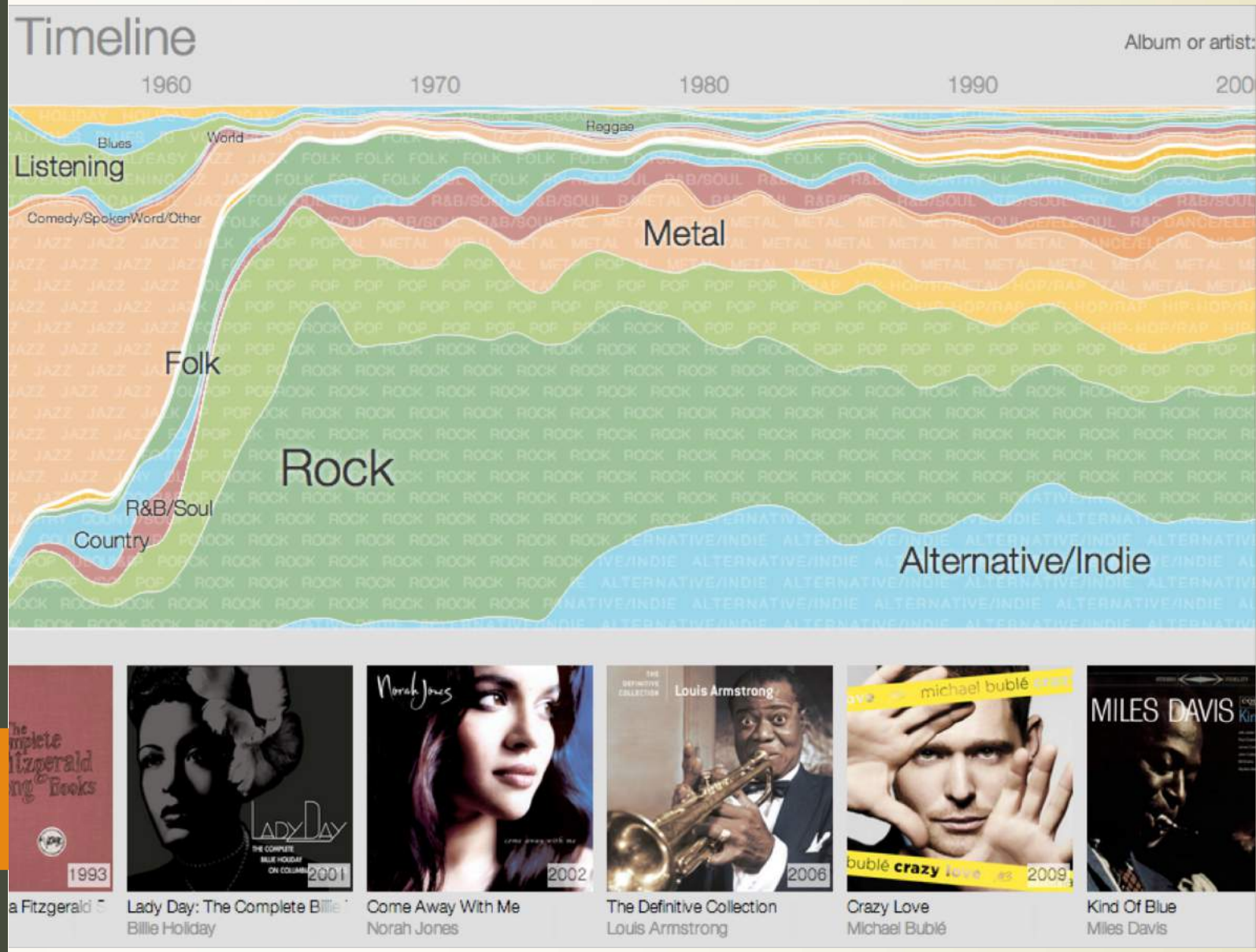
Overview of Information Visualization

What is Information Visualization?

- Information Visualization is the communication of abstract information using graphical representations

- Why visualize data?
  - A single picture can communicate a wealth of information
  - A picture can be processed more quickly than a comparable page of text or a table of numbers
  - Image interpretation is performed in parallel within the human mind, whereas text analysis is limited by the sequential process of reading
  - when done well, people can readily perceive, interpret, and make sense of the information being conveyed

