Principles of effective data visualization

Chris Belter, NICHD

christopher.belter@nih.gov

Agenda

Part One:

The Process

Part Two:

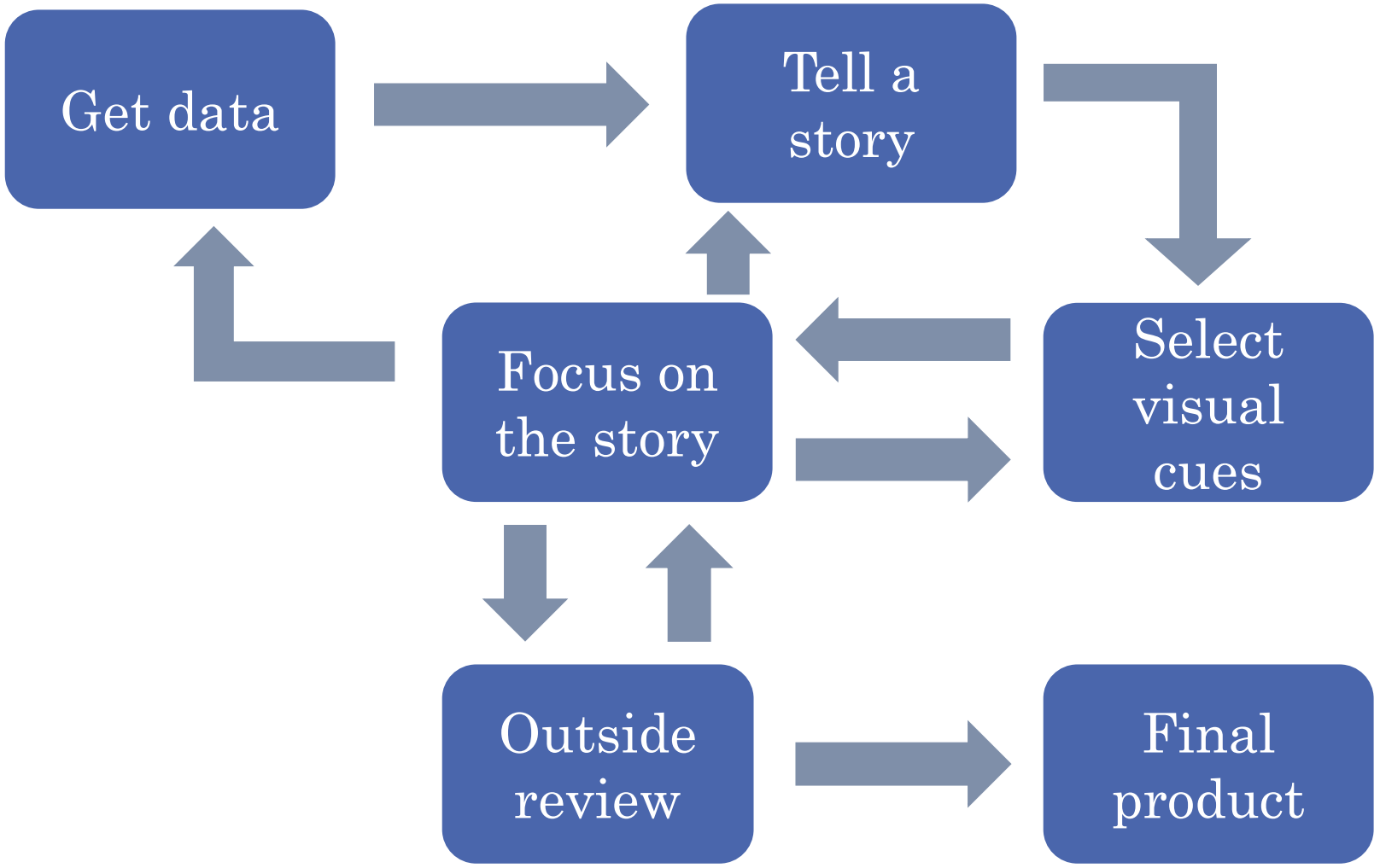
Example

Part Three:

