Model-based Analysis of Concept Maps

Sam K. Hui[*], Yanliu Huang[†] and Edward I. George[‡]

Abstract. A concept map is a data collection tool developed in psychology and education to obtain information about mental representations of concept associations. This methodology has recently been introduced to marketing to study consumers' brand perceptions (John et al. (23); Joiner (24)) and attitudes towards health risk (e.g., Huang et al. (21)). In conjunction with other more established methods (e.g., Multidimensional scaling), concept maps provide an additional valuable tool for researchers to understand consumers' structural knowledge about different important marketing concepts.

Building on the introduction by John et al. (23), we propose a descriptive probability model of concept map formation, along with concept map analyses based on parameter estimates. In particular, we demonstrate how to test hypotheses about differences between two groups of maps, and how to aggregate across individual concept maps to form a "consensus map." To demonstrate our methodology, we apply our model to a dataset that uses concept maps to study college students' perceptions of Sexually Transmitted Diseases (STDs), an important topic of growing interest in health marketing (e.g., Hill (20); LaTour and Pitts (28); Raghubir and Menon (40); Treise and Weigold (46)). Though parsimonious in nature, our model adequately recovers map-level, concept-level, and link-level summary statistics commonly considered by other researchers, yet rarely modeled directly.

Keywords: concept maps, network analysis, Bayesian hypothesis testing

1 Introduction

Concept maps are data collection tools developed in psychology to understand mental representation (Anderson and Bower (3); Rumelhart et al. (41)) of topic knowledge by revealing part of an individual's thinking process (Stuart (43)). Since their introduction, concept maps have widely been used in education as a tool to help learners process information effectively (Anderson (2); Lindsay and Norman (30)), to access people's cognitive structures, and to gauge the effectiveness of interventions (Dansereau et al. (13); Martin (36); Moore and Readence (38); Novak and Gowin (39)). Researchers have also used concept maps in social psychology to study attitudes (Lord et al. (32)), and more recently in marketing as a means to understand consumer brand associations (John et al. (23); Joiner (24)) and health risk perceptions (e.g., Huang et al. (21)). In conjunction with other more established methods (e.g., MDS), concept maps provide a valuable tool for researchers to understand consumers' structural knowledge about

[*]Department of Marketing, The Wharton School, University of Pennsylvania, Philadelphia, PA, mailto:kchui@wharton.upenn.edu

[†]Department of Marketing, The Wharton School, University of Pennsylvania, Philadelphia, PA

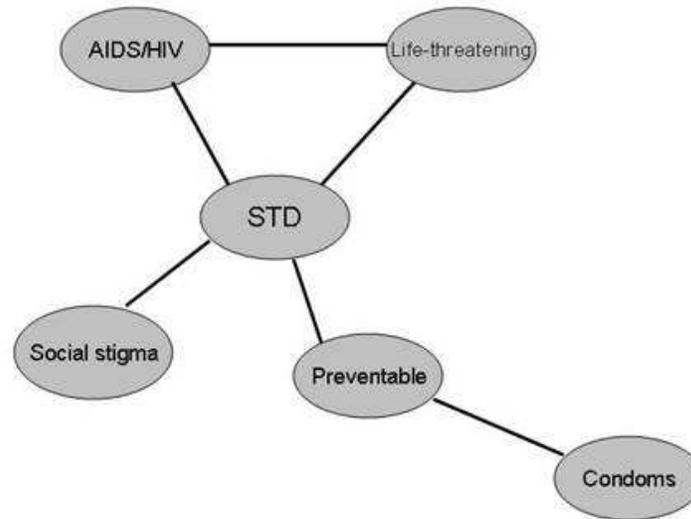[‡]Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA, mailto:edgeorge@wharton.upenn.edu

Figure 1: Example of a concept map.

various important marketing concepts (John et al. (23)).

In a typical "concept mapping" procedure in a marketing context (John et al. (23)), a participant is presented with a "core concept" (e.g., Sexually Transmitted Diseases (STD) as studied here in a health marketing context) and asked to systematically write down all his/her thoughts related to the core concept in a structural manner. Specifically, the participant links the core concept to other concepts (e.g., STD–HIV, STD–Preventable), and subsequently links those concepts to other related concepts that come to mind (e.g., HIV–Life-Threatening, Preventable–Condom). Thus, a concept map, as defined in John et al. (23), is comprised of concepts and undirected links that represent connections between different concepts.[1] Figure 1 presents an example of a concept map that we collected (discussed in Section 4), which describes a particular college student's perception of STDs.

While concept maps provide a valuable tool to study consumer knowledge representation (John et al. (23); Joiner (24); Lord et al. (32)), data in the form such as Figure 1 pose unique challenges for both behavioral and managerial researchers. The key challenge is to extract information from such data to understand how participants organize relevant concepts, and further, how cognitive representations influence people's subsequent attitudes and behaviors (Joiner (24); Lord et al. (32); Martin (36)). Towards this end,

---

[1]In psychology and education, the links in a concept map are sometimes directed, and sometimes labeled with various qualifiers (e.g., Novak and Gowin (39)); in marketing, unlabelled and undirected links are more common. In this paper, we restrict our attention to undirected maps, and return to directed maps in Section 6.

we here propose a probability model for concept maps that allows for the quantification of variation and proper summarization of information across different maps. This model allows us to tackle two particularly important issues. First, we are able to test whether two groups of concept maps (perhaps generated under different experimental conditions or different consumer segments) are different, and if so, in what ways. Second, we are able to summarize information across a group of concept maps by forming a "consensus map" (e.g., John et al. (23); Joiner (24)). Although a formal definition has not been proposed in the literature, a consensus map roughly corresponds to an "average" perception across a set of concept maps. Presumably, a consensus map will provide researchers with a general picture of how consumers, on "average," perceive a brand/concept.

Previous literature have tended to analyze concept maps heuristically or algorithmically (e.g., Chang et al. (11); John et al. (23); Joiner (24); Goldsmith et al. (19); Liu et al. (31); Marshall et al. (35); Takeya et al. (44)), by first extracting a set of summary statistics from the data, and then basing analyses on those statistics. Since different studies generally use different summary statistics and heuristics, hypothesis testing results and consensus maps often depend on the particular choice of heuristics and algorithms. To the best of our knowledge, there has not been any attempt to unify these different heuristics with a consistent statistical methodology that is reasonable with respect to a well-defined probability model. The goal of this paper is to fill this important gap. Specifically, we develop a rich yet parsimonious probability model that provides a unified modeling framework for the analyses of concept maps (e.g., hypothesis testing and consensus map construction).

Our model is based on an elaboration of the uniform graph model (Erdos and Renyi (14)), common in computer science and statistics. Specifically, our model extends the uniform graph model in two directions: (i) allows for non-uniform link-formation probabilities and (ii) introduces a post-realization "pruning" step, discussed in more detail in Section 2. For this model, we propose an estimation procedure, and show that the summary statistics considered in the previous literature (e.g., frequencies of link occurrences) are neither sufficient nor unbiased for our model parameters. In contrast, our parameter estimators adequately summarize the information contained in the data and allow for the recovery of many of the ad hoc summary statistics that have been previously considered. Indeed, analyses based on our model will be adequate to ensure consistency across studies. Furthermore, we demonstrate how consensus maps can be constructed based on model parameter estimates derived from concept map data. To demonstrate our methodology in the marketing context, we apply our model to a dataset that uses concept maps to study college students' perceptions of Sexually Transmitted Diseases (STDs), an important issue in the health marketing literature (e.g., Hill (20); LaTour and Pitts (28); Raghubir and Menon (40); Treise and Weigold (46)). We focused on college students' STD perception because the prevalence of STD on college campuses is steadily increasing (e.g., Mahoney et al. (34); MacDonald et al. (33)) and the data on how college students perceive STD might help design effective educational programs to improve the safe sex and decrease the STD risk. This is consistent with the general goals of health marketing to create, communicate, and deliver health information and

interventions to protect and promote the health of diverse populations, according to the Centers for Disease Control and Prevention.

