The Application of Concept Mapping to Text Analysis:

Examples and Reliability/Validity Issues

Kristin Behfar
University of California, Irvine
December 5, 2008

Concept Mapping

- Variants of the methodology
  - Visual/Drawn (e.g., Novak, 1998; Axlerod, 1976)
  - Statistical (e.g., Carley & Kaufer, 1993)
  - Blend of above variants and hybrid of other analysis approaches
    - Codes (e.g., grounded theory, content analysis, schema analysis, etc.)
    - Words (e.g., KWIC, semantic networks, cognitive maps, etc.)

- Application to open-ended responses
  - Nature of the research question
  - Characteristics of text

Jackson, K., & Trochim, W. (2002). Concept mapping as an alternative approach for the analysis of open-ended survey questions. *Organizational Research Methods*, *5*(4), 307-336.

Concept Mapping

Step 1: Unitizing

Each Statement Contains Only One Idea

E.g.: "*Commitments were made initially but were not followed up or backed up with the required efforts. Opinions strictly related to projects and write-ups were often interpreted by my teammates as personal comments.*"

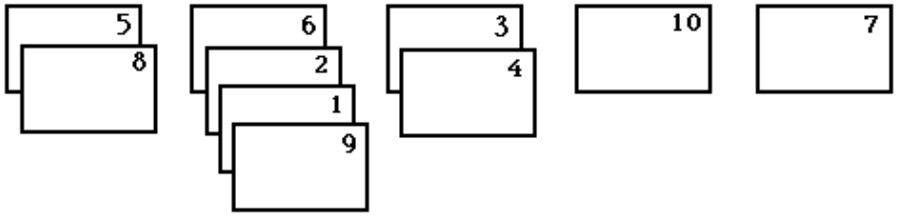
1) Commitments were made initially but were not followed up or backed up with the required efforts.

2) Opinions strictly related to projects and write-ups were often interpreted by my teammates as personal comments.

Concept Mapping

Step 2: Sorting

Hypothetical Sort



Matrix for each sorter

Binary Square Symmetric Similarity Matrix

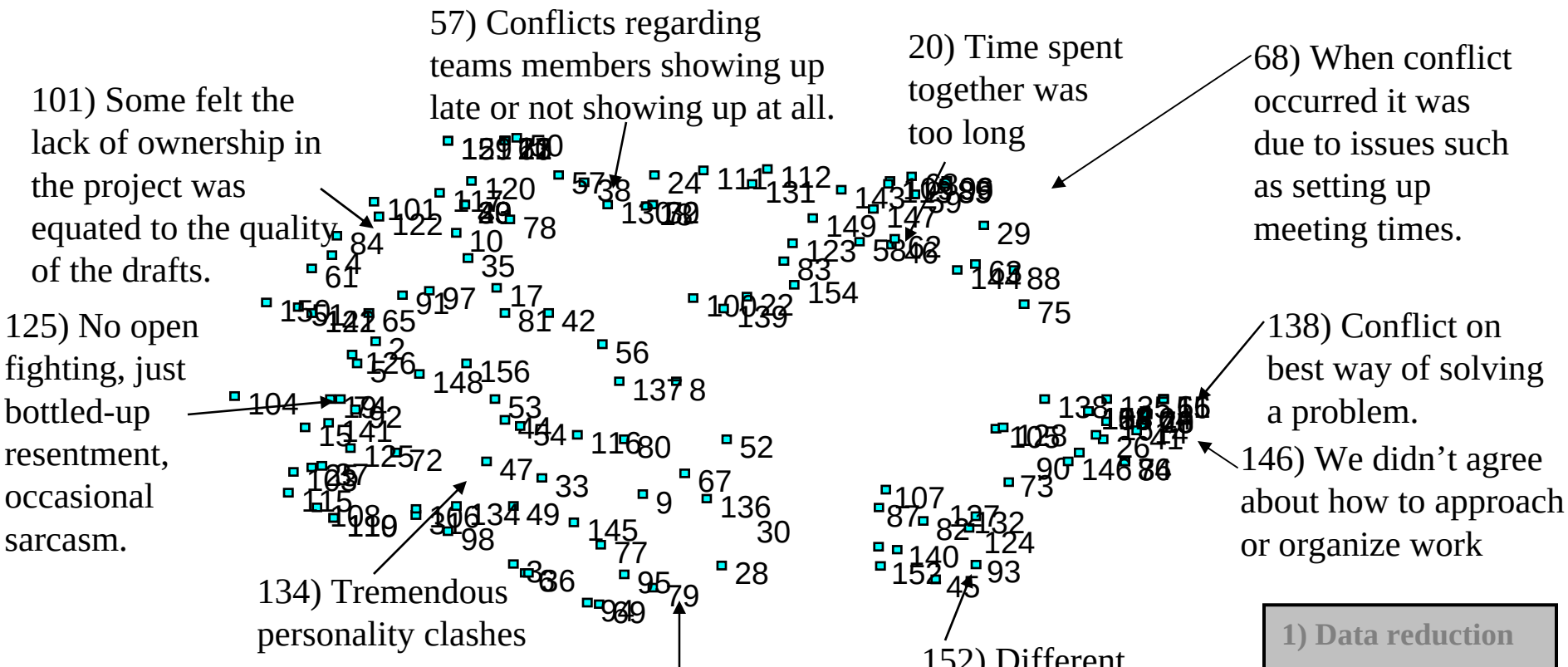
	1	2	3	4	5	6	7	8	9	10
1	1	1	0	0	0	1	0	0	1	0
2	1	1	0	0	0	1	0	0	1	0
3	0	0	1	1	0	0	0	0	0	0
4	0	0	1	1	0	0	0	0	0	0
5	0	0	0	0	1	0	0	1	0	0
6	1	1	0	0	0	1	0	0	1	0
7	0	0	0	0	0	0	1	0	0	0
8	0	0	0	0	1	0	0	1	0	0
9	1	1	0	0	0	1	0	0	1	0
10	0	0	0	0	0	0	0	0	0	1