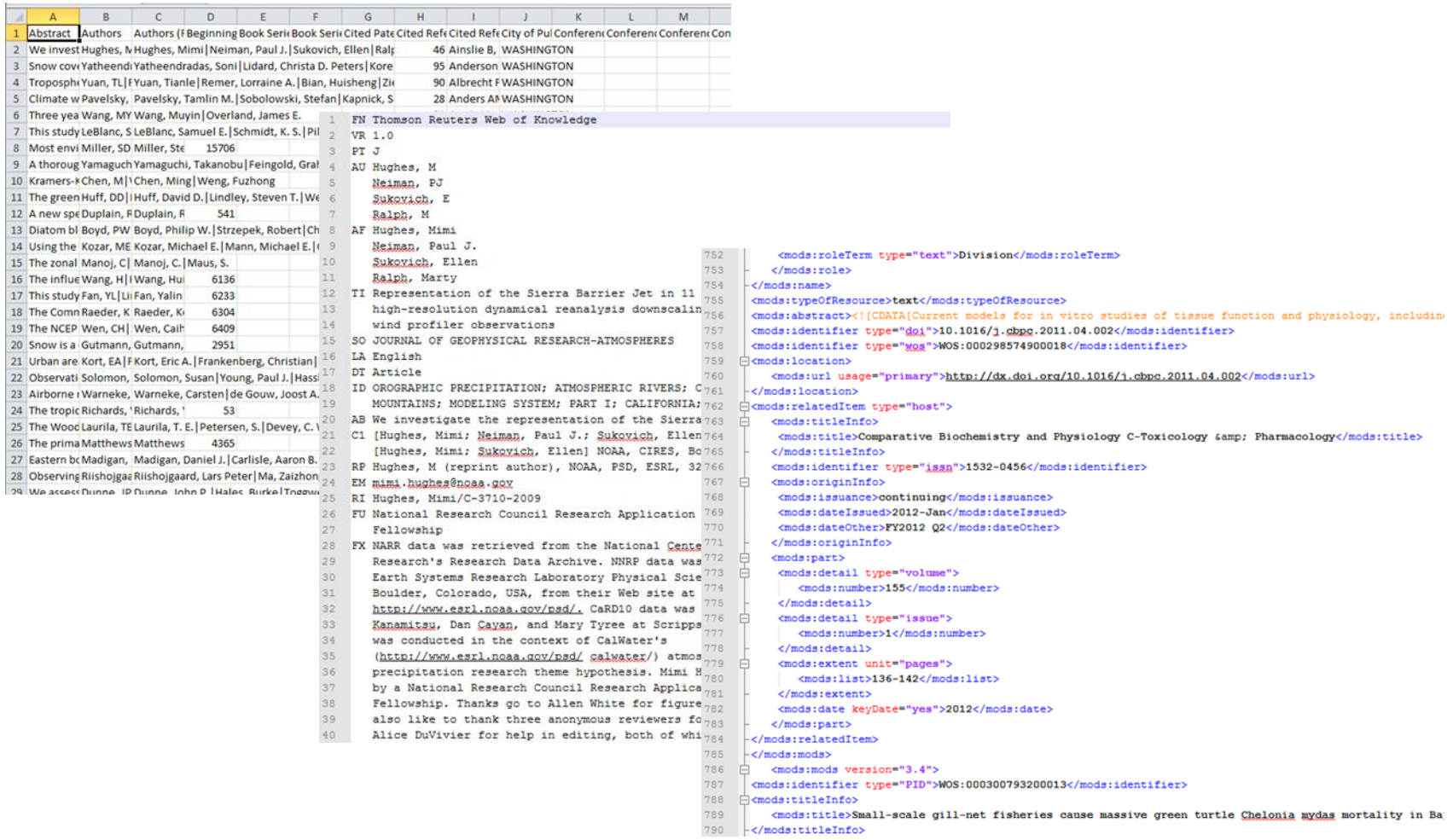
Resources

Part One:
The Process

Process overview



Step 0: get data



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Abstract	Authors	Authors (F	Beginning	Book Seri	Book Seri	Cited Pate	Cited Refe	Cited Refe	City of Pul	Conferen	Conferen	Conferen	Con
2	We invest	Hughes, M	Hughes, Mimi | Neiman, Paul J. | Sukovich, Ellen | Ralp					46	Ainslie B,	WASHINGTON				
3	Snow cov	Yatheendi	Yatheendradas, Soni | Lidard, Christa D. Peters | Kore					95	Anderson	WASHINGTON				
4	Troposphe	Yuan, TL | F	Yuan, Tianle | Remer, Lorraine A. | Bian, Huisheng | Zi					90	Albrecht F	WASHINGTON				
5	Climate w	Pavelsky,	Pavelsky, Tamlin M. | Sobolowski, Stefan | Kapnick, S					28	Anders AI	WASHINGTON				
6	Three yea	Wang, MY	Wang, Muyin | Overland, James E.											
7	This study	LeBlanc, S	LeBlanc, Samuel E. | Schmidt, K. S. | Pil											
8	Most envi	Miller, SD	Miller, Ste	15706										
9	A thoroug	Yamaguch	Yamaguchi, Takanobu | Feingold, Grat											
10	Kramers-k	Chen, M | '	Chen, Ming | Weng, Fuzhong											
11	The green	Huff, DD |	Huff, David D. | Lindley, Steven T. | We											
12	A new spe	Duplain, F	Duplain, F	541										
13	Diatom bl	Boyd, PW	Boyd, Philip W. | Strzepek, Robert | Ch											
14	Using the	Kozar, ME	Kozar, Michael E. | Mann, Michael E. | (											
15	The zonal	Manoj, C |	Manoj, C. | Maus, S.											
16	The influe	Wang, H | I	Wang, Hu	6136										
17	This study	Fan, YL | Li	Fan, Yalin	6233										
18	The Comn	Raeder, K	Raeder, K	6304										
19	The NCEP	Wen, CH |	Wen, Caih	6409										
20	Snow is a	Gutmann,	Gutmann,	2951										
21	Urban are	Kort, EA | F	Kort, Eric A. | Frankenberg, Christian |										
22	Observati	Solomon,	Solomon, Susan | Young, Paul J. | Hass											
23	Airborne s	Warneke,	Warneke, Carsten | de Gouw, Joost A.											
24	The tropic	Richards, '	Richards, '	53										
25	The Wooc	Laurila, TE	Laurila, T. E. | Petersen, S. | Devey, C. \											
26	The prima	Matthews	Matthews	4365										
27	Eastern bc	Madigan,	Madigan, Daniel J. | Carlisle, Aaron B.											
28	Observing	Riishojgaa	Riishojgaard, Lars Peter | Ma, Zaizhon											
29	We assess	Dunne, JP	Dunne, John P. | Hales, Burke | Toggw											

```
FN Thomson Reuters Web of Knowledge
VR 1.0
PT J
AU Hughes, M
   Neiman, PJ
   Sukovich, E
   Ralph, M
AF Hughes, Mimi
   Neiman, Paul J.
   Sukovich, Ellen
   Ralph, Marty
TI Representation of the Sierra Barrier Jet in 11
   high-resolution dynamical reanalysis downscalin
   wind profiler observations
SO JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES
LA English
DT Article
ID OROGRAPHIC PRECIPITATION; ATMOSPHERIC RIVERS; C
   MOUNTAINS; MODELING SYSTEM; PART I; CALIFORNIA;
AB We investigate the representation of the Sierra
C1 [Hughes, Mimi; Neiman, Paul J.; Sukovich, Ellen
   [Hughes, Mimi; Sukovich, Ellen] NOAA, CIRES, Bo
RP Hughes, M (reprint author), NOAA, PSD, ESRL, 32
EM mimi.hughes@noaa.gov
RI Hughes, Mimi/C-3710-2009
FU National Research Council Research Application
   Fellowship
FX NARR data was retrieved from the National Cente
   Research's Research Data Archive. NNRP data was
   Earth Systems Research Laboratory Physical Scie
   Boulder, Colorado, USA, from their Web site at
   http://www.esrl.noaa.gov/psd/. CaRD10 data was
   Kanamitsu, Dan Cayan, and Mary Tyree at Scripps
   was conducted in the context of CalWater's
   (http://www.esrl.noaa.gov/psd/ calwater/) atmos
   precipitation research theme hypothesis. Mimi H
   by a National Research Council Research Applica
   Fellowship. Thanks go to Allen White for figure
   also like to thank three anonymous reviewers fo
   Alice DuVivier for help in editing, both of whi
```

```
      <mods:roleTerm type="text">Division</mods:roleTerm>
    </mods:role>
</mods:name>
<mods:typeOfResource>text</mods:typeOfResource>
<mods:abstract><![CDATA[Current models for in vitro studies of tissue function and physiology, includin
<mods:identifier type="doi">10.1016/j.cbpc.2011.04.002</mods:identifier>
<mods:identifier type="wos">WOS:000298574900018</mods:identifier>
<mods:location>
    <mods:url usage="primary">http://dx.doi.org/10.1016/j.cbpc.2011.04.002</mods:url>
</mods:location>
<mods:relatedItem type="host">
    <mods:titleInfo>
      <mods:title>Comparative Biochemistry and Physiology C-Toxicology &amp;amp; Pharmacology</mods:title>
    </mods:titleInfo>
    <mods:identifier type="issn">1532-0456</mods:identifier>
    <mods:originInfo>
      <mods:issuance>continuing</mods:issuance>
      <mods:dateIssued>2012-Jan</mods:dateIssued>
      <mods:dateOther>FY2012 Q2</mods:dateOther>
    </mods:originInfo>
    <mods:part>
      <mods:detail type="volume">
        <mods:number>155</mods:number>
      </mods:detail>
      <mods:detail type="issue">
        <mods:number>1</mods:number>
      </mods:detail>
      <mods:extent unit="pages">
        <mods:list>136-142</mods:list>
      </mods:extent>
      <mods:date keyDate="yes">2012</mods:date>
    </mods:part>
</mods:relatedItem>
</mods:mods>
    <mods:mods version="3.4">
<mods:identifier type="PID">WOS:000300793200013</mods:identifier>
<mods:titleInfo>
    <mods:title>Small-scale gill-net fisheries cause massive green turtle Chelonia mydas mortality in Ba
</mods:titleInfo>
```