Timeline of Music Genres

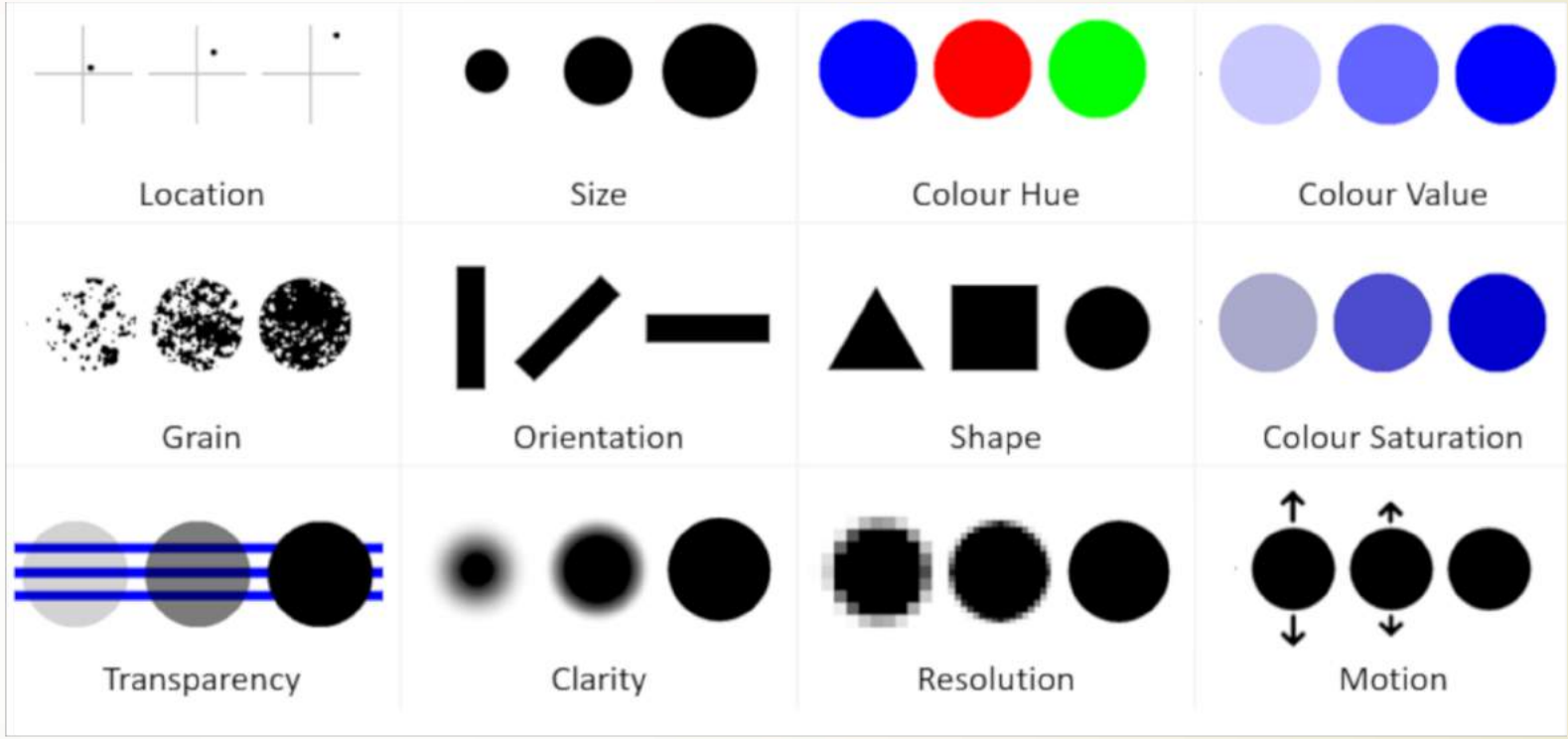


➡http://research.google.com/bigpicture/music/

Visual Variables

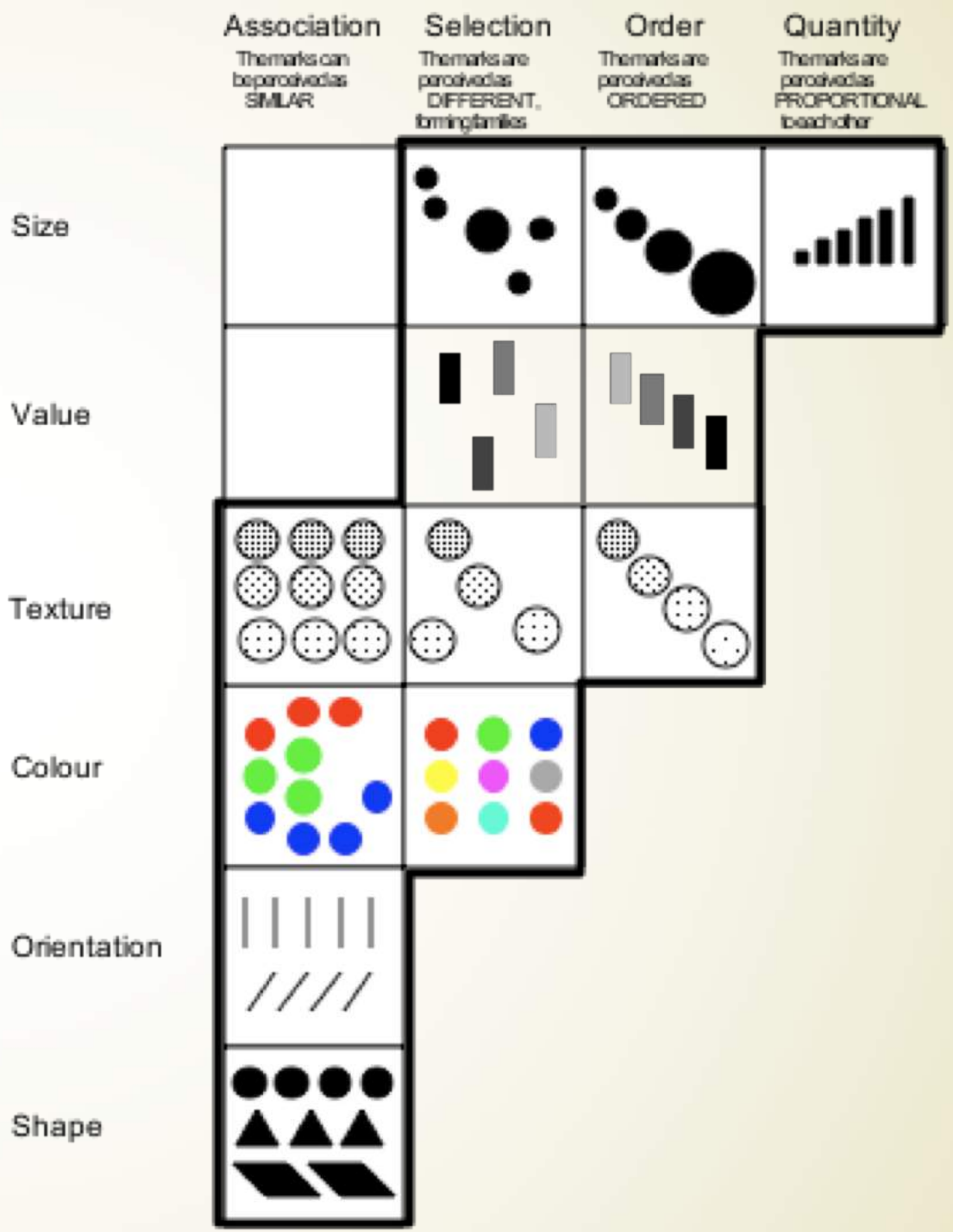
Visual Variables

- In order to create and arrange graphical symbols within a visualization, we must consider graphic primitives (visual variables) and their properties



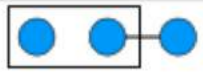
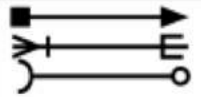
Visual Variables & Perceptual Tasks

- Bertin's Semiology of Graphics (1967) discusses the encoding of information based on the user's primary task:
  - association
    - can elements be perceived as similar
  - selection
    - can elements be perceived as different
  - order
    - can elements be perceived as ordered
  - quantity
    - can elements be perceived as proportional to one another