Aggregate all matrices

1) Data reduction
2) Sorting
3) MDS
4) Cluster analysis
5) Cluster labeling

Concept Mapping

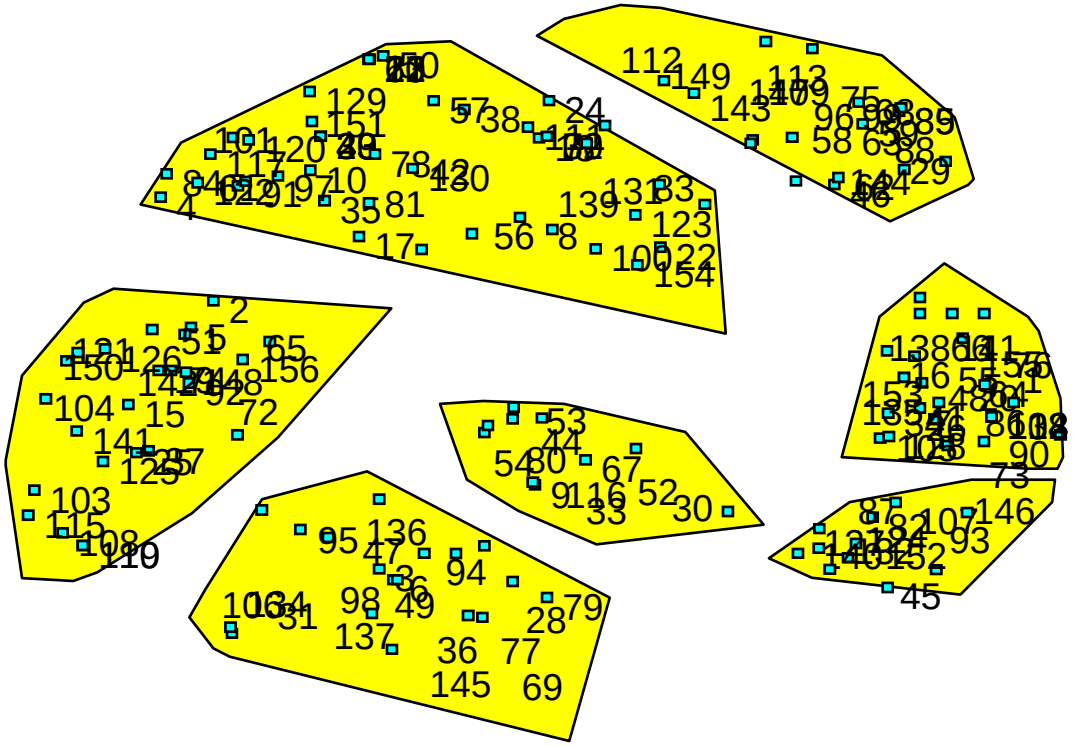
Step 3: Multidimensional Scaling



1) Data reduction
2) Sorting
3) MDS
4) Cluster analysis
5) Cluster labeling

Concept Mapping

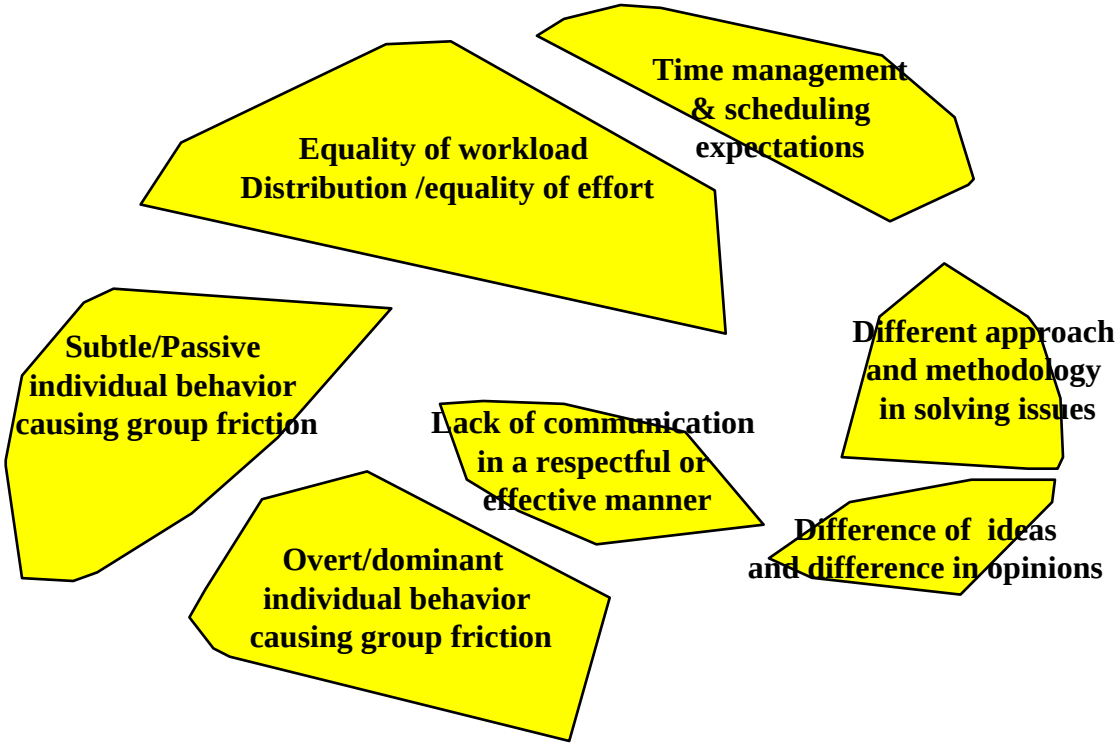
Step 4 : Cluster Analysis



1) Data reduction
2) Sorting
3) MDS
4) Cluster analysis
5) Cluster labeling

Concept Mapping

Step 5: Labeling the Final Solution



1) Data reduction
2) Sorting
3) MDS
4) Cluster analysis
5) Cluster labeling

Assessing Reliability

- **Stability**: same coder at different times codes the same data in a similar manner
  - Repeat individual sorts and assess correlation
- **Reproducibility**: similar results can be reproduced, different times/locations/coders
  - Correlate individual matrix with aggregate ('errorless standard') (e.g., Trochim, 1993)
  - Choice of sorters and proxies is key
- **Accuracy**: amount of error (intra-and inter-observer disagreement, systematic deviation)
- **Other**
  - Some units are harder to code than others
  - Some categories are harder to understand than others
  - Subsets of categories can sometimes be confused with larger categories
  - Individual coders may be careless, inconsistent, or interdependent

(Krippendorf, 1980)