Step 1: tell a story

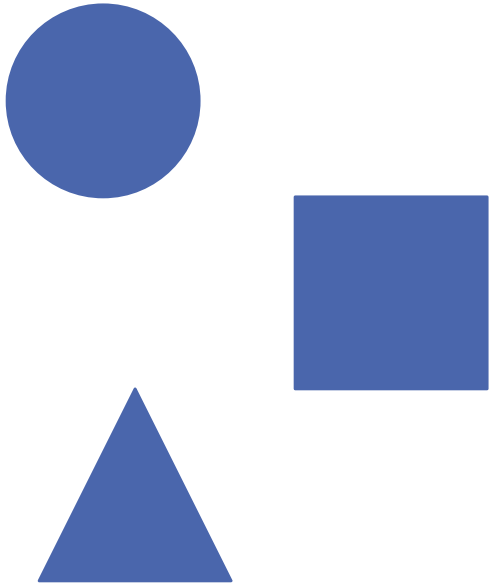
```
752      <mods:roleTerm type="text">Division</mods:roleTerm>
753    </mods:role>
754  </mods:name>
755  <mods:typeOfResource>text</mods:typeOfResource>
756  <mods:abstract><![CDATA[Current models for in vitro studies of tissue function and physiology, includin
757  <mods:identifier type="doi">10.1016/j.cbpc.2011.04.002</mods:identifier>
758  <mods:identifier type="wos">WOS:000298574900018</mods:identifier>
759  <mods:location>
760      <mods:url usage="primary">http://dx.doi.org/10.1016/j.cbpc.2011.04.002</mods:url>
761  </mods:location>
762  <mods:relatedItem type="host">
763      <mods:titleInfo>
764        <mods:title>Comparative Biochemistry and Physiology C-Toxicology &amp;amp; Pharmacology</mods:title>
765      </mods:titleInfo>
766      <mods:identifier type="issn">1532-0456</mods:identifier>
767      <mods:originInfo>
768        <mods:issuance>continuing</mods:issuance>
769        <mods:dateIssued>2012-Jan</mods:dateIssued>
770        <mods:dateOther>FY2012 Q2</mods:dateOther>
771      </mods:originInfo>
772      <mods:part>
773        <mods:detail type="volume">
774          <mods:number>155</mods:number>
775        </mods:detail>
776        <mods:detail type="issue">
777          <mods:number>1</mods:number>
778        </mods:detail>
779        <mods:extent unit="pages">
780          <mods:list>136-142</mods:list>
781        </mods:extent>
782        <mods:date keyDate="yes">2012</mods:date>
783      </mods:part>
784  </mods:relatedItem>
785  </mods:mods>
786      <mods:mods version="3.4">
787  <mods:identifier type="PID">WOS:000300793200013</mods:identifier>
788  <mods:titleInfo>
789      <mods:title>Small-scale gill-net fisheries cause massive green turtle Chelonia mydas mortality in Ba
790  </mods:titleInfo>
```

Step 2: select visual cues

Visual Cue	Category	Quantity
Shape	Yes	Yes
Position	Yes	Yes
Angle	No	Yes
Color	Yes	Yes

Visual cues: shape

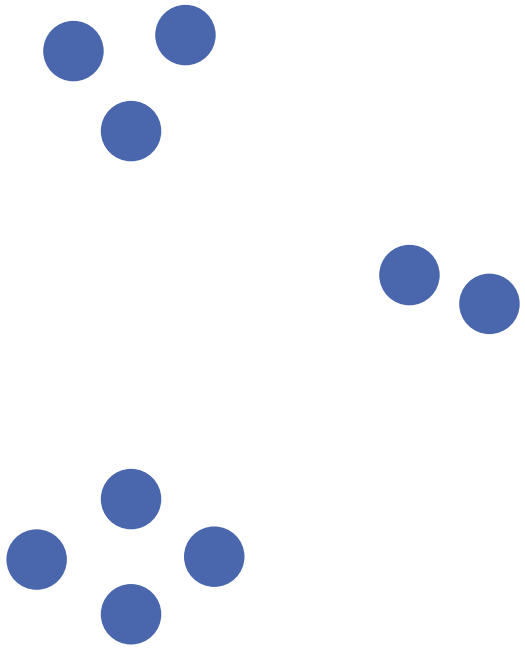
Category



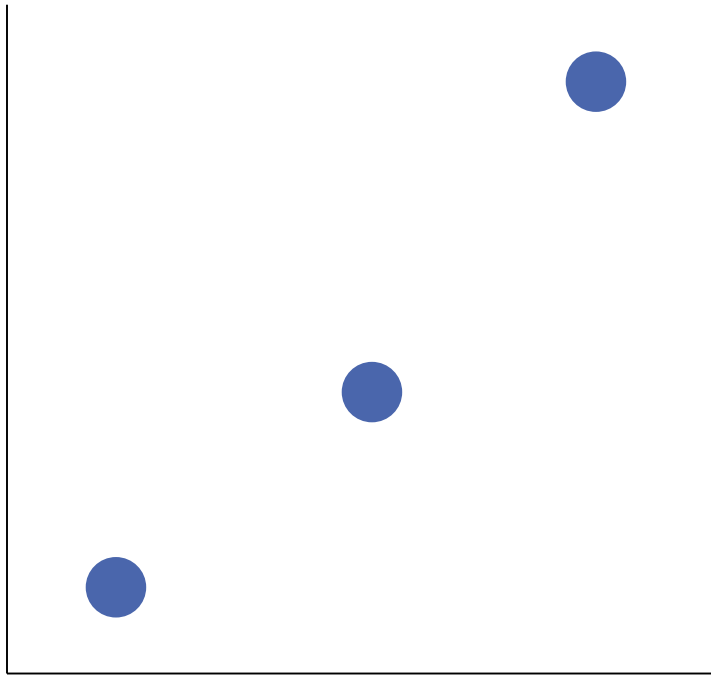
Quantity

Visual cues: position

Category



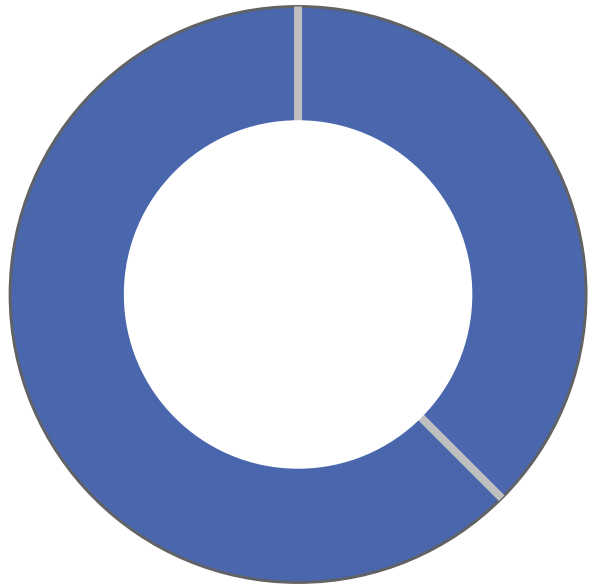
Quantity



Visual cues: angle

Quantity (1)