Visual Variables & Data Types

Properties and Best Uses of Visual Encodings

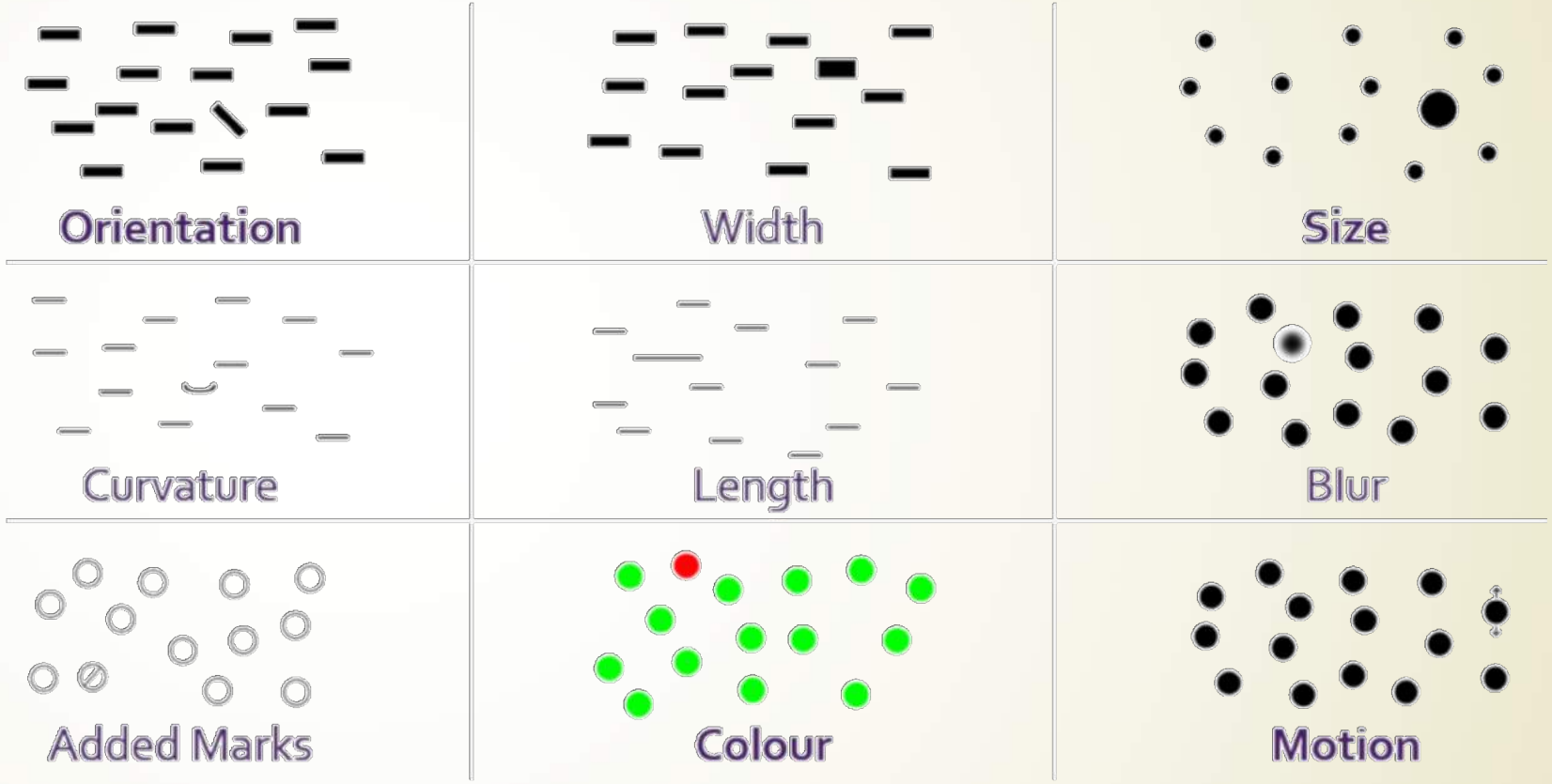
Example	Encoding	Ordered	Useful values	Quantitative	Ordinal	Categorical
	position, placement	**yes**	**infinite**	Good	Good	Good
1, 2, 3; A, B, C	text labels	**optional** (alphabetical or numbered)	**infinite**	Good	Good	Good
	length	**yes**	**many**	Good	Good	
	size, area	**yes**	**many**	Good	Good	
	angle	**yes**	medium/few	Good	Good	
	pattern density	**yes**	few	Good	Good	
	weight, boldness	**yes**	few		Good	
	saturation, brightness	**yes**	few		Good	
	color	no	few (< 20)			Good
	shape, icon	no	medium			Good
	pattern texture	no	medium			Good
	enclosure, connection	no	**infinite**			Good
	line pattern	no	few			
	line endings	no	few			
	line weight	**yes**	few		Good	

Noah Iliinsky • ComplexDiagrams.com/properties • 2012-06

http://complexdiagrams.com/properties

Pre-Attentive Processing

- Our brain is hard-wired to process certain visual stimuli
  - allows fast, parallel processing of what we see
  - enables us to quickly identify differences in certain types of visual features without cognitive effort

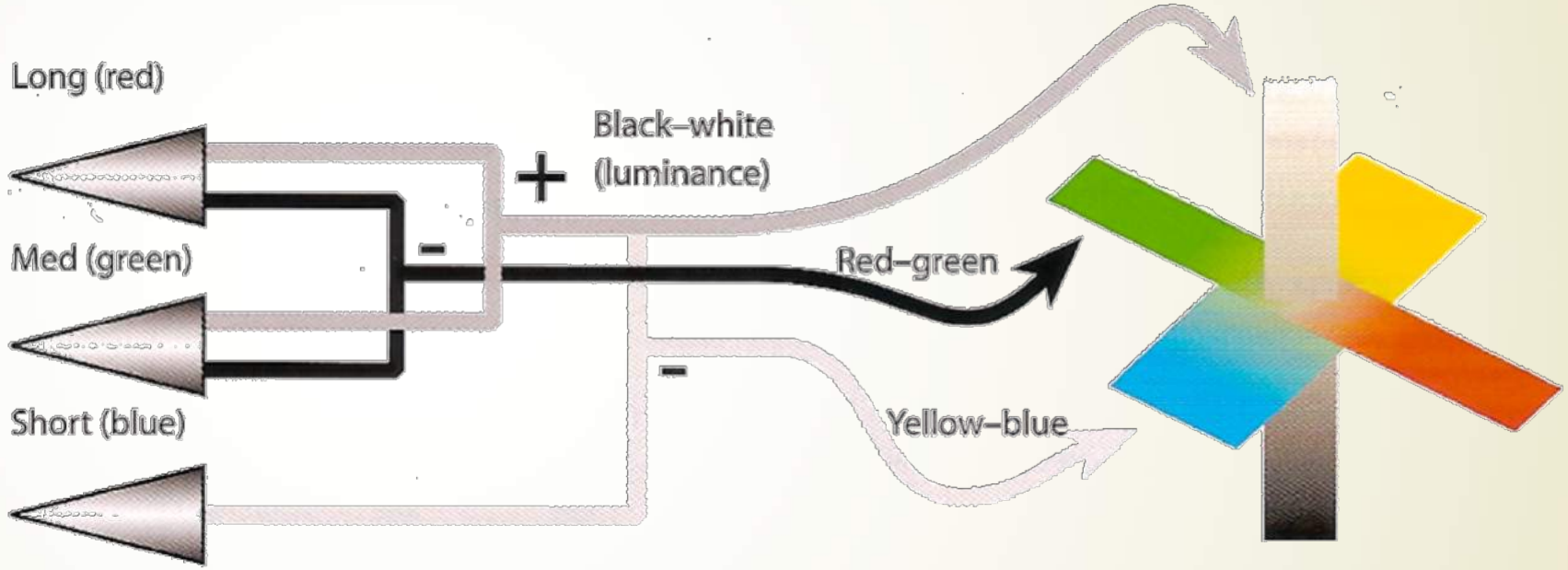


Colour Theory

- When using colour, there are some important properties that we often would like to have:
  - perceptual balance
    - a unit step anywhere in the colour scale produces a perceptually-uniform difference in colour
  - distinguishability
    - within a discrete set of colours, each is equally distinguishable from the others
  - opposites
    - for colour scales that represent data with a natural zero, negative colours are perceived as opposite from positive colours