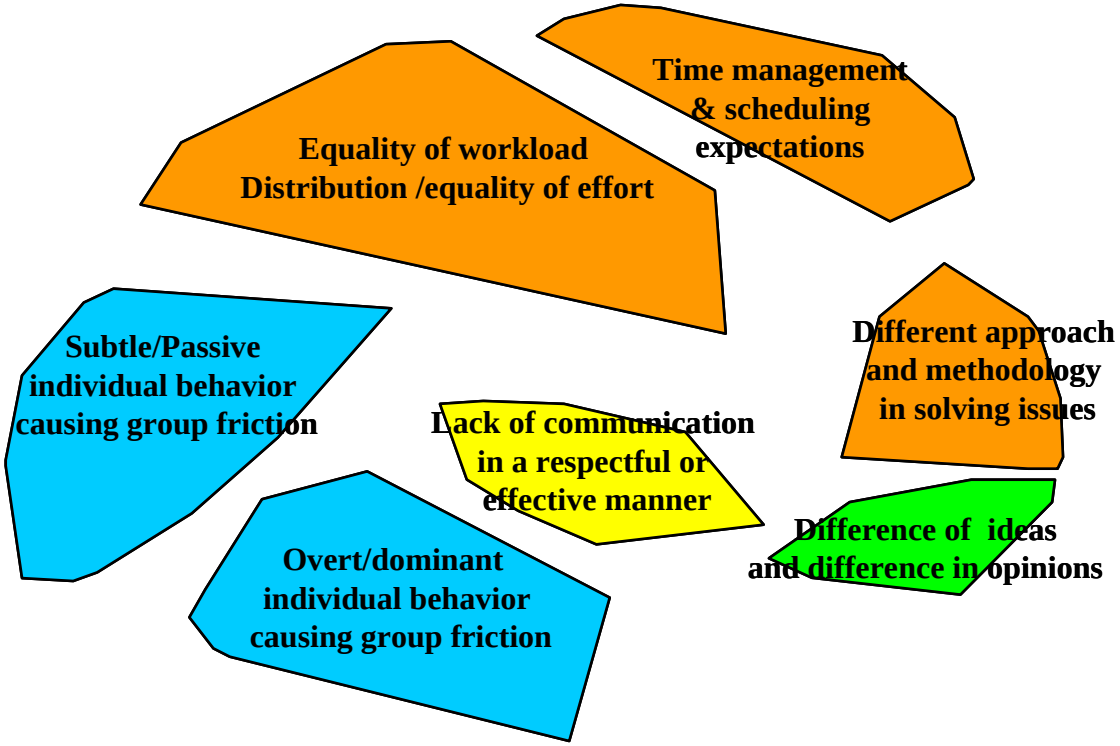
Validity

Depends on choices at each step of the analysis

- Construct
  - Population generating units (sampling)
  - Retaining context
    - Unitizing (internal)
    - Choice of sorters (external)
- Internal
- External: interpretation (theoretical) (e.g., Johnson & Turner, 2003)

Extending the Core Analysis, Ia

comparison to "expert" judgment



Purpose

- Comparison to expert Likert ratings of each unit
- Generation of new scale content
- Generate follow-up interview questions
- Clarify confusion around taxonomy and associated mixed empirical results

(Behfar, Mannix, Peterson, & Trochim, 2008)