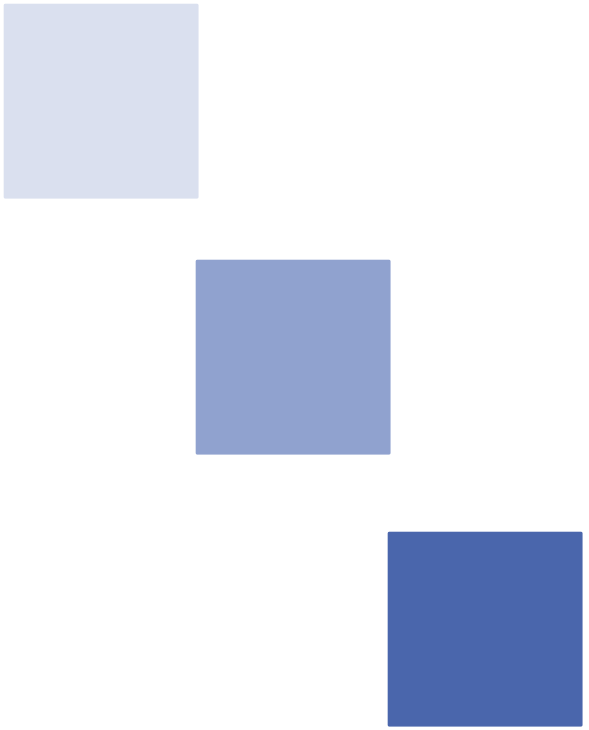
Quantity (2)