The remainder of this paper is organized as follows. In Section 2, we develop our statistical model of concept maps and describe our proposed procedures for hypothesis testing and consensus map formation. Section 3 describes a simulation experiment that ensures that our estimation procedure is able to recover the model parameters. Section 4 outlines our data collection procedure and provides an overview of our dataset. In Section 5 we apply our model to actual data, perform model validation, and present parameter estimates. Finally, Section 6 concludes with directions for future research.

2   Model

We describe our model in detail in this section. Section 2.1 defines our notation. Section 2.2 begins with the simple uniform model of graph formation (Erdos and Renyi ([14])), and argues that it is inadequate for the description of concept maps; we then introduce two elaborations to develop an adequate model of concept maps. Specifically, we allow for non-uniform probabilities of edge formation between each pair of concepts, and introduce a (latent) "pruning" step to ensure that our model generates maps that are fully connected, consistent with the connectedness property of concept maps. Section 2.3 describes an empirical Bayes estimation procedure (Carlin and Louis ([8])), Section 2.4 extends this to empirical Bayes hypothesis testing, and Section 2.5 proposes a method of aggregation across concept maps to form a "consensus map" using the estimated model parameters.

2.1   Definitions and notations

The two hypothetical concept maps shown on the left in Figure 2 illustrate our definitions and notation. As seen in the figure, a concept map is comprised of two inter-related components: concepts (labeled by the oval shapes), and links between concept pairs (denoted by lines connecting the two corresponding ovals). The concept at the center (concept 1 in Figure 2) denotes the "core concept" from which the entire map is generated. The existence of a link between a concept pair indicates a conceptual linkage between them in the mind of the respondent.

Formally, a concept map that is used in marketing, as defined by John et al. ([23]), is an *undirected* and *connected* mathematical graph.[2] Its undirected links indicate which concept pairs are connected, but do not inform us about the direction of such linkages.[3] It is connected because there exists a path (a series of consecutive links) that connects it

---

[2]In general, a concept map used in psychology/education does not have to be fully connected; the ones collected in marketing are fully connected due to the mechanism of the data collection (participants are asked to think in terms of associations (i.e., links) from the core concept, thus every concept has to be connected in some way to the core concept).

[3]While some extensions of the concept map methodology are available (e.g., Joiner ([24])) that record the order in which concepts are formed, we focus on undirected concept maps as defined in John et al. ([23]) throughout this paper.

Concepts (X_{ij})		Links (y_{ijk})					
			1	2	3	4	5
X_{i1}	1	1	0	0	0	1	1
X_{i2}	0	2	0	0	0	0	0
X_{i3}	1	3	0	0	0	1	0
X_{i4}	1	4	1	0	1	0	0
X_{i5}	1	5	1	0	0	0	0

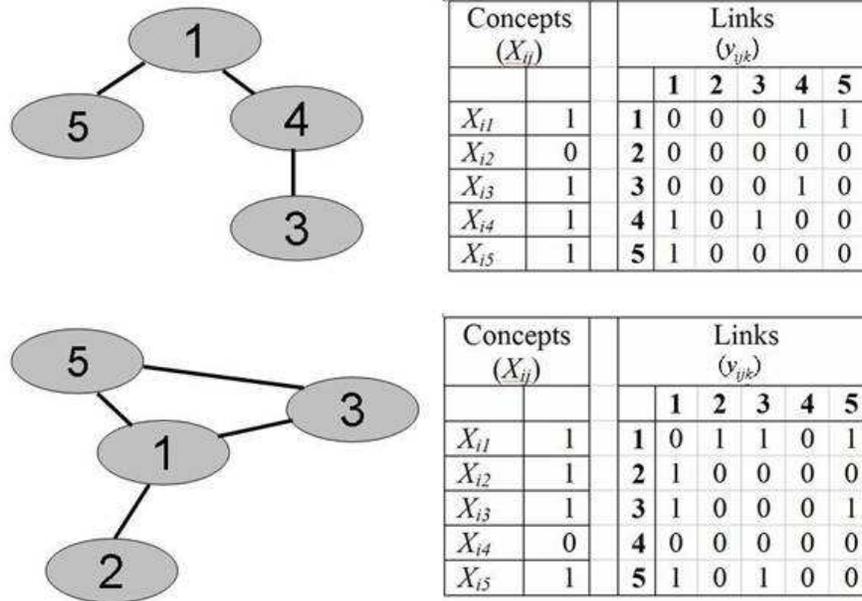| Concepts (X_{ij})		Links (y_{ijk})					
			1	2	3	4	5
X_{i1}	1	1	0	1	1	0	1
X_{i2}	1	2	1	0	0	0	0
X_{i3}	1	3	1	0	0	0	1
X_{i4}	0	4	0	0	0	0	0
X_{i5}	1	5	1	0	1	0	0

Figure 2: Examples of concept maps to illustrate our notations.

back to the core concept. Thus, no concept can be disconnected from the core concept.

Throughout this paper, we use i ($i = 1, 2, \ldots, I$) to index people, and j and k ($j = 1, 2, \ldots, N; k = 1, 2, \ldots, N$) to index concepts. We use I for the total number of people, and N for the total number of observed (unique) concepts.[4] We define X_{ij} as an indicator that equals 1 if concept j appears in the i-th person's concept map, and 0 otherwise. In addition, we define y_{ijk} as an indicator that equals 1 if there is a link between concept j and k in the i-th person's map, and 0 otherwise. Since concept maps are undirected (based on our definition), y_{ijk} is symmetric, i.e., $y_{ijk} = y_{ikj}$. Clearly, any graph (with N nodes) can be represented by an $N \times N$ matrix of 0-1 entries which identify the presence/absence of links (Bollobas (7)). Thus, our representation of a concept map through y_{ijk} (and X_{ij}) is completely general. As an example of our notational representation, the tables on the right in Figure 2 identify the two concept maps on the left.

2.2   A probability model of concept map

Our goal in this research is to develop a descriptive probability model of concept map that is parsimonious, yet able to recover key summary statistics that describe the data

---

[4]In this paper, we assume that N is fixed, and is equal to the total number of unique concepts that appear in the dataset. Alternatively, one could think of N as a random variable and infer it from the data. We leave this issue for future research.
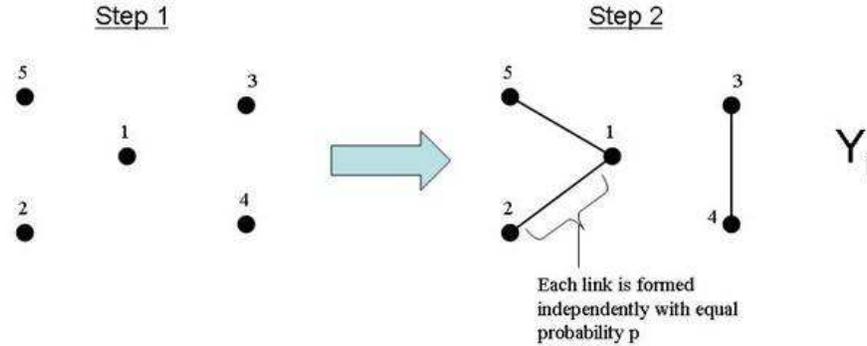
Figure 3: Illustration of the uniform graph model (Erdos and Renyi (14)).

(as shown in Section 3). To this end, we start with the uniform graph model (Erdos and Renyi (14)), a parsimonious model of graph formation, and introduce two elaborations that account for the unique characteristics of concept maps, which lead to our proposed $\tilde{G}(N, \Theta)$ model.

The uniform graph model

The uniform graph model has been widely applied in the theory of random graphs (Bollobas (7)). Due to its simplicity, the application of the uniform graph model is usually not on providing a flexible description of graph. Rather, it has typically been used to derive the null distributions of graphical characteristics of special interest (e.g., degree distribution, clustering coefficient, average path length; see, for instance, Albert and Barabasi (1)), or as a "test bed" for specialized graph algorithms (e.g., Levy et al. (29)).