Extending the Core Analysis, Ib

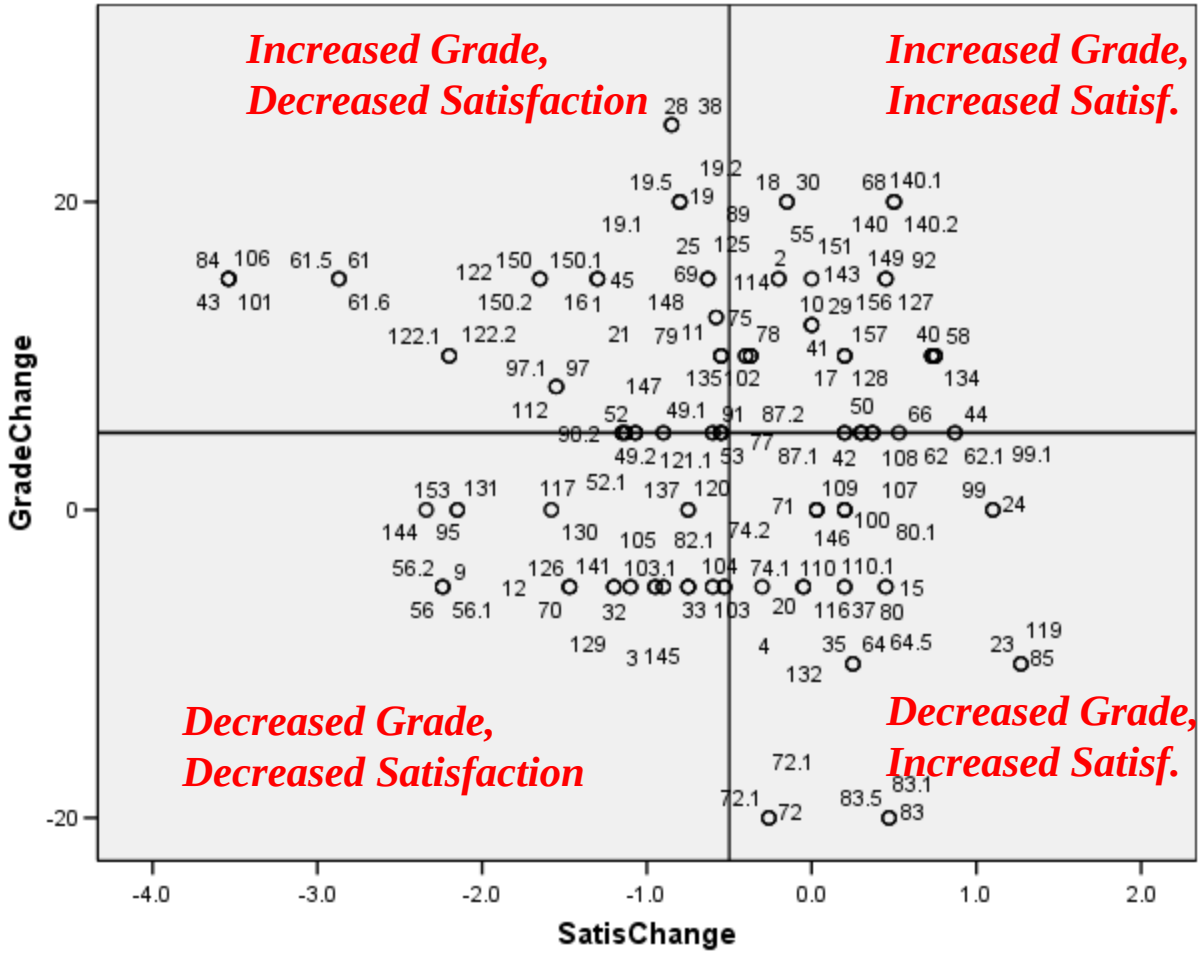
comparing two different demographic groups

Shared Challenges	Unique to Multicultural Teams
Direct vs. Indirect Confrontation Stylistic differences in publicly expressing and/or confronting different points of view, resulting in escalation of interpersonal tension.	**Violations of Respect & Hierarchy** Challenges stemming from different expectations for respecting hierarchy and other status indicators.
Norms for Problem Solving & Decision Making Differences in preference for a more slow-paced analytical problem solving and relationship building process versus a more efficiency focused approach.	**Inter-group Prejudices** Challenges stemming from innate or pre-existing stigma, prejudices, and judgments spilling over into the workplace.
Time, Urgency, & Pace Differences in time estimates to deliver products and the definition of "on-time" delivery.	**Lack of Common Ground (Language, Credit)** Challenges stemming from perceived favoritism or lack of recognition for contribution based how or how well members expressed themselves.
Differences in Work Norms & Behaviors Differences about what is acceptable workplace etiquette stemming either from national customs or national norms for separating personal time and work time.	**Fluency (Accents & Vocabulary)** Challenges caused by negative reactions and/or misunderstandings due to language issues such as heavy accents and words with different connotations.
	Thought You Had Agreement? Implicit versus Explicit Communication Challenges stemming from differences in interpretation about the level of commitment and/or agreement reached.