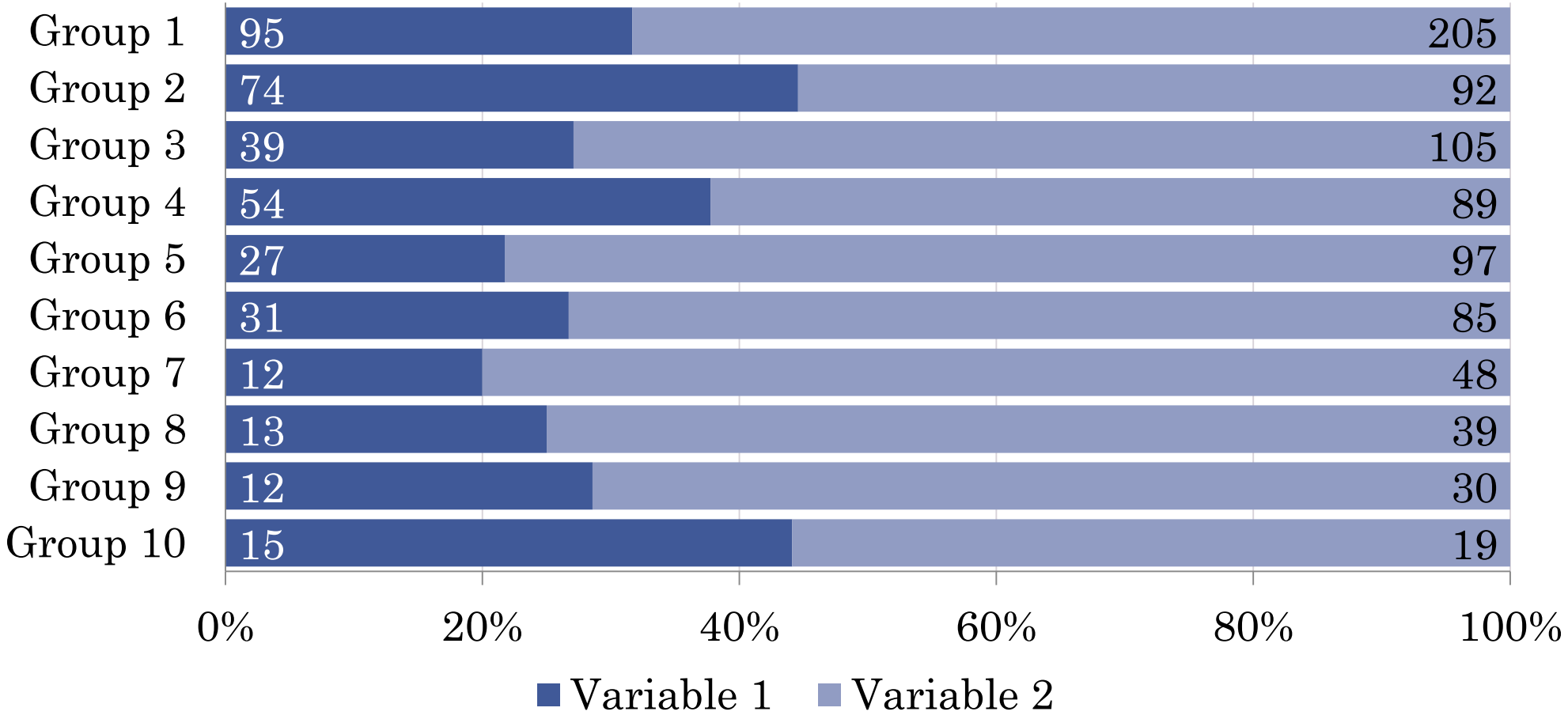
Visual cues: color

Category

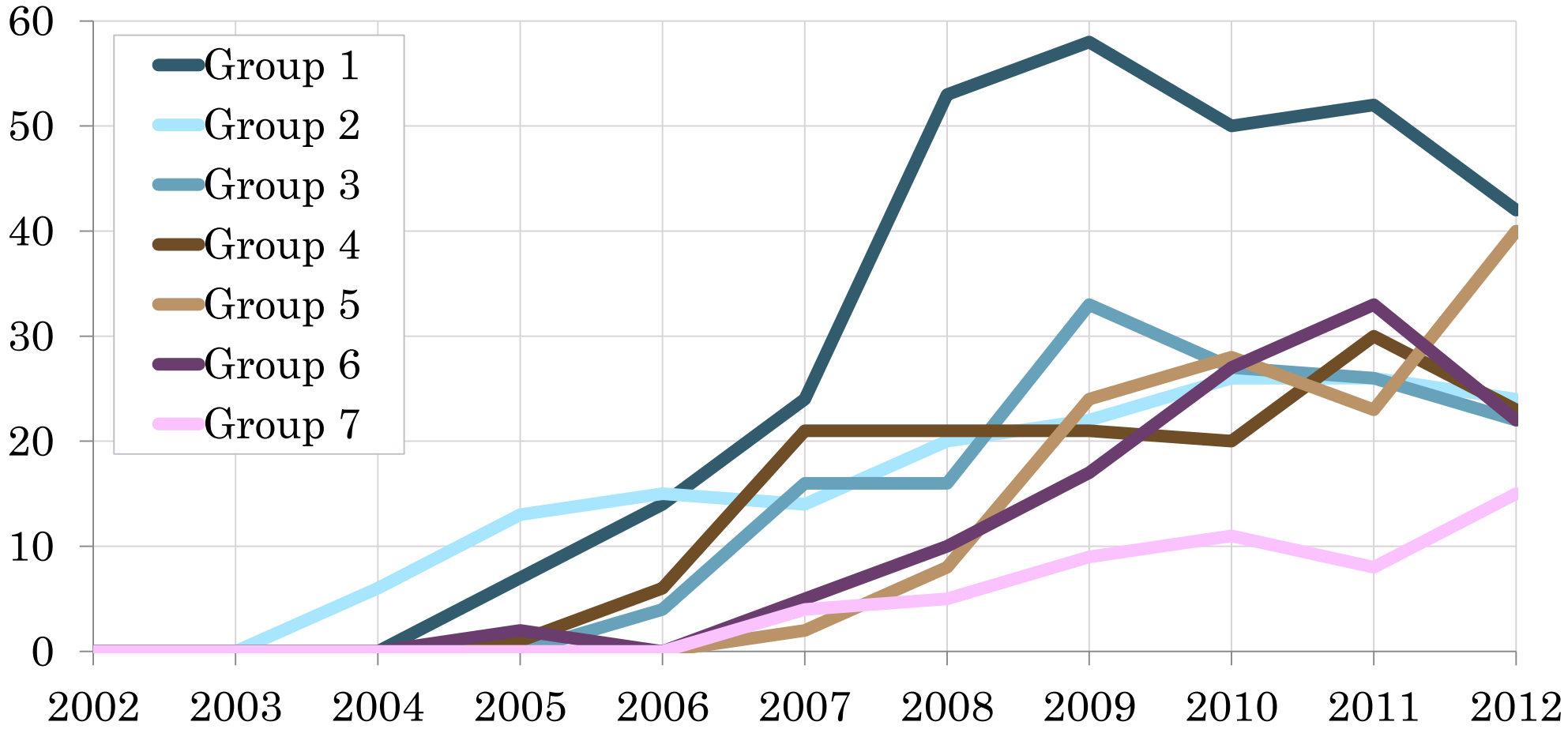
Quantity



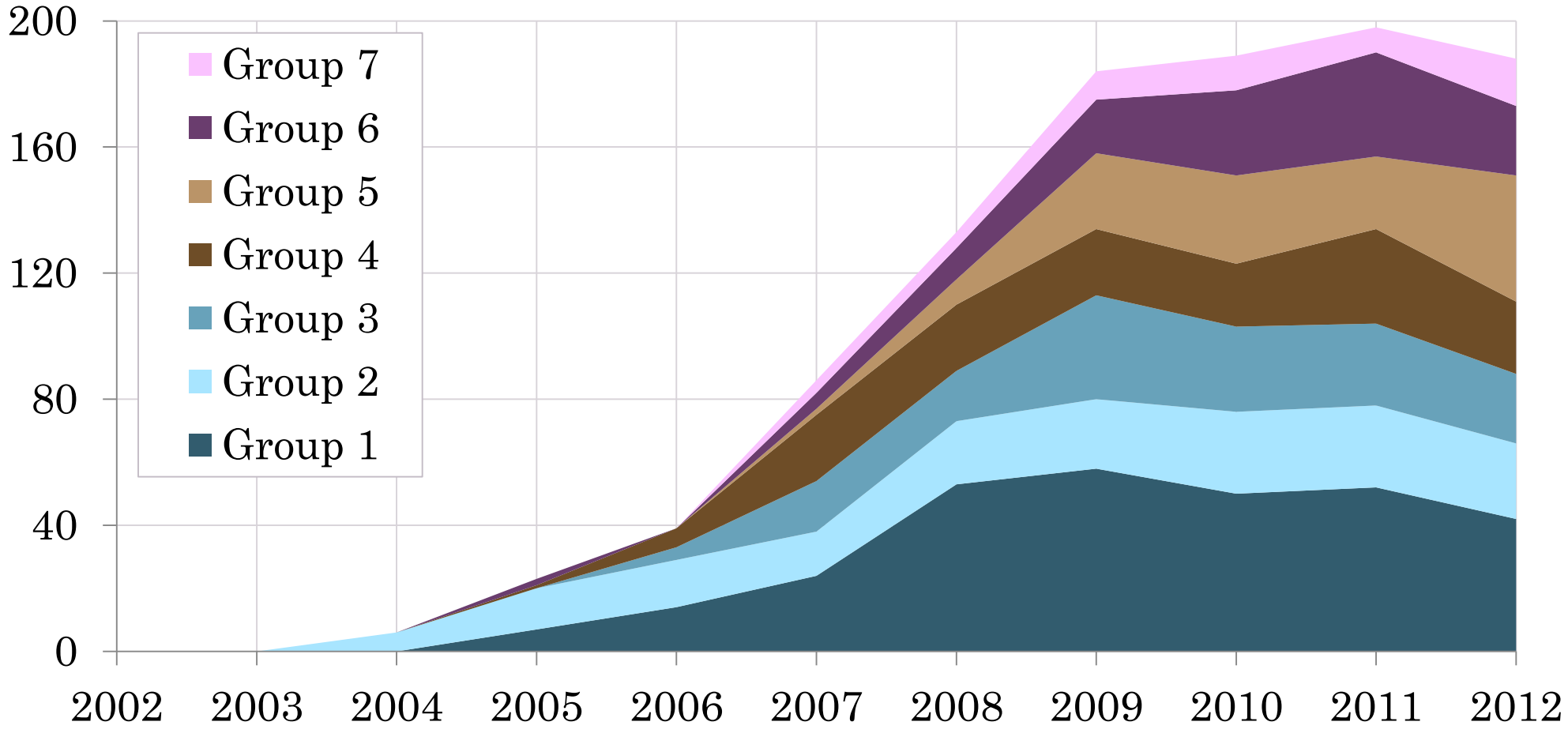
Combining cues (1)