Formally, the uniform graph model is denoted $G(N, p)$, where N denotes the number of nodes (concepts) and p denotes the probability of edge (link) presence. It corresponds to the following two-step stochastic process governing the generation of graphs (Bollobas (7)), as depicted graphically in Figure 3 (the corresponding concept map terminology is placed in parentheses):

Step 1 Start with N nodes (concepts) and position them on an observed graph Y_i.

Step 2 On Y_i, add edges (links) between pairs of nodes (links) at random. Each possible edge (link) is included independently with probability p.

The above stochastic process gives rise to the following likelihood function of the i-th map Y_i (with links denoted by y_{ijk}):

$$f(Y_i | p, N) = p^{\sum_{j<k} y_{ijk}} (1-p)^{\frac{N(N-1)}{2} - \sum_{j<k} y_{ijk}} \tag{1}$$

The uniform graph model, as defined by the likelihood function in Equation [1], is an inadequate descriptive model of concept maps for three reasons. First, the equal probability assumption for link occurrence (in Step 2) between any concept pair is too restrictive because links will typically occur with very different frequencies, for example in the data we discuss in Section 4. Second, in the concept map setting, Step 1 assumes that all concepts appear in every map, which will generally not be true. Typically, a person mentions only a small subset of all the possible concept, as is illustrated in our data. Third, the graph generated under $G(N, p)$ may have nodes (concepts) that are disconnected from the designated core concept, i.e., not reachable from the core concepts by following the links, as shown in the right panel of Figure 3. This violates the connectedness property of concept maps.

Proposed model: the $\tilde{G}(N, \Theta)$ model

Our proposed $\tilde{G}(N, \Theta)$ model is a three-stage, partially observed stochastic process that addresses all of the three issues above. Our overall approach, as shown graphically in Figure 4, can be summarized as follows:

Step 1' Start with N nodes (concepts) and position them on a *latent* graph Z_i.

Step 2' On the latent graph Z_i, add edges (links) between pairs of nodes (links) at random. The edge (link) between concept j and concept k is included with probability θ_{jk}.

Step 3' Finally, the latent graph Z_i is "pruned" (defined in Equation [3] later) to arrive at the *observed* map Y_i. That is, any concepts in Z_i not connected in some way to the core concept are removed to obtain Y_i.

Note that the stochastic process that underlies our model is only partially observed. That is, in contrast to the uniform graph model, we assume that the first two steps in our model (Step 1' and Step 2') take place in latent space, which then induce an observed map Y_i in Step 3'.

The first step (Step 1') of our model is the same as that of the uniform graph model, where the only difference is that we begin by creating a *latent* graph Z_i. In Step 2', we extend the uniform graph model by allowing each link to have a different occurrence probability, therefore replacing the parameter p by a (symmetric) matrix $\Theta_{N \times N}$, where θ_{jk} denotes the probability that a link is present between concepts j and k.[5] Since concept maps are undirected and no concept is linked to itself, $\Theta_{N \times N}$ will be symmetric ($\theta_{jk} = \theta_{kj}$ for all j and k) with diagonal entries equal to zero. Formally, we have:

$$P(z_{ijk} = 1) = \theta_{jk} \tag{2}$$

---

[5]While it is possible to make $\Theta_{N \times N}$ an individual specific parameter, we stay with population-level parameters due to identification issues which we discuss later. The possibility of specifying individual-level parameters are discussed in more detail in Section 6.
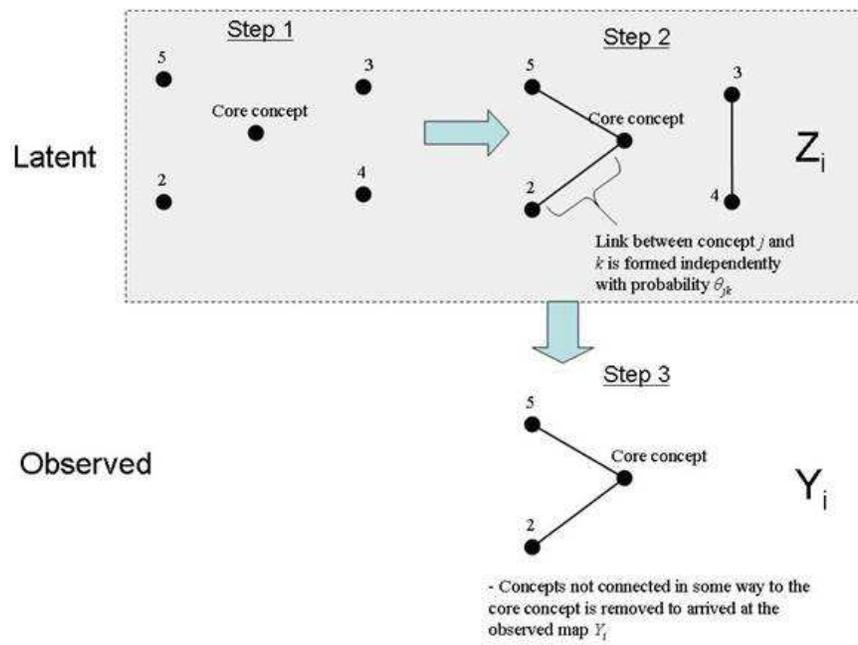
Figure 4: Illustration of our proposed model of concept map.

where z_{ijk} is an indicator variable that represents the presence/absence of an edge between concepts j and k in the latent map Z_i.

After Step 1' and 2', the latent map Z_i may not be fully connected. To ensure consistency with the connectedness property of concept maps (as we defined earlier), we introduce Step 3', a latent "pruning" step where any concepts in Z_i that are not connected in some way to the core concept are removed, to arrive at the *observed* map Y_i. Figure 4 illustrates our pruning process with an example. The structure of the (latent) graph Z_i before pruning is shown in the upper right panel of Figure 4. Since concepts 3 and 4 are not connected to the core concept (i.e., it is impossible to reach them from the core concept), they are removed ("pruned") from the latent graph, thus resulting in the pruned graph (which correspond to the *observed* concept map Y_i) in the lower panel. Although concepts 3 and 4 are connected among themselves, they are not related to the core concept; thus, when asked to elicit her preference, the participant may not put concepts 3 and 4 on her concept map, resulting in the observed concept map in the lower panel of Figure 4. Mathematically, we denote the process pruning by $f_{prune}(\cdot)$, a (deterministic) function that operates on latent map Z_i to get the observed concept map Y_i. That is,

$$Y_i = f_{prune}(Z_i) \tag{3}$$

where the jk-th element of Y_i, namely y_{ijk}, equals 1 if both $z_{ijk} = 1$ and concept j or concept k are connected to the core concept in map i; otherwise, $y_{ijk} = 0$.

Together, our model, which we refer to as $\tilde{G}(N, \Theta)$ (where Θ represents the distinct linkage probabilities and the tilda ($\sim$) indicates pruning), leads to the following likelihood function for map Y_i:

$$f(Y_i | \Theta, N) = \big(\prod_{j<k} \theta_{jk}^{y_{ijk}} (1 - \theta_{jk})^{1-y_{ijk}} \big) H_i \tag{4}$$

where H_i is an indicator variable which takes value 1 if Y_i is fully connected, and 0 otherwise.

Before moving on to describe our inferential procedures, we note three potential limitations of our model $\tilde{G}(N, \Theta)$. First, each participant's map is assumed to be drawn from the same set of population-level parameters Θ, so that heterogeneity across individual parameters is not accommodated. Since each person produces only a single map, modeling such parameter heterogeneity across participants would require individual-level covariate information, an issue that we return to in Section 6. Second, $\tilde{G}(N, \Theta)$ is a model of the final concept map, rather than a sequential model which describes the process of concept map formation. As we mentioned earlier, most concept map applications collect data in the form of undirected graphs such as in Figure 1, where the direction of links is not specified. Our model is perhaps better suited for such data, whereas a model for the process of map formation would require directed link data. Third, our model treats all link formations as mutually independent, conditional on Θ. Although

one could extend our model by introducing more parameters to allow for dependence between link formations, our framework provides a surprisingly adequate description of concept maps, allowing for the recovery of many summary statistics, as will be seen in Section 5. Future directions for incorporating correlations into our model may include approaches developed in testlet theory (Wainer et al. (47)).

