Social Networks 54 (2018) 162–167

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Network visualization and problem-solving support: A cognitive fit study

Brooke Foucault Welles[a,*], Weiai Xu[b]

[a] Northeastern University, 360 Huntington Avenue, 204 Lake Hall, Boston, MA, 02115, USA
[b] University of Massachusetts at Amherst, N308 Integrative Learning Center, 650 N. Pleasant Street, Amherst, MA, 01003-1100, USA

ARTICLE INFO

Article history:
Received 2 May 2017
Received in revised form
22 November 2017
Accepted 5 January 2018

Keywords:
Network visualization
Problem-solving
Cognitive fit
Data visualization

ABSTRACT

This study examines the relative effectiveness of four different social network representations for improving human problem-solving accuracy and speed: node-link diagrams, adjacency matrices, tables, and text. Results suggest that visual network representations improve problem-solving accuracy and speed, compared with text. Among the visual representations, tables produced superior problem-solving outcomes for symbolic tasks and link-node diagrams produced superior problem-solving outcomes for spatial tasks. These results partially support a cognitive fit model of problem-solving support. There is not "one best way" to represent network data. Instead, it is important to match network representations and problem-solving tasks.

Introduction

Network visualizations are used to present complex information to researchers, practitioners, and – increasingly – the general public (Foucault Welles and Meirelles, 2015). These network visualizations could, theoretically, enhance a whole host of problem-solving tasks that people face, from the mundane, such as whom to invite to a dinner party, to the consequential, such as whom to reach out to for help in a crisis. However, very little is known about how people interpret network visualizations, nor whether or when network visualizations enhance human problem-solving quality and/or efficiency. In this study, we compare the relative effectiveness of four different network representations – node-link diagrams, adjacency matrices, tables, and text descriptions – for aiding human problem-solving. Drawing on the theoretical framework of *cognitive fit,* we test whether matching network representations to problem-solving tasks enhances problem-solving quality when people consider relational information.

Network visualizations

There are a number of ways networks can be represented including, node-link diagrams, matrices, or summary tables. Here, we focus on a specific subset of *social networks,* where network nodes are people connected by various types of relationships (friends, family, co-workers, classmates, etc.). These networks are endemic to everyday life, and are often implicated in problem-solving tasks. For instance, if someone is looking for a job, they might interact with a picture of their LinkedIn network to identify contacts in target companies or industries. Or, a military analyst might consult a network matrix to determine if a detainee is affiliated with a terrorist organization.

There is a large mathematical literature on how to visualize social networks (e.g. Battista et al., 1998), including studies how network layouts or aesthetics can facilitate accurate interpretations of network graphs (e.g. Dunne and Shneiderman, 2009; Purchase, 1997; Purchase, 1998). However, few studies examine how visualization choices influence interpretations of the context or relational qualities that network visualizations are designed to represent. A series of studies by McGrath and colleagues (Blythe et al., 1996; McGrath et al., 1996, 1997) are among the limited few that explore how variations in network graphs influence inferences about the social contexts graphs are meant to illustrate. For example, McGrath and colleagues (1997) investigated how various graph layouts depicting the same professional network influenced students' inferences about social grouping and the importance of various actors within the network. Even among these relatively knowledgeable students who just finished a graduate course in organizational theory that emphasized network structure, manipulating graph layouts significantly influenced the students' inferences about the importance and connectivity of individual actors and the graph as a whole. Although McGrath and colleagues did not test resulting effects on problem-solving, it follows that any judgments stem-

* Corresponding author.
E-mail addresses: b.welles@northeastern.edu (B. Foucault Welles), weiaixu@umass.edu (W. Xu).

https://doi.org/10.1016/j.socnet.2018.01.005
0378-8733/

ming from different inferences drawn from various graph layouts would likewise be influenced.

Cognitive fit theory

Despite evidence that graph layout may influence solutions derived from relational data, we know little about precisely how various graph visualizations influence problem-solving outcomes, and less still about how alternate presentations of relational data (such as text or matrices, both common representations of relational data) compare to network graphs as problem-solving aids. The present study aims to explore these differences, drawing specifically on Paivio's (1991) theory of *dual coding*, which suggests that verbal (words) and non-verbal (images) information are understood through distinct, albeit interrelated, processes. A large body of work has explored the consequences of the words/images divide, much of which focuses on the problem-solving consequences of data represented in tables (words) or graphs (images), with few consistent results (Vessey, 1991).