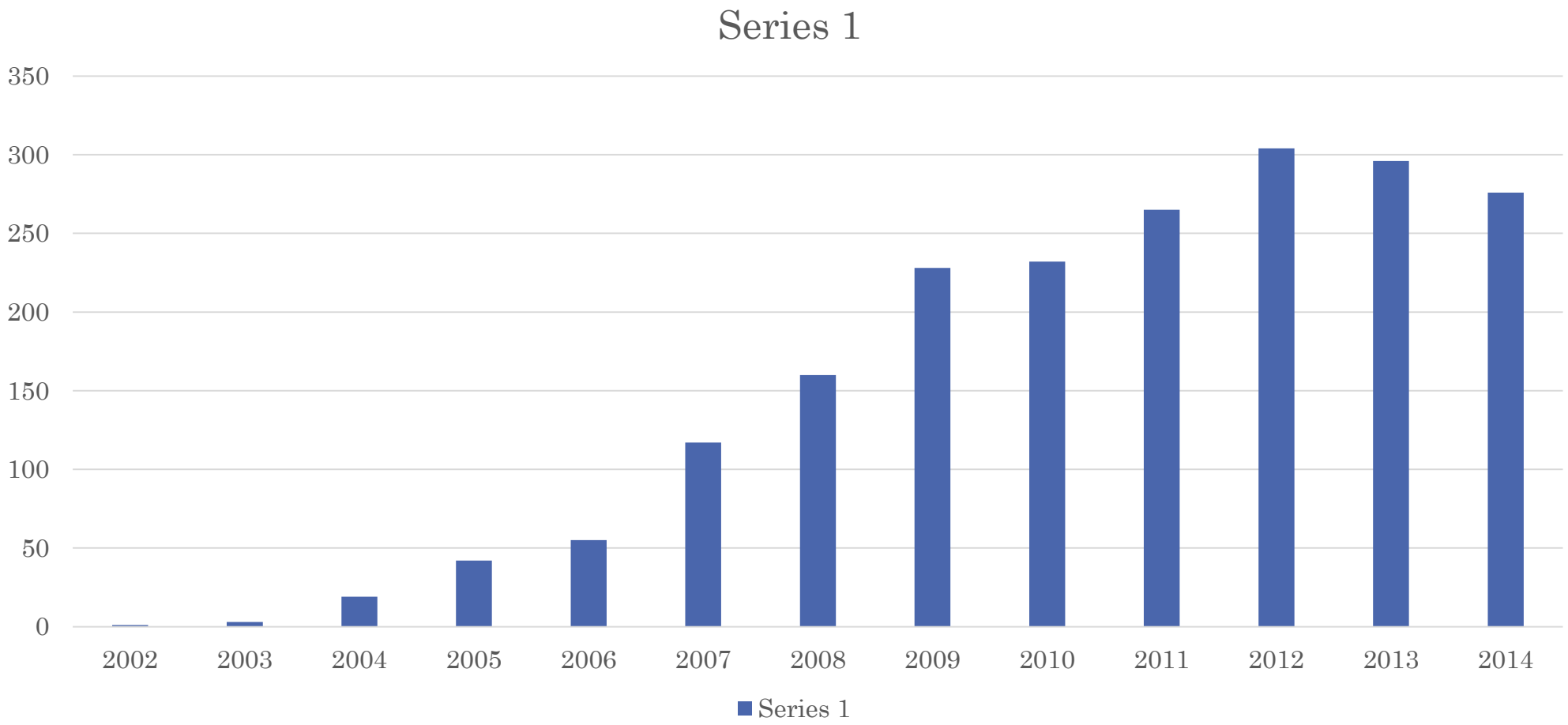
Combining cues (2)



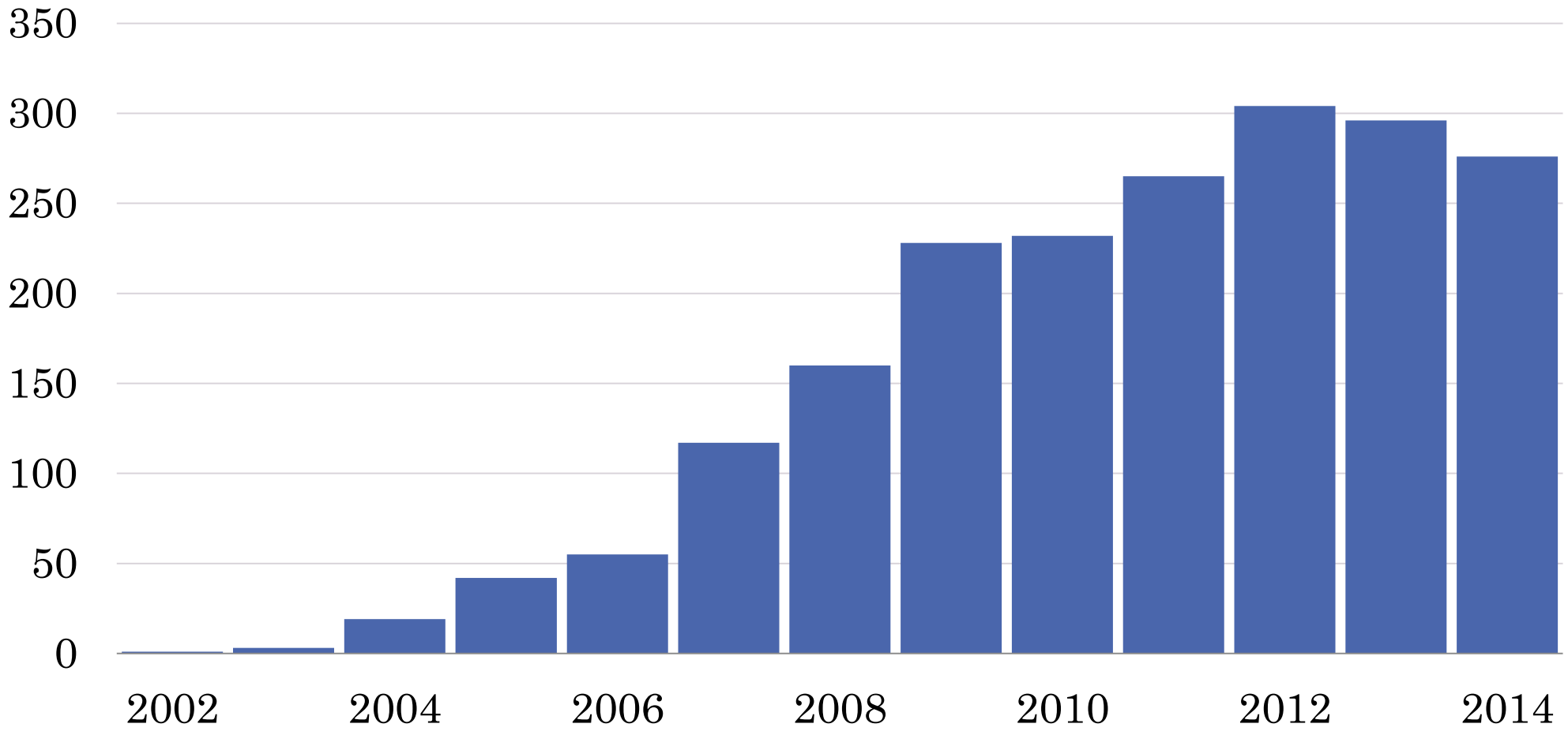
Combining cues (3)