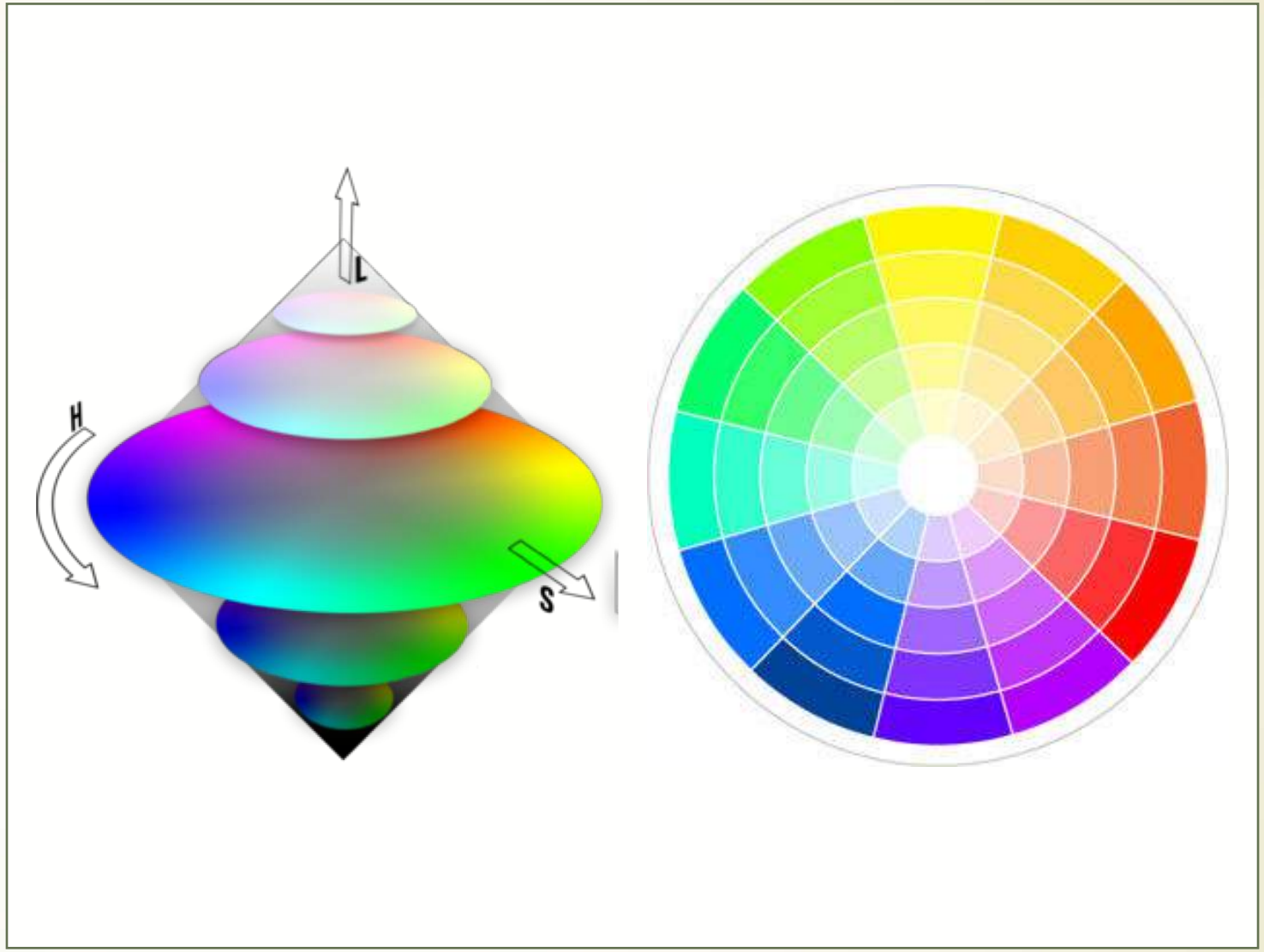
Opponent Process Theory of Colour

- Our brain combines the stimulus from our three cones (L, M, S) before processing them
- Results in three colour channels
  - red-green
  - yellow-blue
  - luminance (black-white)



Perceptually Organized Colour

- This colour theory has lead to the creation of colour scales that are perceptually organized (rather than being based on the wavelength of light)



Use of Colour for Encoding Data

- How we use colour to represent data depends on the features of the data
  - qualitative
    - discrete colours, as unique from one another as possible

  - ordinal
    - discrete colours, perceptually ordered

  - quantitative
    - continuous, perceptually ordered colour scale

  - data with a natural zero (ordinal, quantitative)
    - zero as a neutral colour (white, grey)
    - positive and negative values as independent perceptually ordered colour scales

http://colorbrewer2.org/

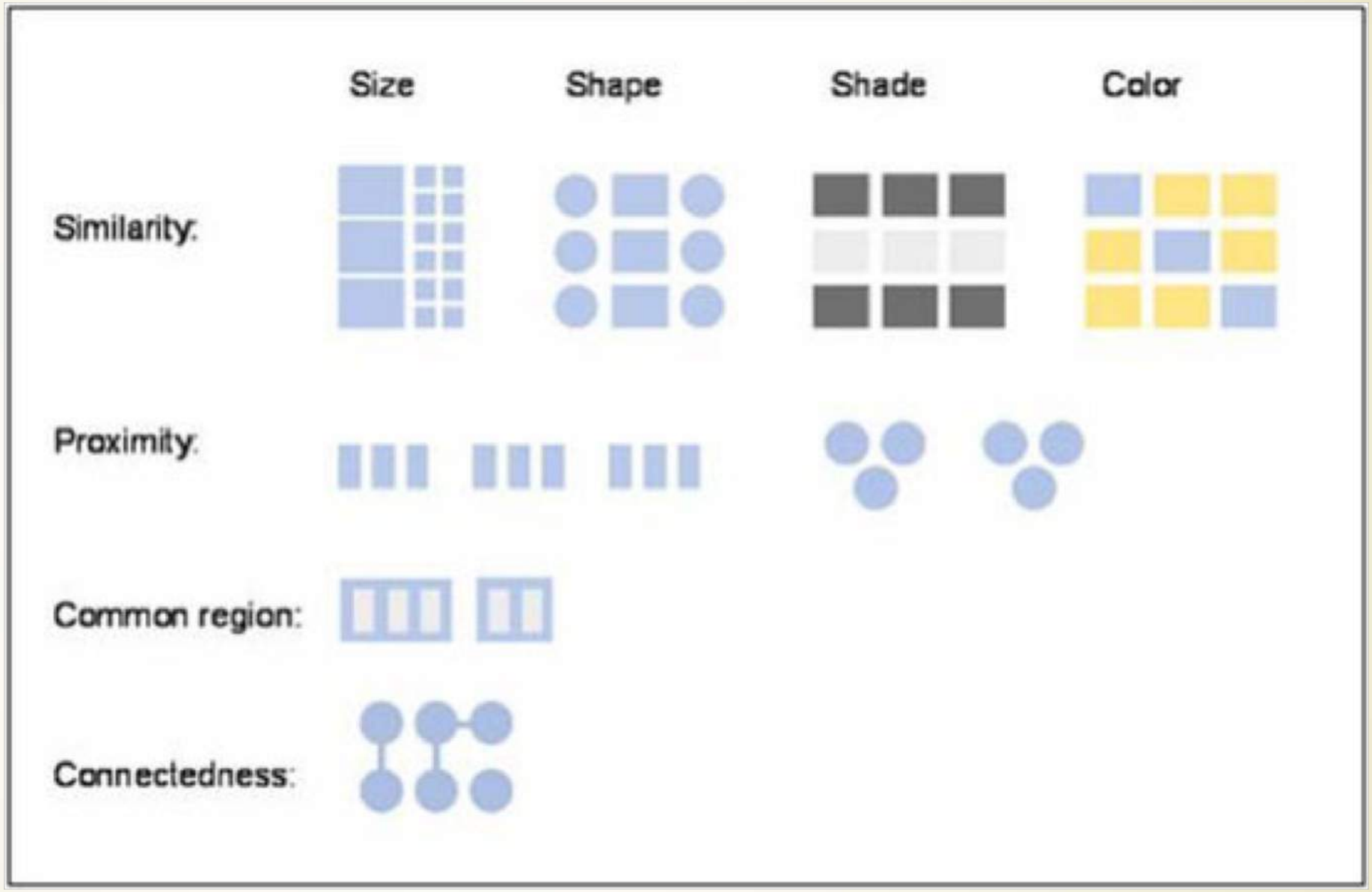
Visual Interpretation & The Gestalt Principles