2.3   Estimation of $\Theta_{N \times N}$

If the latent map Z_i is observed, making inference about $\Theta_{N \times N}$ should be straightforward. However, in reality, inference about our model parameters $\Theta_{N \times N}$ is complicated by the fact that we only observe Y_i, a pruned version of the (latent) map Z_i; some of the links present in Z_i may be unobserved in Y_i because of pruning, essentially a form of censoring. Thus, inference based on Y_i must contend with such censoring, a common issue in other observational studies, particularly survival analysis (Klein and Moeschberger (27)). Adjustments similar to those made when estimating models based on censored data must be made when estimating $\Theta_{N \times N}$.

Previous analyses of concept maps have been typically based on ad hoc summary statistics, such as the percentage of times a link occurs across all maps. In particular, a key summary statistic previous researchers (e.g., Joiner (24)) have relied on is

$$\tilde{\theta}_{jk} = \frac{m_{jk}}{I} = \frac{1}{I} \sum_{i=I}^{I} y_{ijk} \tag{5}$$

where m_{jk} denotes the total number of concept maps containing a link between concept j and k, I is the total number of people in the sample.

Statistical inference based on the $\tilde{\theta}_{jk}$ in Equation [5] seems undesirable for two reasons. First, $\tilde{\theta}_{jk}$ is a biased estimator which tends to underestimate θ_{jk} , i.e., $E(\tilde{\theta}_{jk}) < \theta_{jk}$ because Equation [5] does not compensate for censoring; see the Appendix for an outline of a proof with ideas similar to those in Klein and Moeschberger (27). Second, and perhaps more importantly, $\tilde{\theta}_{jk}$ is not a sufficient statistic (Fisher (15)) under our model. Thus, any statistical procedure based solely on $\tilde{\theta}_{jk}$ will typically be inadmissible in a statistical decision theory context (Berger (4)). In Section 5, we show that when our model is properly calibrated, the statistic $\tilde{\theta}_{jk}$ and others like it (described later) can be recovered from a *realization* from our model. Indeed, our model parameter estimates incorporate the information contained in those summary statistics and more.

As opposed to $\tilde{\theta}_{jk}$, the Maximum Likelihood Estimate (MLE) θ_{jk}^{MLE} under our model adjusts for the pruning or censoring step. This MLE can be written as

$$\theta_{jk}^{MLE} = \frac{m_{jk}}{n_{j|k}} = \frac{\sum_{i=1}^{I} y_{ijk}}{\sum_{i=1}^{I} 1_{(X_{ij}=1 \ \ OR \ \ X_{ik}=1)}} \tag{6}$$

where $n_{j|k}$ denotes the number of times that either concept j or concept k is observed

(see the Appendix for a derivation). Note that the MLE is based on the ratio quantities $\frac{m_{jk}}{n_{j|k}}$ instead of the ratio quantities $\frac{m_{jk}}{I}$ in Equation [5]. Looking at $n_{j|k}$ rather than I is similar in spirit to the adjustments made in the analysis of right-censored data in survival analysis (e.g., Klein and Moeschberger (27)), which focuses only on the number of observed cases.

Beyond this MLE, inference about $\Theta_{N \times N}$ can be further enhanced by treating the elements of $\Theta_{N \times N}$ as realizations from a hyperdistribution, in effect embedding $\tilde{G}(N, \Theta)$ within a hierarchical model. This elaboration allows us to conveniently borrow strength across the links by using empirical Bayes estimation (Carlin and Louis (8)). In particular, we consider the simple case where the parameters θ_{jk} are assumed to be i.i.d. draws from a common Beta hyperdistribution, i.e.,

$$\theta_{jk} \sim Beta(a, b) \quad (j = 1, 2, \ldots, N; k > j) \tag{7}$$

In this empirical Bayes framework, the hyperparameters a and b are assumed to be unknown fixed constants that we estimate by maximizing the marginal likelihood of a and b based on the data. Since the probability of observing m_{jk} (conditional on θ_{jk}) is Binomial($n_{j|k}, \theta_{jk}$), the marginal distribution of the data is Beta-Binomial. Thus, a and b can be easily estimated using routine non-linear maximization methods. [6] Once we have estimated a and b, the posterior distribution of θ_{jk} (given the data and the hyperparameters) is given by:

$$\theta_{jk} | \text{data}, \widehat{a}, \widehat{b} \sim Beta(m_{jk} + \widehat{a}, \ n_{j|k} - m_{jk} + \widehat{b}) \tag{8}$$

where $\widehat{a}$ and $\widehat{b}$ denote the maximum marginal likelihood estimates of the hyperparameters a and b, respectively. Thus, the posterior mean (i.e., the empirical Bayes estimator) for θ_{jk} is given by

$$\widehat{\theta}_{jk} = \frac{m_{jk} + \widehat{a}}{n_{j|k} + \widehat{a} + \widehat{b}} = \frac{\sum_i y_{ijk} + \widehat{a}}{\sum_i 1_{(X_{ij}=1 \ OR \ X_{ik}=1)} + \widehat{a} + \widehat{b}} \tag{9}$$

and the corresponding $(1\text{-}\alpha)\%$ posterior interval estimate [7] is given by

$$\left(B(\frac{\alpha}{2}; m_{jk} + \widehat{a}, n_{j|k} - m_{jk} + \widehat{b}), B(1 - \frac{\alpha}{2}; m_{jk} + \widehat{a}, n_{j|k} - m_{jk} + \widehat{b}) \right) \tag{10}$$

---

[6]We note that in some cases, the MLE for a and b may not exist or may be spurious (Takeya and Young (45)). Thus, we recommend that researchers first examine a 3D-plot of the likelihood function before applying non-linear maximization methods.

[7]One can also report HPD intervals and/or make adjustments for the uncertainty of the hyperparameters (a,b), e.g., using methods described in Carlin and Louis (8). We find that the intervals generated using Equation [10], the HPD intervals, and the fully Bayes intervals are very similar. Details are available upon request.

where $B(\alpha; r_1, r_2)$ denotes the α quantile of a Beta(r_1, r_2) distribution. In Section 3, we further demonstrate the performance of our estimation procedure by conducting a simulation experiment.

It should be mentioned that these empirical Bayes estimates may be regarded as an approximation to a fully Bayes analysis. For example, we also considered a fully-Bayes analysis by putting noninfluential diffuse priors on the hyperparameters a and b (e.g., $\log(a) \sim N(0, 1000^2)$ and $\log(b) \sim N(0, 1000^2)$,[8] and then using MCMC sampling to compute posterior quantities of interest. Such a fully Bayes approach may be preferable because it properly accounts for hyperparameter uncertainty, for example by yielding slightly larger posterior interval estimates. However, the fully Bayes results we obtained were very similar to the empirical Bayes results presented here. Furthermore, the empirical Bayes closed-form solutions above will allow other researchers to use our method within their studies more easily.

Once our model is calibrated, concept map analyses can be based completely on the parameter estimates. Use of the same unifying model in this way ensures that all such analyses will be conceptually consistent. In the two subsections below, we demonstrate how to test for differences between groups of concept maps based on our model, and how to generate consensus maps that summarize information across individual concept maps.

2.4   Hypothesis testing

In some cases, researchers may want to test for differences between two groups of participants with different demographic characteristics (e.g., gender, age group) or whose maps were constructed under different experimental conditions. Under our Bayesian framework, we do so by comparing the hypothesis H_0 that there is no difference between the groups, with the hypothesis H_1 that the groups are different. Formally, we specify the two hypotheses as:

$$H_0 : \theta_{jk}^{(A)} = \theta_{jk}^{(B)} = \theta_{jk} \qquad \forall j, k$$

$$H_1 : \theta_{jk}^{(A)} \neq \theta_{jk}^{(B)} \; \; for \; some \; j, k$$

where A and B denote each of the two groups, respectively.