Step 3: focus on the story



Step 3: focus on the story



Step 4: outside review

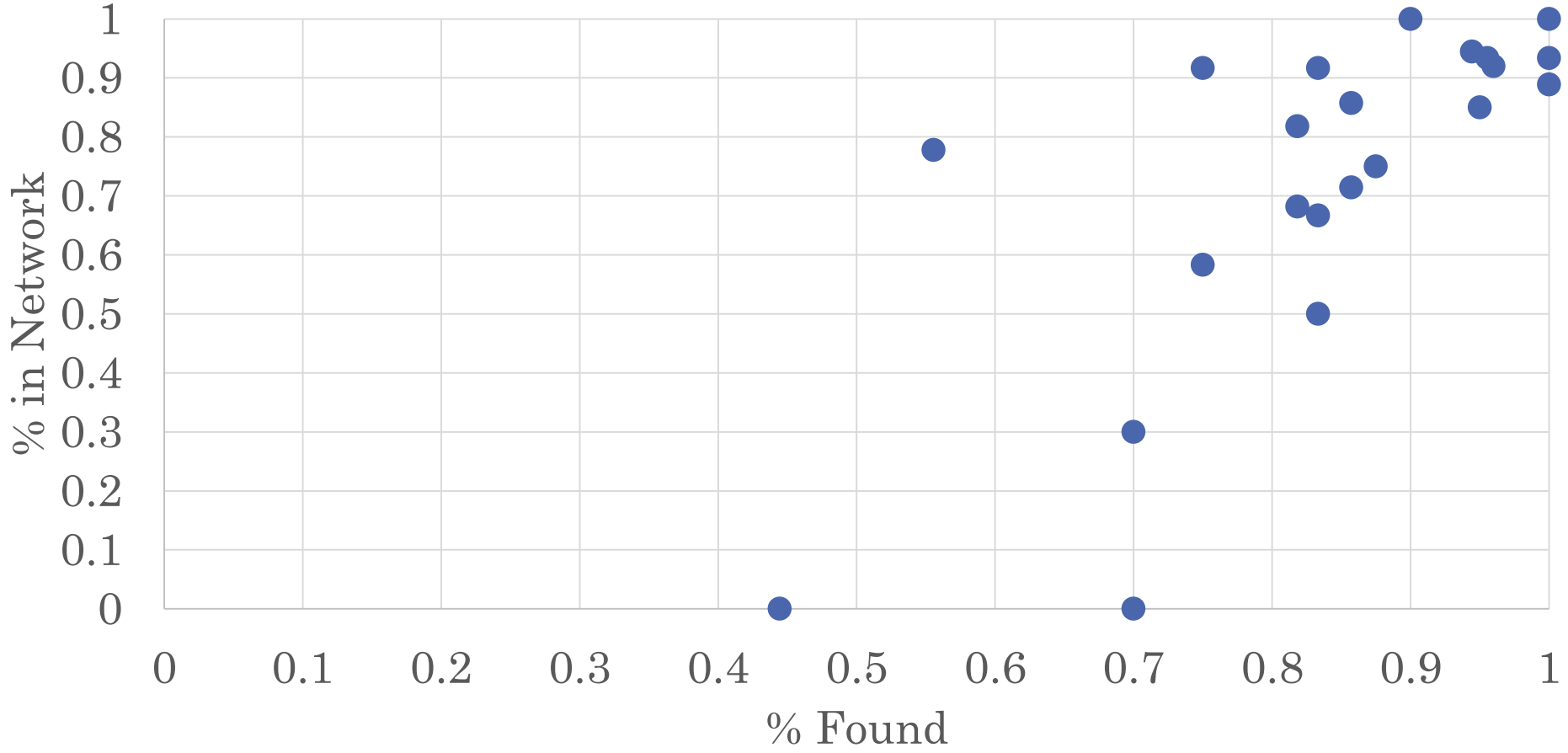
Part Two: Example

Steps 0 and 1

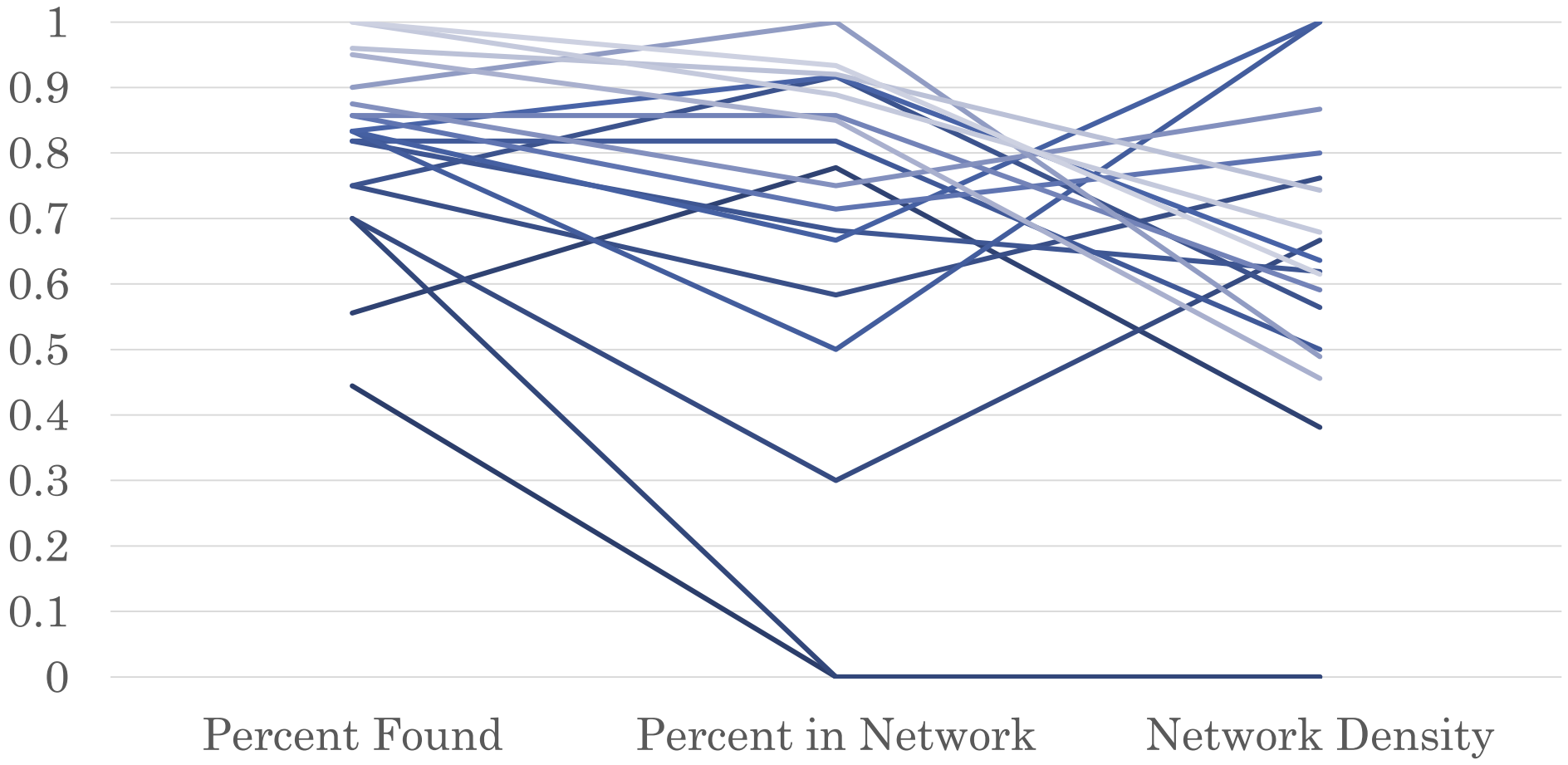
Data: 4 different variables for 23 systematic reviews