- Theories from cognitive psychology can explain how we interpret what we see
- In 1912, Westheimer, Koffka, and Kohler founded the Gestalt school of psychology
  - attempt to understand pattern perception
  - outcome of their research: laws that describe how humans perceive patterns
- These Gestalt Laws easily translate into a set of design principles for information displays, which we call the Gestalt Principles

Gestalt Principles

- The Gestalt Principles that are of primary interest to information visualization work are:
  - similarity
    - visually similar objects will be viewed as being related
  - proximity
    - objects that are close together are perceptually grouped
  - closure/common region
    - objects within a closed/common region will be viewed as being part of a group
  - connectedness
    - connecting lines between objects will be interpreted as related

Gestalt Principles



- http://connect.humanfactors.com/profiles/blogs/eye-insights-evaluating-the

Summary of Information Visualization

- We have outlined some key elements of information visualization
  - what is information visualization?
  - visual variables
  - pre-attentive processing
  - colour theory
  - The Gestalt Principles

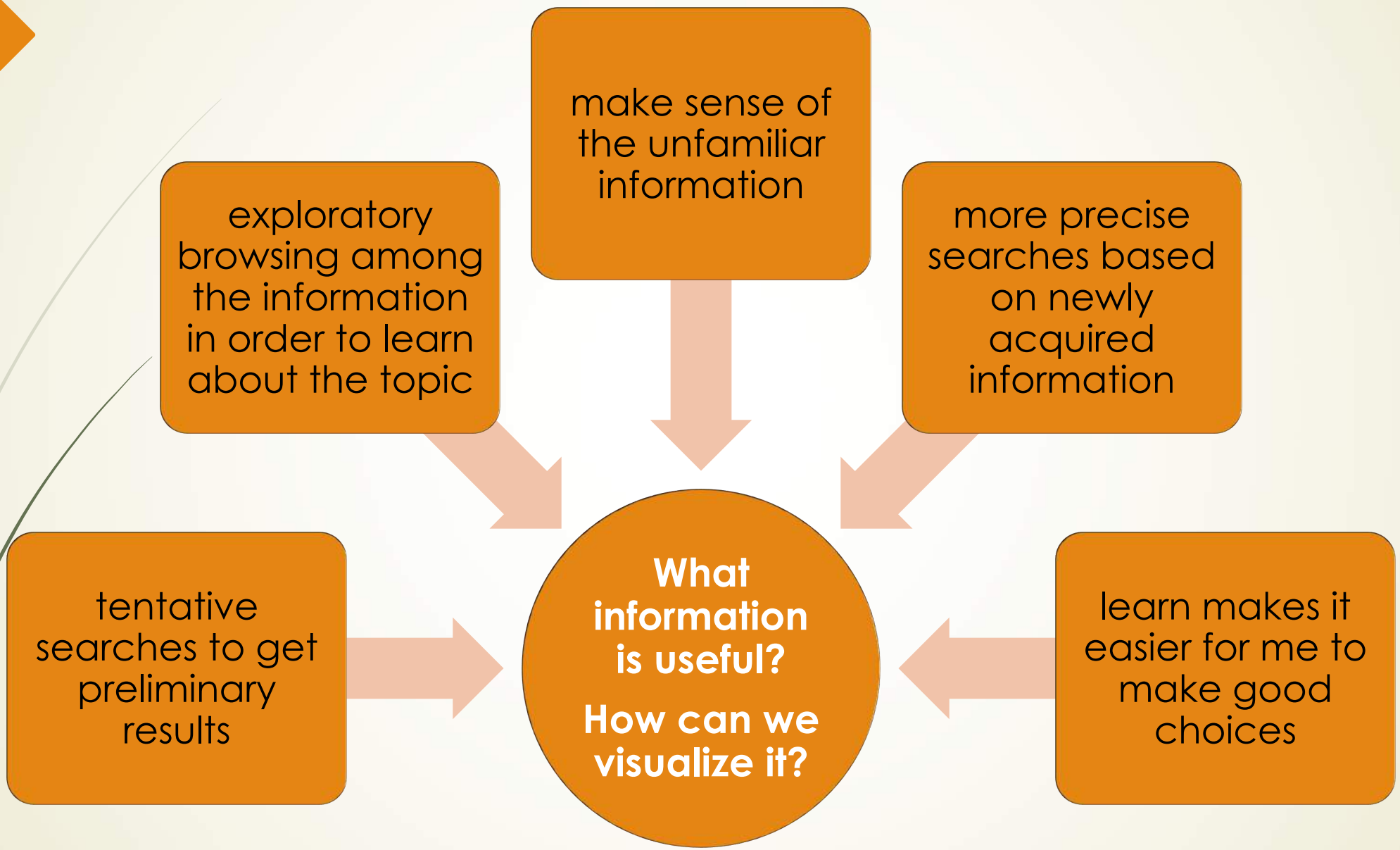
Search Behaviour & Strategies

Exploratory Search

Exploratory Search

- Exploratory search processes may be employed in various scenarios:
  - the information need is open-ended (White & Roth, 2009)
  - there are multiple facets to simultaneously consider (White & Roth, 2009)
  - the task is complex in nature (Awadallah et al., 2014)

- What makes exploratory search unique is the highly interactive nature of the search process
  - start with a vague or ill-defined search goal
  - develop knowledge & understanding through interactive querying
  - new knowledge supports the refinement of queries and the development of new queries


make sense of the unfamiliar information
exploratory browsing among the information in order to learn about the topic
more precise searches based on newly acquired information
tentative searches to get preliminary results
What information is useful?
How can we visualize it?
learn makes it easier for me to make good choices