Frustrated by the lack of consistency in studies of information representation and human problem solving, Vessey (1991) proposed the theory of *cognitive fit,* which suggests that neither words nor images are *inherently* better at facilitating problem solving, but rather a good *fit* between the problem-solving task and the information representation is most important that is, according to cognitive fit theory, problem solving is enhanced when the *problem-solving task* and the *problem representation* are aligned. Focusing on comparisons of graphical versus tabular information representations, Vessey argues that data can usually be represented by either a graph or a table, each of which emphasizes different informational features,

> "Graphs are spatial problem representations since they present spatially related information. ... they emphasize information about relationships in the data. Tables are symbolic data representations since they present information that is symbolic in nature. Tables represent discrete data values," (Vessey, 1991, p. 225)

Zhu and Watts (2010) extended this distinction between graphical and tabular information to include network representations, arguing that network graphs (node-link graphs) are consistent with Vessey's definition of graphical (spatial) information representations, while matrices and table representations of relational data are consistent with tabular (symbolic) information representations.

Classifying information representations as either graphical or tabular (spatial or symbolic, respectively) is useful, not because one representation is inherently better than the other, but because problem-solving tasks can similarly be classified as spatial or symbolic. *Spatial tasks* require observers to make inferences about relationships within data, making comparisons, observing trends, and/or detecting patterns or deviations from patterns. In contrast, *symbolic tasks* require observers to extract specific values from the data, either through direct observation or carrying out computations on discrete values observed within the data.

Cognitive fit expects that problem-solving often includes both spatial and symbolic tasks, and predicts that problem solving will be optimized when problem representations are matched to problem-solving tasks (Vessey, 1991). That is, problems that require judgments about the relationships or patterns within data will be solved quickest and most accurately when data are represented in graphs, while problems that require judgments about discrete values within data will be solved quickest and most accurately when data are represented in tables. Cognitive fit theory is illustrated in Fig. 1, a diagram depicting the problem-solving process:

Extending to Zhu and Watts' (2010) work on network information representations, cognitive fit theory suggests that node-link graphs may be most effective for problems involving an explicit assessment of relationships, while data tables, may be more effective for problems that involve extracting discrete values from networked data.

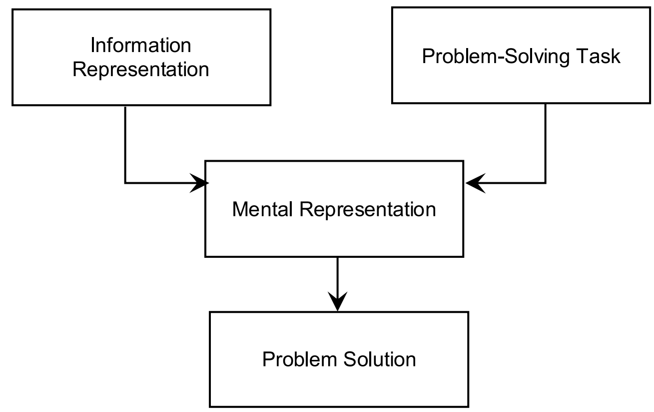

Fig. 1. Illustration of Vessey's (1991) Cognitive Fit Theory.

The present study aims to empirically test cognitive fit theory using two network data representations: node-link graphs, tables, adjacency matrices, and text, and two types of problem-solving tasks: spatial and symbolic. Based on the theory of cognitive fit, we expect the following:

H1. For symbolic tasks, problem solving speed and accuracy will be superior when networks are represented as data tables.

H2. For spatial tasks, problem solving speed and accuracy will be superior when networks are represented as node-link graphs.

In addition, because prior research neither explores the effectiveness of two other common relational data representations – adjacency matrices and text – nor classifies them as spatial or symbolic, we propose the following research questions:

RQ1: How effective are adjacency matrices for supporting spatial and symbolic problem-solving tasks?

RQ2: How effective are text descriptions for supporting spatial and symbolic problem-solving tasks?

Methods

Platform

To test the proposed hypotheses and research question, we designed an online experiment called NetIQ. The NetIQ experiment was hosted on *Volunteer Science* (volunteerscience.com). *Volunteer Science* is an integrated website and development platform for deploying web-based single-user and multi-user behavioral experiments. *Volunteer Science* is a scalable, open, and flexible platform for conducting online experiments (Radford et al., 2015). It has been validated as scientifically and technically accurate for tasks similar to the NetIQ study (Radford et al., 2016). The site is built on top of free open source development tools, including Django, Bootstrap, HTML5, and JavaScript, enabling it to work on any device with modern web-browsing technologies. It uses Amazon Web Services to provide high-performance and scalable on-demand hosting, and seamlessly ties into Amazon's Mechanical Turk for participant recruitment and compensation.

Participants and procedure

We recruited 223 unique participants via Mechanical Turk and directed them to the Volunteer Science website where they consented to participate in research prior to beginning any research