The relevant Bayesian post-data quantity for comparing H_0 and H_1 is the posterior odds $\frac{P(H_1|\text{data})}{P(H_0|\text{data})}$ which by Bayes rule is the product of the Bayes Factor

$$BF_{H_1:H_0} = \frac{P(\text{data}|H_1)}{P(\text{data}|H_0)} \tag{11}$$

---

[8]Alternatively, one can put a reference prior on $\phi = \frac{a}{a+b}$ and $\omega = a+b$, where $p(\phi, \omega)$ is proportional to $\phi^{(-0.5)}(1 - \phi)^{(-0.5)}\omega^{(-0.5)}(1 + \omega)^{(-1)}$.

and the prior odds $\frac{P(H_1)}{P(H_0)}$. Thus, the Bayes factor $BF_{H_1:H_0}$ is the ratio of the posterior odds to the prior odds, and is usually interpreted as the evidence given by the data in favor of H_1 over H_0. When $BF_{H_1:H_0} > 1$, the data indicate some support for H_1 ; and when $BF_{H_1:H_0} < 1$, the data indicate some support for H_0. When $BF_{H_1:H_0} > 3$, Jeffreys (22) suggested we conclude that the data offer "substantial" support for H_1 over H_0. Note that when the two hypotheses are deemed to be a priori equally probable, i.e., the prior odds $\frac{P(H_1)}{P(H_0)} = 1$, the Bayes factor $BF_{H_1:H_0}$ is equal to the posterior odds. As a result, Bayesian testing of H_1 versus H_0 is typically carried out by simply reporting $BF_{H_1:H_0}$.

From Equation [11], we can compute the Bayes factor $BF_{H_1:H_0}$ in our setting as simply the ratio of the marginal densities of the data under the two hypotheses. Since our model uses conjugate Beta priors (Equation [7]), we can calculate the Bayes factor in closed form as follows ($m^{(A)}$ denotes the whole collection of $m_{jk}^{(A)}$):

$$P(\text{data}|H_1) = \Big(\int f(m^{(A)}|\theta^{(A)})\pi(\theta^{(A)})d\theta^{(A)} \Big)\Big(\int f(m^{(B)}|\theta^{(B)})\pi(\theta^{(B)})d\theta^{(B)} \Big) \quad (12)$$

$$= \prod_{j<k} BetaBinom(m_{jk}^{(A)}|n_{j|k}^{(A)},\widehat{a}^{(A)},\widehat{b}^{(A)})BetaBinom(m_{jk}^{(B)}|n_{j|k}^{(B)},\widehat{a}^{(B)},\widehat{b}^{(B)}) \quad (13)$$

$$P(\text{data}|H_0) = \int f(m^{(A)},m^{(B)}|\theta)\pi(\theta)d\theta \quad (14)$$

$$= \int f(m^{(B)}|\theta)f(m^{(A)}|\theta)\pi(\theta)d\theta \quad (15)$$

$$= \int f(m^{(B)}|\theta)f(\theta|m^{(A)})d\theta \quad (16)$$

$$= \prod_{j<k} BetaBinom(m_{jk}^{(B)}|n_{j|k}^{(B)},\widehat{a} + m_{jk}^{(A)},\widehat{b} + n_{j|k}^{(A)} - m_{jk}^{(A)}) \quad (17)$$

Note that to go from [14] to [15], we use the property that $m^{(A)}$ and $m^{(B)}$ are conditionally independent given θ. Also, we note that Equation [15] is symmetric with respect to $m^{(A)}$ and $m^{(B)}$; i.e., we can also flip the order of $m^{(A)}$ and $m^{(B)}$ in [15] and hence write [16] as $\int f(m^{(A)}|\theta)f(\theta|m^{(B)})d\theta$.

Using Equation [13] and [17], the Bayes factor is given by:

$$BF_{H_1:H_0} = \frac{\prod_{j<k} BetaBinom(m_{jk}^{(A)}|n_{j|k}^{(A)},\widehat{a}^{(A)},\widehat{b}^{(A)})BetaBinom(m_{jk}^{(B)}|n_{j|k}^{(B)},\widehat{a}^{(B)},\widehat{b}^{(B)})}{\prod_{j<k} BetaBinom(m_{jk}^{(B)}|n_{j|k}^{(B)},\widehat{a} + m_{jk}^{(A)},\widehat{b} + n_{j|k}^{(A)} - m_{jk}^{(A)})}$$
$$(18)$$

We can also take a step further and assess the difference in the parameters governing the formation of each link between group A and B, i.e., to evaluate the following pair of hypotheses for each pair of concepts j and k:

$$H_0(j,k) : \theta_{jk}^{(A)} = \theta_{jk}^{(B)}$$

$$H_1(j,k) : \theta_{jk}^{(A)} \neq \theta_{jk}^{(B)}$$

Similar to the derivation of [18], the corresponding Bayes factor (of $H_1(j,k)$ over $H_0(j,k)$) can be written as:

$$BF_{H_1:H_0}(j,k) = \frac{BetaBinom(m_{jk}^{(A)}|n_{j|k}^{(A)}, \widehat{a}^{(A)}, \widehat{b}^{(A)}) BetaBinom(m_{jk}^{(B)}|n_{j|k}^{(B)}, \widehat{a}^{(B)}, \widehat{b}^{(B)})}{BetaBinom(m_{jk}^{(B)}|n_{j|k}^{(B)}, \widehat{a} + m_{jk}^{(A)}, \widehat{b} + n_{j|k}^{(A)} - m_{jk}^{(A)})} \tag{19}$$

2.5   Consensus map generation

The aggregation of individual concept maps to form a "consensus map" is an important open issue (John et al. (23); Joiner (24)). As briefly mentioned in the introduction, an aggregate consensus map represents (roughly) the "average" perception among the participants. Intuitively, links with higher probabilities of occurrence should be included in the consensus map, while those with lower probabilities of occurrence should be excluded. We formalize this idea with the following procedure. First, we specify a "cutoff" value c, to be used as the criterion to determine whether the posterior mean $\widehat{\theta}_{jk}$ is "large enough" to be included in the consensus map. Second, we define a (latent) intermediate consensus map Z^* , where $z_{jk}^* = 1$ if $\widehat{\theta}_{jk} > c$, and 0 otherwise. Finally, Z^* is pruned to obtain the final consensus map, Y^* , i.e., $Y^* = f_{prune}(Z^*)$.

Note that here c is a tuning parameter to be set by the researcher. A larger value of c will lead to a smaller consensus map with stronger linkages between concepts, while a smaller value of c will lead to a more complicated consensus map. Ideally, the researcher can vary c and observe how the resulting consensus map evolves as c is gradually decreased. Alternatively, the automatic selection of c is briefly discussed in Section 6 as a future research direction.

3   Simulation experiment

Before applying our model to actual data, we conduct a simulation experiment to assess whether our estimation procedure is able to recover the "true" parameter values. Setting N, the total number of concepts to be 5, we simulated the "true" parameters in $\Theta_{N \times N}$ using a Beta(a=1, b=1) distribution; the "true" value of each θ_{jk} simulated is shown in Table 1. With these parameters, we simulated $I = 20$ concept maps under the $\tilde{G}(N, \Theta)$

model that we specified through Step [1'], [2'], and [3'] in Section 2. The simulated maps are shown in Figure 5.

Parameter	"True value"	(20 simulated maps)		(200 simulated maps)	
		Post. mean	95% Post. interval	Post. mean	95% Post. interval
θ_{12}	0.522	0.680	(0.474, 0.856)	0.529	(0.459, 0.596)
θ_{13}	0.308	0.262	(0.104, 0.462)	0.351	(0.286, 0.417)
θ_{14}	0.884	0.911	(0.764, 0.989)	0.899	(0.855, 0.937)
θ_{15}	0.271	0.262	(0.104, 0.462)	0.261	(0.203, 0.324)
θ_{23}	0.044	0.035	(0.000, 0.151)	0.032	(0.012, 0.062)
θ_{24}	0.140	0.080	(0.007, 0.227)	0.152	(0.106, 0.206)
θ_{25}	0.517	0.442	(0.236, 0.660)	0.550	(0.476, 0.622)
θ_{34}	0.472	0.421	(0.222, 0.634)	0.435	(0.366, 0.505)
θ_{35}	0.748	0.885	(0.699, 0.986)	0.751	(0.684, 0.813)
θ_{45}	0.042	0.129	(0.025, 0.300)	0.041	(0.018, 0.072)

Table 1: Parameter recovery in the simulation experiment.