Behfar, Kern, & Brett (2006)

Extending the Core Analysis, II

associating qualitative responses with measured outcomes



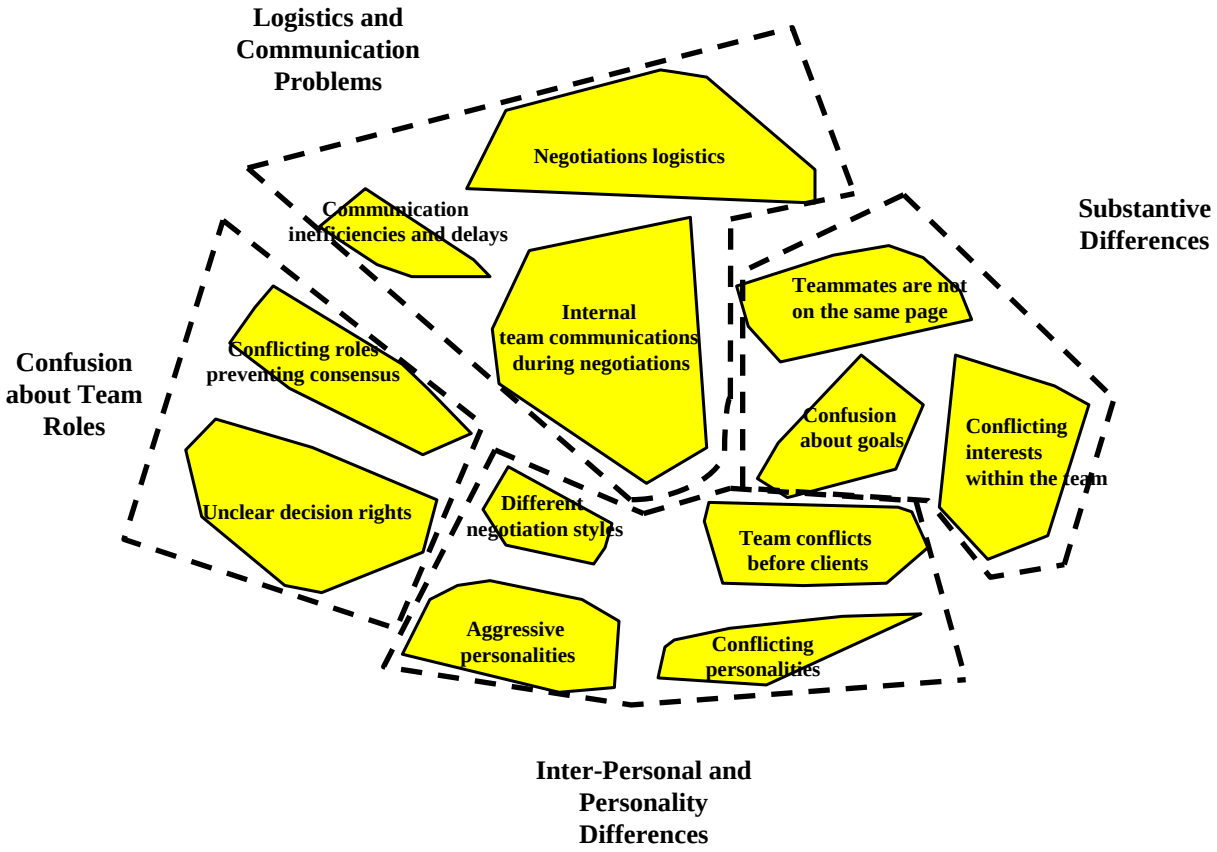
Purpose

- Build theory about trade-offs in decision making about conflict
- Show how the meaning of the code varied across outcomes—i.e., why current coding schemes were misrepresenting meaning

Behfar, Peterson, Mannix, and Trochim (2008)

Extending the Core Analysis, III

predicting behavior and attitudes from cluster matches and Likert-ratings



Purpose

- To theoretically ground concepts in a new context (pattern coding)
- To demonstrate interdependencies between issues
- To predict attitudes based on interdependencies

Behfar, Friedman, & Brett (2007)

Practical Issues

Advantages

- Good fit for the nature of open ended responses
- Does not rely on preconceived coding schemes
- Preserves the context of the concept in the unit of analysis
- Considerable time savings and expandable analysis

Challenges

- Access to units of analysis
- Sorter burden
- Complex text

  e.g. Statements with direction or conditions
- Dense text