Exploratory Search

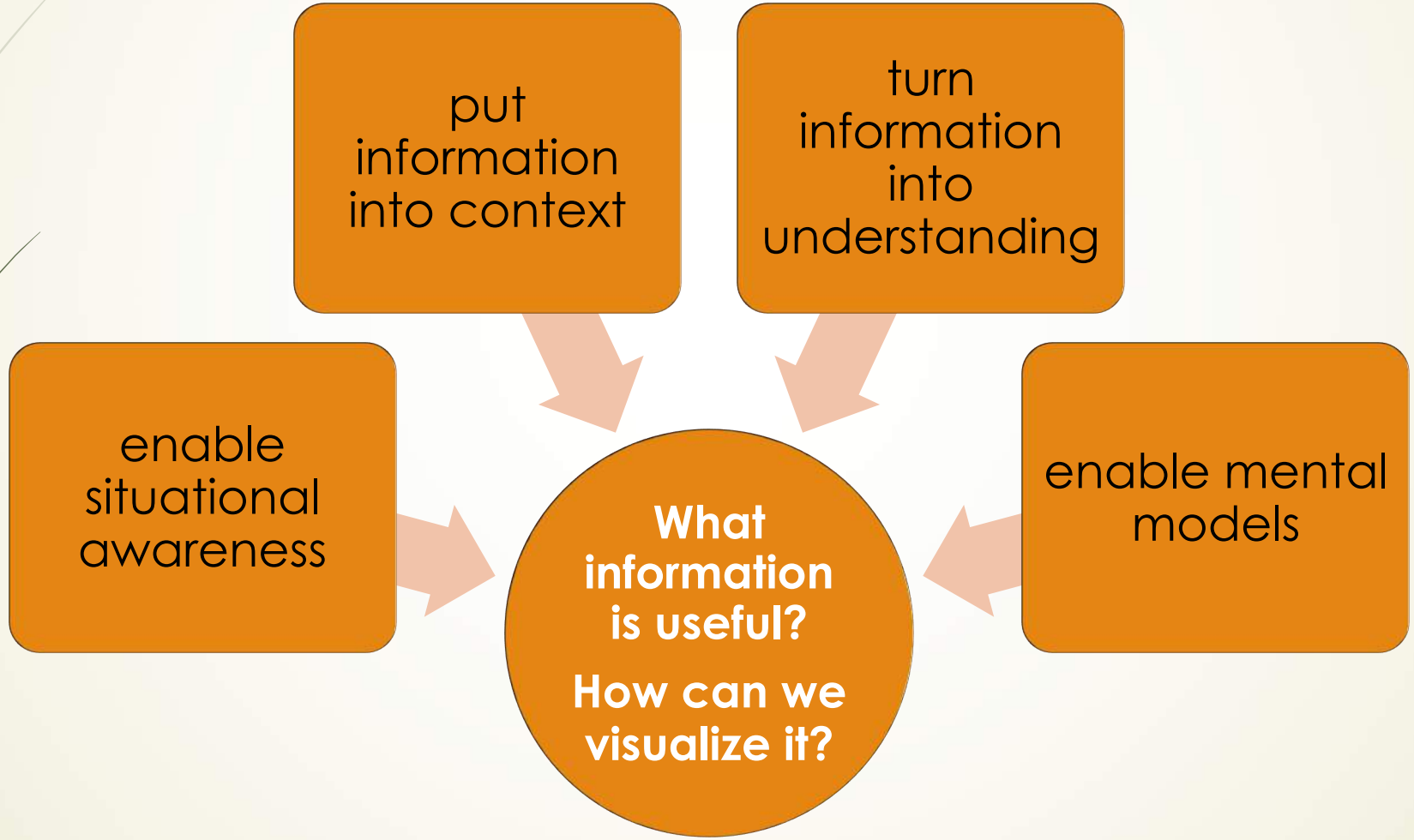
- Search interfaces that support exploratory search must encourage the interactive nature of search, much of which can be enabled with information visualization
  - interactive query refinement
  - interactive search results exploration
  - keep track of what has been seen
  - identify & show what is new
  - support knowledge development
  - enable the dynamic development of information needs
  - encourage inductive analysis

Sensemaking

Sensemaking

- Sensemaking is the cognitive process of developing an understanding of complex information in relation to the world around us (Pirolli & Russell, 2011)

- Two critical aspects of sensemaking are:
  - situational awareness
    - putting information into context (local and global)
  - turning information into an understanding of the situation
    - synthesis
    - integration into a mental model


put information into context
turn information into understanding
enable situational awareness
What information is useful?
How can we visualize it?
enable mental models