We then applied our estimation procedure to the simulated data. As can be seen in Table 1, all of our point estimates (posterior means) were close to their corresponding true values, and all of our 95% posterior intervals covered the true values of the corresponding parameters. Thus, our estimation procedure is able to adequately recover the model parameters, even when the sample size is reasonably small. We then repeated the above experiment by simulating a much larger dataset of $I = 200$ maps. As can be seen in Table 1, as we would expect, the mean estimates became closer to the true values and the posterior intervals became narrower for this larger I.

4 Data

4.1 Data overview

To demonstrate our approach, we collected concept maps describing perceptions of Sexually Transmitted Diseases (STDs) from a total of 223 undergraduate students at a large northeastern university. The study of how people perceive different diseases (e.g., Chandran and Menon (10); Keller et al. (26); Menon et al. (37); Raghubir and Menon (40)) is an important issue in health marketing, a growing research area that has recently attracted a huge amount of interest among marketing researchers (e.g., Block and Keller (5), Block and Keller (6); Huang et al. (21); Keller (25); LaTour and Pitts (28); Treise and Weigold (46)). Although we focus here on STDs, we should emphasize that our framework and model is completely general and can be applied to a wide variety of settings in marketing, e.g., consumers' perception of other important health-related concepts, people's reactions to marketing communication, or consumer brand associations as studied in John et al. (23) and Joiner (24).
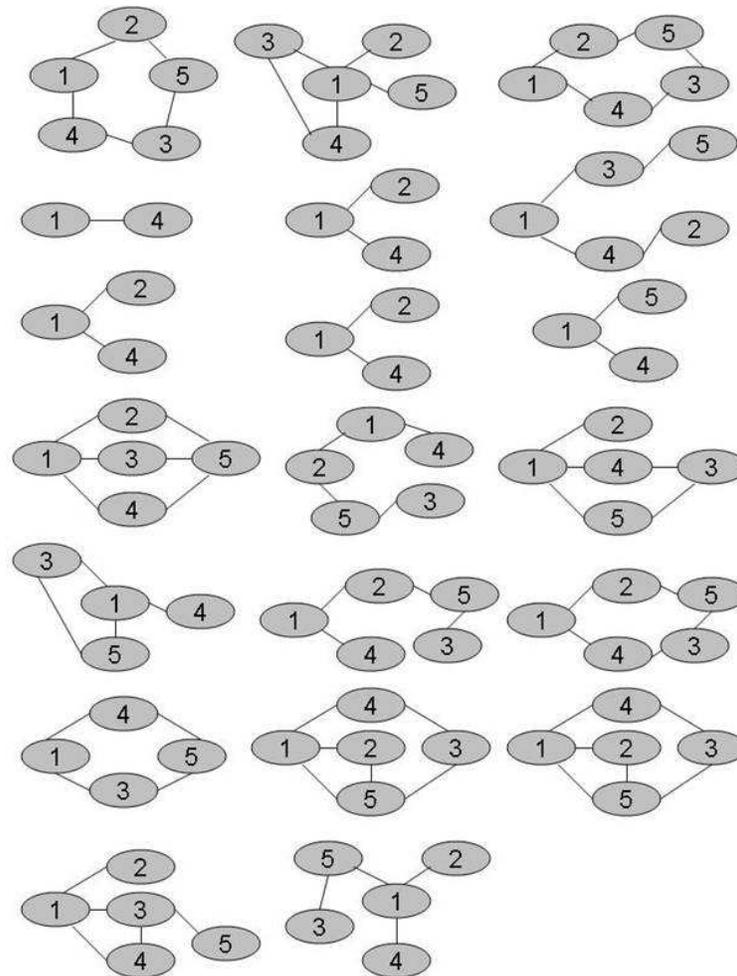
Figure 5: Twenty simulated concept maps generated in the simulation experiment. The core concept is labeled as concept 1.

Participants were told that we wanted to understand their thoughts and how they organize information when they consider the concept of STD. They were given a concept map example and were asked to construct their own maps starting with the core concept (i.e., STD), using a procedure similar to that in John et al. (23). Participants were given as much time as they needed to draw this concept map. Along with the concept map, we also asked each participant to record his/her gender (although we find no significant difference between the genders later in Section 5.3). In our sample, 47% of the participants are male.

After collecting the data, we compiled a list of all the unique concepts listed in any of the individual concept maps. This resulted in a total of 29 unique concepts,[9] listed in Table 2.[10] Interestingly, this number of unique concepts was roughly the same size as that reported in other concept map studies, e.g., John et al. (2006) reported 25 unique concepts when studying brand associations of Mayo Clinic using concept maps.

Concept	Occurence	Concept	Occurence
STD(core concept)	100.0%	Sex	21.5%
AIDS/HIV	54.3%	Fear	20.2%
Different kinds of STDs	54.3%	Social stigma	19.3%
Condoms	46.2%	Everyone can get it	18.8%
Safe sex	39.5%	Disgusting	17.9%
Life-threatening	38.1%	Medical treatment	14.3%
Dangerous	36.3%	Education	14.3%
Physical symptoms	31.8%	Sexual partner	11.2%
Undersirable consequences	30.9%	Prevalent	11.2%
Preventable	27.8%	Bacterial/viral/parasitic	11.2%
Risky lifestyle	26.9%	Get tested	9.9%
Trasmittable	23.8%	Not for me	7.6%
Permanent	23.3%	Disease	5.8%
Unsafe sex	22.4%	Health cost	5.4%
Global problem	22.0%		

Table 2: A list of 29 unique concepts along with their % occurence in decreasing order of occurence (concept-level summary statistics).

---

[9]Similar to the approach in John et al. (23), concepts that are very similar in meaning, e.g., gross/disgusting, are combined. While this process admittedly involve some subjective human judgment, similar procedures are involved in most studies where researchers elicit verbal/textual responses from participants (e.g., Chattopadhyay and Alba (12); Lord et al. (32); Shavitt and Brock (42)).

[10]A very few of our participants included an inappropriately offensive epithet and/or inappropriate personal information with their concepts. We decided that it would be best to simply exclude these epithets from our analysis. Because of their rarity, their inclusion in our analysis would not have had any noticeable effect on our model's performance and results.

4.2   Summary statistics

We extracted three sets of summary statistics from our data: map-level, concept-level, and link-level summary statistics that are also primary measures considered by other concept maps researchers (e.g., John et al. (23); Joiner (24)). In Section 5, we demonstrate that these three sets of summary statistics can be adequately recovered using our model.

Map-level summary statistics

We extract the following map-level summary statistics from each map.

(i) Total number of unique concepts: $\sum_j X_{ij}$, where X_{ij}, as defined earlier, is an indicator variable that equals 1 if i-th map contains the j-th concept, and 0 otherwise.

(ii) Total number of links: $\sum_{k>j} y_{ijk}$.

(iii) Number of first order links: A first order link is defined as a link that is directly connected to the core concept. Thus, the number of first order links (for the i-th map) is $\sum_k y_{i1k}$.

(iv) Number of second (and above) order links: This measures the number of links that are at least two links away from the core concept.

These summary statistics are shown in Table 3. The average number of concepts per map is around 7.66, while the median is 8.00. Across the data, the average number of links per map is about 7.54. The median number of links is 7.00, while the median number of first order links and second (and above) order links is 5.00 and 2.00, respectively.

	Mean	Median	S.D.	Min	Max
Total number of concepts	7.66	8.00	2.03	2.00	14.00
Total number of links	7.54	7.00	2.45	1.00	15.00
Number of 1st order links	4.59	5.00	1.43	1.00	9.00
Number of 2nd and higher order links	2.74	2.00	2.14	0.00	11.00

Table 3: Map-level summary statistics.