Story: recall increases as paper coherence increases

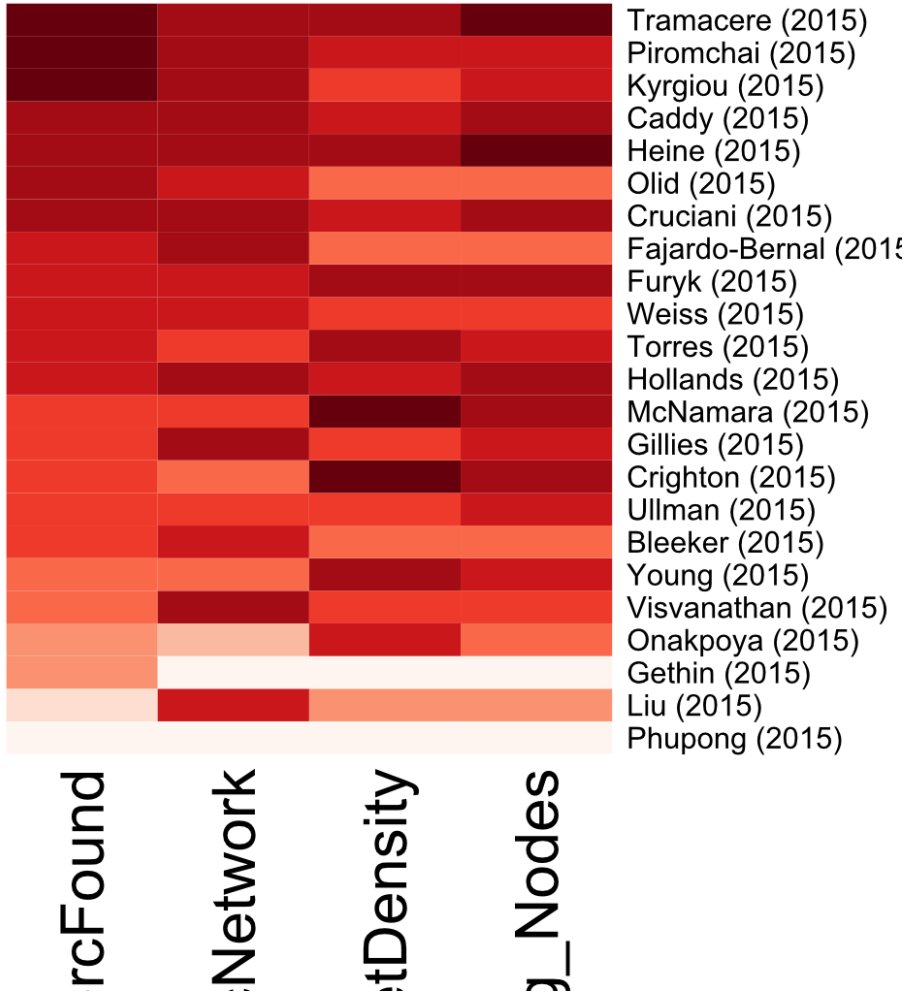
Step 2: trial one



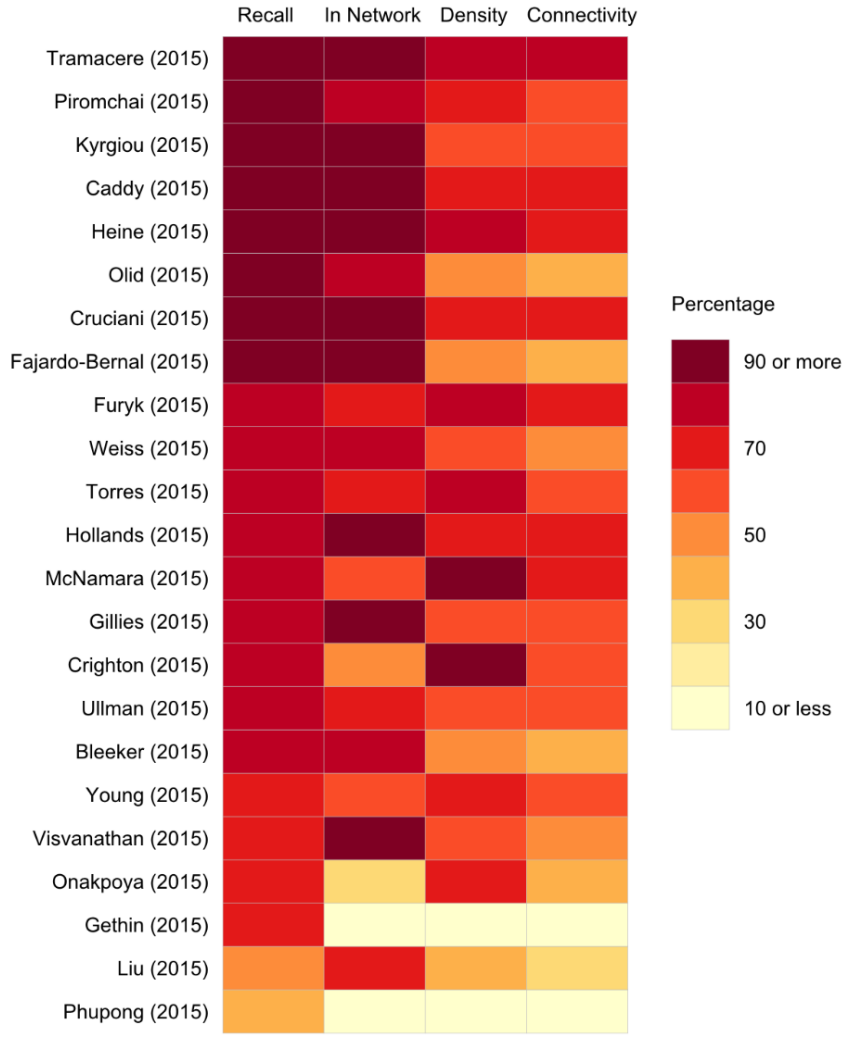
Step 2: trial two



Step 2: trial three



Step 3


Recall
In Network
Density
Connectivity
Tramacere (2015)
Piromchai (2015)
Kyrgiou (2015)
Caddy (2015)
Heine (2015)
Olid (2015)
Cruciani (2015)
Fajardo-Bernal (2015)
Furyk (2015)
Weiss (2015)
Torres (2015)
Hollands (2015)
McNamara (2015)
Gillies (2015)
Crighton (2015)
Ullman (2015)
Bleeker (2015)
Young (2015)
Visvanathan (2015)
Onakpoya (2015)
Gethin (2015)
Liu (2015)
Phupong (2015)
Percentage
90 or more
70
50
30
10 or less

Part Three: Resources

Online resources

Blogs and Galleries

- https://flowingdata.com
- https://informationisbeautiful.net
- https://www.nytimes.com/section/upshot

Color Palette Tools

- www.colorbrewer2.org
- https://color.adobe.com

Open-Source software

R: data cleaning, analysis, visualization, etc., etc., etc.

Sci2: network analysis and geospatial visualization

Gephi: network visualization

VOSviewer: network analysis and visualization

Inkscape: visualization polishing and editing

Questions?