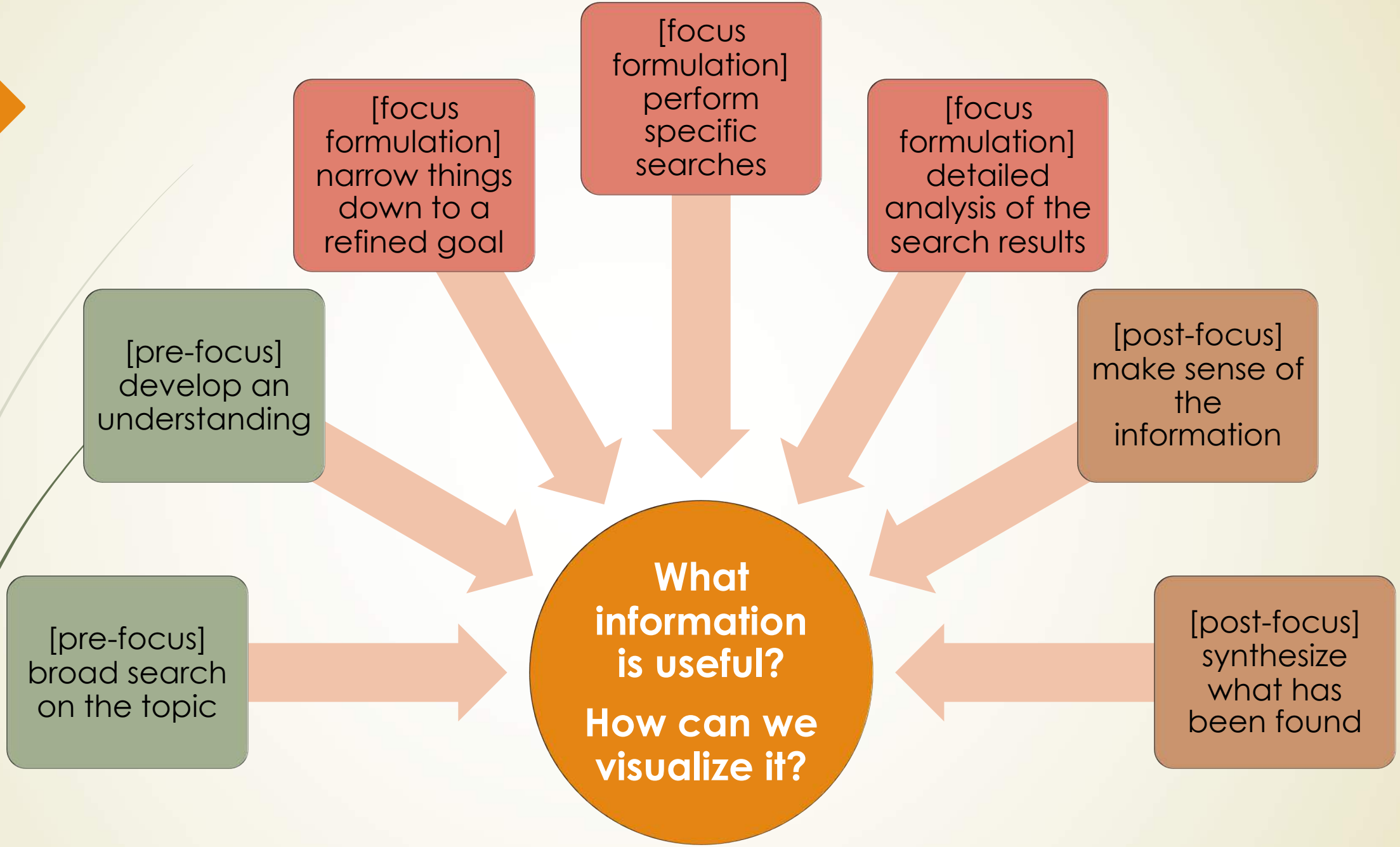
Sensemaking

- Pirolli & Russell (2011) suggested that sensemaking can be enabled by:
  - providing meaningful representations
    - information scent (Pirolli, 2007)
    - information visualization (Ware, 2008)
  - supporting rich interaction mechanisms
    - exploratory search (White & Roth, 2009)
    - taking an active role in the information seeking process (Hoeber, 2014)

Vakkari's Three-Stage Model of Information Seeking (Vakkari, 2003)

Vakkari's Model

Pre-focus	undertake a series of broad searches to assess the breadth of information in the collection
Focus formulation	refine the goal of the search activity and seek relevant information
Post-focus	collect, organize, and make use of the information

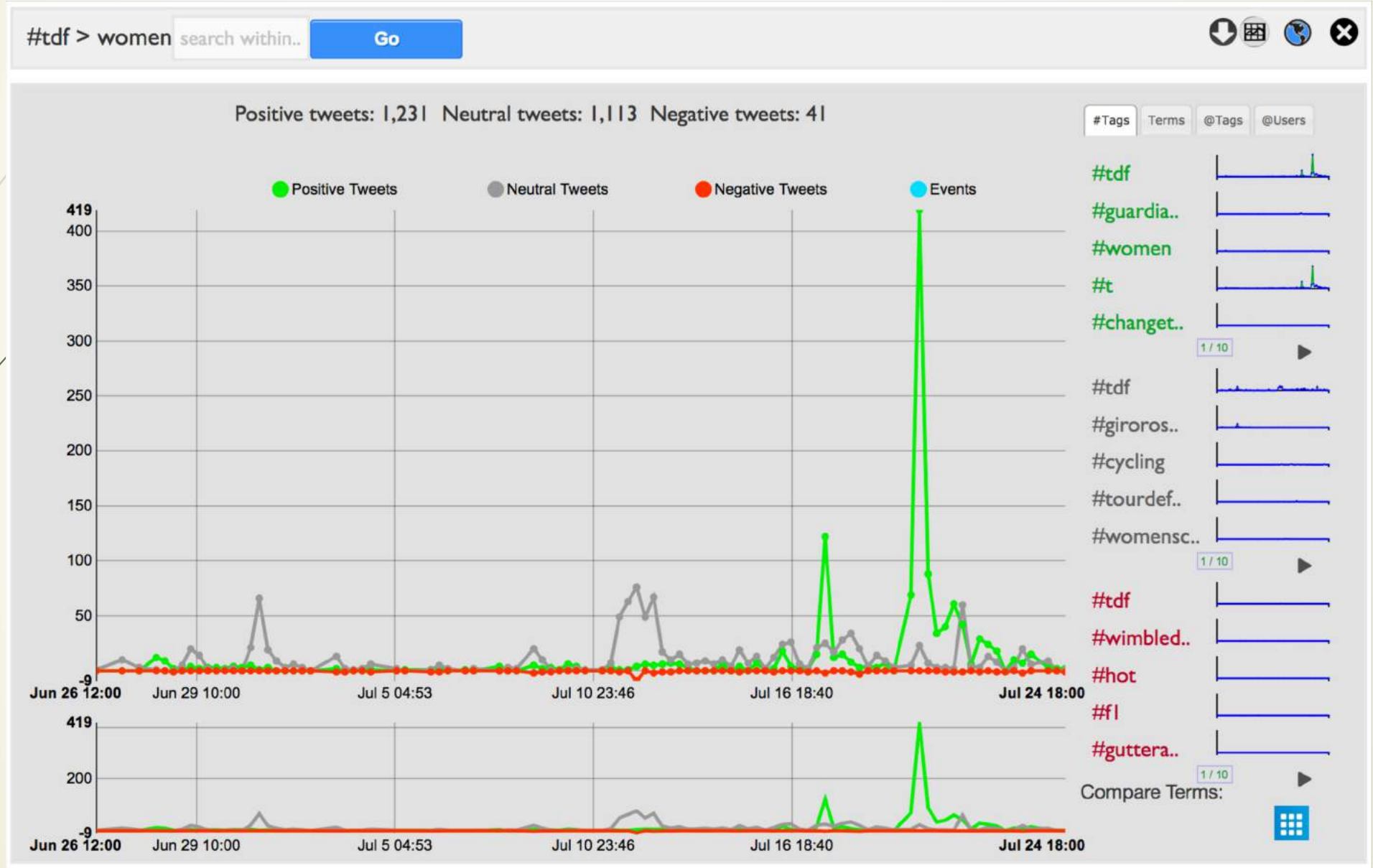

[focus formulation] narrow things down to a refined goal
[focus formulation] perform specific searches
[focus formulation] detailed analysis of the search results
[pre-focus] develop an understanding
[post-focus] make sense of the information
[pre-focus] broad search on the topic
What information is useful?
How can we visualize it?
[post-focus] synthesize what has been found

Summary

- We have outlined two important aspects of information seeking behaviour:
  - exploratory search
  - sensemaking
- These processes are all intertwined, enabling searchers to undertake complex information seeking tasks
- We have also discussed a specific model that is often employed when a search task is complex:
  - Vakkari's three-stage model of information seeking

Example

Temporal Patterns: Vista (Hoeber et al., 2016)



Discussion

Information visualization techniques

- timeline
- colour encoding (positive, neutral, negative)
- small multiples sparklines

Information to be conveyed to searcher

- overall temporal patterns
- top #tags/terms/@tags/users and their specific temporal patterns