Concept-level summary statistics

We record the number of times each concept is present across our data. The % occurrence of each concept, sorted in decreasing order, is shown in Table 2. The concepts mentioned most frequently are "STD" (the core concept that must appear by default), "AIDS/HIV," "Different kinds of STDs," and "Condoms." These concepts appear in more than 40% of the concept maps in our data.

Link-level summary statistics

Finally, we record the percentage of times each link occurred across all the maps. The top 20 links that are observed most frequently in the data are shown in Table 4. While this list is, unsurprisingly, dominated by first order links, some second-level links are also frequently observed (e.g., AID/HIV–Different kinds of STDs, AIDS/HIV–Life-threatening).

Concept 1	Concept 2	Occurrence
STD (core concept)	Different kinds of STDs	45.7%
STD (core concept)	AIDS/HIV	34.1%
STD (core concept)	Dangerous	30.0%
STD (core concept)	Condoms	25.1%
STD (core concept)	Preventable	23.3%
STD (core concept)	Sex	21.5%
STD (core concept)	Safe Sex	20.6%
STD (core concept)	Life-threatening	19.7%
STD (core concept)	Physical symptoms	19.3%
STD (core concept)	Risky life style	18.8%
STD (core concept)	Undesirable consequences	18.4%
STD (core concept)	Fear	17.9%
STD (core concept)	Social stigma	16.6%
STD (core concept)	Unsafe sex	15.7%
STD (core concept)	Disgusting	15.2%
STD (core concept)	Everyone can get it	15.2%
AIDS/HIV	Different kinds of STDs	14.3%
STD (core concept)	Permanent	14.3%
STD (core concept)	Global problem	13.9%
AIDS/HIV	Life-threatening	11.2%

Table 4: Link-level summary statistics: Top 20 links that occur most frequently are listed.

5   Application

We apply our model to our dataset on STDs. In Section 5.1, we assess our model's performance by recovering the key map-level, concept-level, and link-level summary statistics considered in Section 3. In Section 5.2, we interpret our parameter estimates. In Section 5.3, we test whether the concept maps are significantly different across the two genders. In Section 5.4, we demonstrate how to generate a consensus map from our dataset using our model.

5.1   Model validation

We estimated the model parameters with the procedure proposed in Section 2.3. Using the estimated parameters, we simulated 100 datasets, each containing 223 concept maps (which replicates the size of our actual data) from the model. We then computed, for each simulated dataset, all the map-level, concept-level, and link-level summary statistics considered in Section 3, and compared them with the summary statistics from the actual dataset. This procedure is similar to the posterior predictive checks proposed in Gelman et al. (16). If our model provides an adequate description of the data, the summary statistics from the simulated datasets should look similar to the observed summary statistics. In contrast, if the observed summary statistics look very different from those of the simulated datasets, inadequate model fit is suggested. By comparing observed and simulated summary statistics, we can validate our model performance and gain a deeper understanding of the limitations of our model.

Figure 6 shows the comparisons between the map-level summary statistics from the observed data and the simulated datasets. In each of the panels, the solid lines represent the observed data summary statistics, while the histograms represent the distribution of corresponding summary statistics calculated from the 100 simulated datasets (simulated using our model and the estimated parameters). As can be seen, each of the solid lines is located around the center of the corresponding distribution, suggesting that the map-level summary statistics considered are recovered very well by our model.

Figure 7 allows us to assess our model's performance in recovering concept-level and link-level summary statistics. The left panel of Figure 7 compares the % occurrence of each concept in the observed dataset (x-axis) and the mean across the 100 simulated datasets (y-axis). Similarly, the right panel of Figure 7 plots the observed % occurrence of each link (x-axis) vs. the mean % occurrence of each link across the 100 simulated datasets (y-axis). From the figure, we see that our model provides a fairly adequate fit to both concept-level and link-level summary statistics; we should note, however, that the simulated % occurrences are somewhat higher than the observed ones over the lower half of the range of observed values (and vice versa for the upper half). This suggests that the model here generates maps that are somewhat "more uniform" than the observed maps, and that it may be worth here considering elaborations of our model that allow for slightly more heterogeneity of concepts and links.

5.2   Parameter estimates and posterior intervals

We now look at the estimated model parameters more closely. Figure 8 shows a histogram and Table 5 shows the largest 20 of the posterior means $\widehat{\theta}_{jk}$ estimates. To a large extent, this list mirrors that of the observed % occurrences in Table 4. There are, however, a number of exceptions, particularly for the links between non-core concepts (i.e., second- or above order links). For instance, the link between the concepts "Condom" and "Safe Sex" occurs in only 11.2% of the maps, but the estimated $\widehat{\theta}_{jk}$ between them is 0.16; likewise, the link between "Condom" and "Preventable" occurs in 9.0% of the maps, while the estimated $\widehat{\theta}_{jk}$ between them is around 0.14. As we discussed
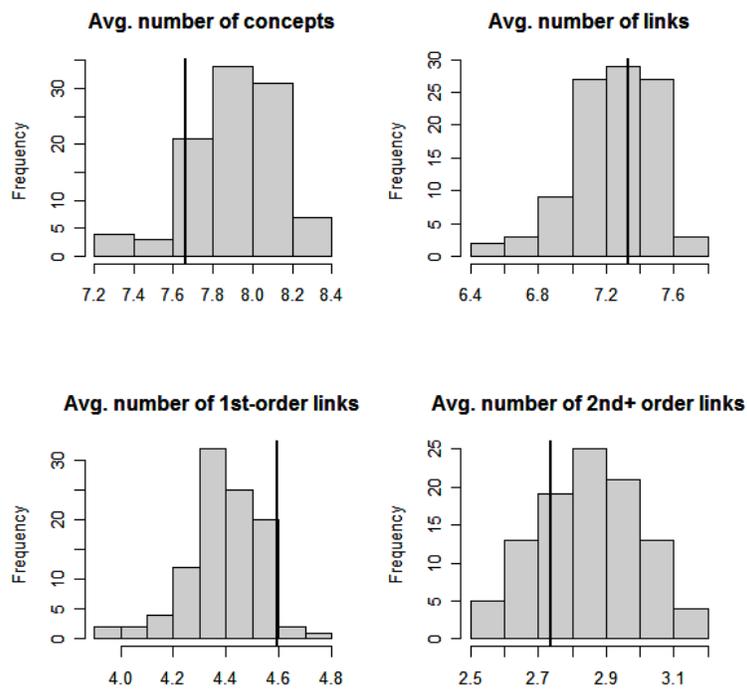
Figure 6: Comparison of map-level summary statistics between the observed and simulated datasets. The solid line denotes the summary statistics from the observed dataset, while the histograms show the corresponding distribution of the corresponding summary statistics generated from the 100 simulated datasets.
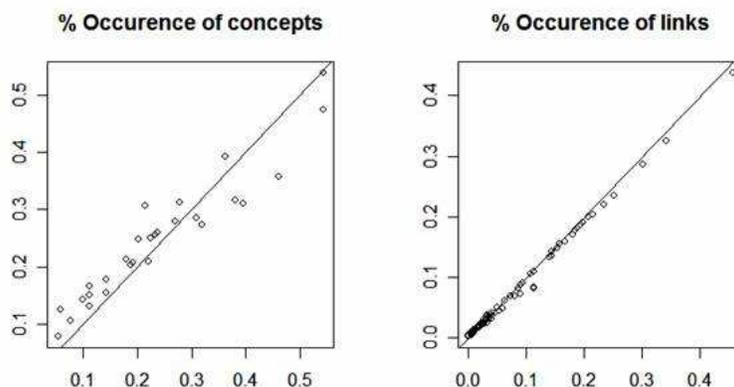
Figure 7: Comparison of concept-level (left panel) and link-level (right panel) summary statistics between the observed and simulated datasets. In each panel, the observed % occurrence of a concept/link is plotted on the x-axis, while the corresponding % occurrence of the concept/link across the 100 simulated datasets is plotted on the y-axis.

earlier in Section 2.3, the observed proportion of occurrences of links between concepts j and k underestimates θ_{jk} under our model (when neither j and k is the core concept), because the issue of censoring/pruning is not accounted for. Thus, focusing only on the observed occurrences without adjusting for censoring tends to overstate the importance of first-order links and understate the importance of the second (and above) order links.