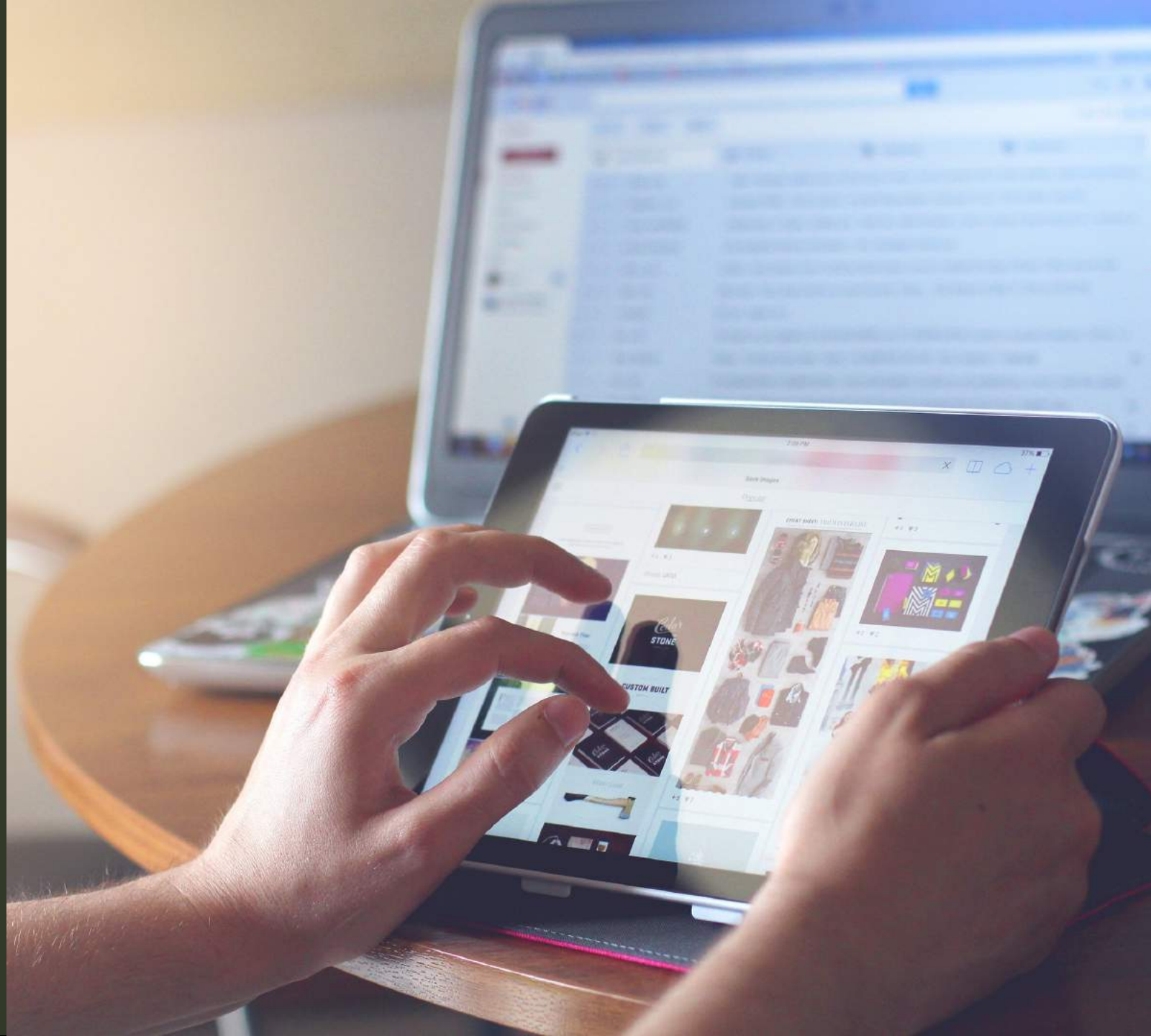
Information seeking behaviour support

- exploratory search
- sensemaking

Vakkari's three stage model of information seeking

- pre-focus – visual overview; recognition of #tags/terms/@tags/users
- focus formulation - sub-querying and comparisons
- post-foucs – save timelines (subqueries); export the data

But wait.
Sometimes I don’t need all of this complexity.



Not a Silver-Bullet Solution

- These approaches for adding visualization to information retrieval interfaces should not be viewed as a silver-bullet solution to all search problems
- There is a computational cost for performing information extraction, and a cognitive cost in learning how to interpret and interact with the visual interfaces
- In simple search settings, the query box and search results list work well
- In complex search settings that are already imposing large cognitive costs on the searcher, InfoVis + IIR may be explored as a means for alleviating the low-level cognitive work, allowing the searcher to perform high-level exploration, reasoning, and decision-making

Recap

- Presentation Goals:
  1. understand the fundamental principles and theories of information visualization
  2. Appreciate how information visualization can enable and support various interactive information retrieval processes
  3. example of a visual search interface that uses information visualization to enable information seeking

- The goal is to leverage the power of automatic and intelligent information processing approaches, using these as the basis for providing visual and interactive support for searchers
  - move search beyond a simple filtering mechanism, making it a fundamental tool for for analyzing and understanding the information
  - turn the problems of information overload and big data into opportunities for enhancing human-centred search

References

- Information Visualization
  - J. Bertin (1983). Semiology of Graphics. Madison, WI: University of Wisconsin Press.
  - E. Hering (1964). Outlines of a Theory of Light Sense (Grundzge der Lehr von Lichtsinn, 1920). Cambridge, MA: Harvard University Press.
  - K. Kofka (1935). Principles of Gestalt Psychology. New York, NY: Harcourt-Brace.
  - M. Ward, G. Grinstein, and D. Keim (2015). Interactive Data Visualization: Foundations, Techniques, and Applications, second edition. Natick, MA: A K Peters/CRC Press,.
  - C. Ware (2008). Visual Thinking for Design. Waltham, MA: Morgan Kaufmann.

References

- Information Seeking Behaviour
  - A. H. Awadallah, R. W. White, P. Pantel, S. T. Dumais, & Y.-M. Wang (2014). Supporting complex search tasks. In Proceedings of the ACM International Conference on Information and Knowledge Management, p. 829-838. doi: 10.1145/2661829.2661912
  - O. Hoeber (2014). Visual search analytics: Combining machine learning and interactive visualization to support human-centred search. In Proceedings of Beyond Single-Shot Text Queries: Bridging the Gap(s) Between Research Communities Workshop, p. 37–43.
  - P. Pirolli (2007). Information Foraging Theory: Adaptive interaction with information. New York: Oxford University Press.
  - P. Pirolli & D. M. Russell (2011). Introduction to this special issue on sensemaking. Human- Computer Interaction, 26(1-2), 1–8. doi: 10.1080/07370024.2011.556557
  - R. W. White & R. A. Roth (2009). Exploratory search: Beyond the query-response paradigm. San Rafael, CA: Morgan & Claypool Publisher. doi: 10.2200/S00174ED1V01Y200901ICR003

References

- Example
  - O. Hoeber, L. Hoeber, M. El Meseery, K. Odoh, and R. Gopi (2016). Visual Twitter analytics (Vista): Temporally changing sentiment and the discovery of emergent themes within sport event tweets. Online Information Review 40(1): 25–41.
  - O. Hoeber, R. Snelgrove, L. Hoeber, L. Wood (2017). A Systematic Methodology for Preserving the Whole in Large-Scale Qualitative-Temporal Research, Journal of Sport Management, 31(4):378-400.
  - O. Hoeber (2018). Information visualization for interactive information retrieval, In Proceedings of the ACM SIGIR Conference on Human Information Interaction and Retrieval, 371-374.