Many other analyses can be performed based on the estimates of our model parameters $\Theta_{N \times N}$. For instance, the row sums of $\Theta_{N \times N}$ roughly correspond to how interrelated each concept is to all the other concepts, and is similar to the notation of "centrality," a key measure in social network analysis (e.g., Wasserman and Faust (48)). Since $\Theta_{N \times N}$ is symmetric by definition, the row sums and the column sums are equivalent. The centralities of each of the concepts, which we define as the row sums $\sum_k \theta_{jk}$ for the j-th concept, are listed in Table 6 (sorted in decreasing centrality) along with their corresponding 95% posterior intervals, calculated using the posterior distributions of θ_{jk}'s. By design, the core concept "STD" has the highest centrality (4.38). The concepts with the next highest centralities are "AIDS/HIV" (1.18), "Different kinds of STDs" (1.15), and "Condoms" (0.89). Thus, our model not only allows researchers to introduce tools/measures from network analysis to study concept maps, but also to generate interval estimates through the posterior distribution of our model parameters.

5.3   Hypothesis testing

Using the procedure described in Section 2.4, we assess whether there are any differences between the concept maps produced by male and female participants. Formally, we compare the following two hypothesis, which correspond to an omnibus test of homogeneity
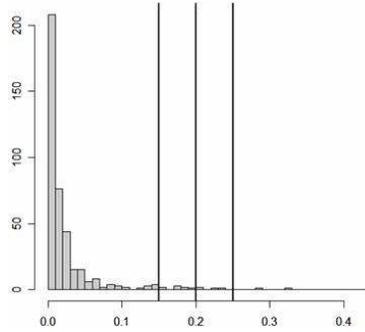
Figure 8: Histogram of posterior means $\widehat{\theta}_{jk}$. The vertical lines marks the cutoff values of c used (0.15, 0.20, 0.25) when generating consensus maps in Figure 9.

Concept 1	Concept 2	$\widehat{\theta}_{jk}$	95% Post. interval	Occurrence
STD (core concept)	Different kinds of STDs	0.43	(0.37, 0.50)	45.7%
STD (core concept)	AIDS/HIV	0.32	(0.27, 0.38)	34.1%
STD (core concept)	Dangerous	0.29	(0.23, 0.34)	30.0%
STD (core concept)	Condoms	0.24	(0.19, 0.30)	25.1%
STD (core concept)	Preventable	0.22	(0.17, 0.28)	23.3%
STD (core concept)	Sex	0.21	(0.16, 0.26)	21.5%
AIDS/HIV	Different kinds of STDs	0.20	(0.14, 0.27)	14.3%
STD (core concept)	Safe sex	0.20	(0.15, 0.25)	20.6%
STD (core concept)	Life-threatening	0.19	(0.14, 0.24)	19.7%
STD (core concept)	Transmittable	0.18	(0.14, 0.24)	19.3%
STD (core concept)	Physical symptoms	0.18	(0.13, 0.23)	18.8%
STD (core concept)	Risky life style	0.18	(0.13, 0.23)	18.4%
STD (core concept)	Undesirable consequences	0.17	(0.13, 0.22)	17.9%
Condoms	Safe sex	0.16	(0.11, 0.22)	11.2%
STD (core concept)	Fear	0.16	(0.11, 0.21)	16.6%
STD (core concept)	Social stigma	0.15	(0.10, 0.20)	15.7%
AID/HIV	Life-threatening	0.15	(0.10, 0.19)	11.2%
STD (core concept)	Unsafe sex	0.15	(0.10, 0.19)	15.2%
STD (core concept)	Disgusing	0.15	(0.10, 0.20)	15.2%
Condoms	Preventable	0.14	(0.09, 0.20)	9.0%

Table 5: The 20 largest $\widehat{\theta}_{jk}$ estimates.

Concept	Estimated centrality	95% Posterior interval
STD	4.38	(4.15, 4.63)
AID/HIV	1.18	(1.03, 1.33)
Different kinds of STDs	1.15	(1.01, 1.29)
Condoms	0.89	(0.76, 1.03)
Undesirable consequences	0.87	(0.74, 1.01)
Disease	0.76	(0.63, 0.90)
Life-threatening	0.75	(0.62, 0.88)
Dangerous	0.73	(0.60, 0.89)
Medical treatment	0.73	(0.59, 0.88)
Get tested	0.70	(0.58, 0.84)
Fear	0.64	(0.52, 0.79)
Everyone can get it	0.64	(0.52, 0.78)
Transmittable	0.63	(0.51, 0.77)
Sex	0.60	(0.47, 0.74)
Social stigma	0.60	(0.47, 0.75)
Unsafe sex	0.54	(0.43, 0.68)
Not for me	0.49	(0.38, 0.62)
Sexual partner	0.48	(0.35, 0.62)
Preventable	0.47	(0.36, 0.61)
Bacteria/viral/parasitic	0.45	(0.32, 0.61)
Safe sex	0.45	(0.31, 0.60)
Permanent	0.45	(0.34, 0.58)
Risky life style	0.44	(0.33, 0.56)
Health cost	0.40	(0.27, 0.56)
Global problem	0.39	(0.28, 0.51)
Disgusting	0.38	(0.28, 0.51)
Physical symptoms	0.36	(0.25, 0.49)
Education	0.30	(0.19, 0.44)
Prevalent	0.29	(0.18, 0.44)

Table 6: Estimated "centrality" (i.e., $\sum_k \widehat{\theta}_{jk}$) of each concept and their 95% posterior intervals.

across genders:

$$H_0 : \theta_{jk}^{(male)} = \theta_{jk}^{(female)} \qquad \forall j, k$$

$$H_1 : \theta_{jk}^{(male)} \neq \theta_{jk}^{(female)} \quad for \; some \; j, k$$

The Bayes factor $BF_{H_1:H_0}$, computed using Equation [18], is smaller than 0.001, indicating that the data strongly supports H_0, i.e., the model in which male and female participants are pooled (Jeffreys 1961).

We also separately compute the Bayes factors which compare whether male and female participants are different with respect to each particular link. That is, we evaluate the following pairs of hypothesis for each pair of links between the concepts j and k:

$$H_0(j, k) : \theta_{jk}^{(male)} = \theta_{jk}^{(female)}$$

$$H_1(j, k) : \theta_{jk}^{(male)} \neq \theta_{jk}^{(female)}$$

Most Bayes factors ($H_1(j, k)$ over $H_0(j, k)$) are smaller than 1, suggesting that there is no difference between male and female participants with respect to most of the links. For some links, however, we obtain Bayes factors larger than 1, suggesting some specific differences between male and female participants. The links with Bayes factors larger than 2 (i.e., in favor of $H_1(j, k)$) are listed in Table 7. As can be seen, the two largest observed Bayes factors are around 6.11 (for the link "Dangerous"–"Physical symptoms") and 4.09 (for the link "Life-threatening"–"Disease"), which offers some evidence that male and female participants tend to have different strengths of association on those links. In particular, the data suggest that male participants are more likely to exhibit the above two links in their concept maps than female participants.

Concept	Concept	$\widehat{\theta}_{jk}^{male}$	$\widehat{\theta}_{jk}^{female}$	Bayes Factor (in favor of $H_1(j,k)$)
Dangerous	Physical symptoms	0.064	0.004	6.11
Life-threatening	Disease	0.073	0.005	4.09
Transmittable	Permanent	0.061	0.005	2.93
Fear	Safe sex	0.052	0.004	2.89
AIDS/HIV	Unsafe sex	0.042	0.003	2.34
Life-threatening	Undesirable consequences	0.044	0.004	2.16

Table 7: Testing for difference between male and female participants using Bayes Factor.

5.4   Consensus map

As another application of our modeling framework, we demonstrate how a consensus map can be constructed based solely on our model parameters by using the procedure

outlined in Section 2.4. The male and female participants are pooled to form a single consensus map for the data. Figure 9 shows the consensus map generated with $c = 0.25$ (top panel), 0.20 (middle panel), 0.15 (bottom panel). These values of c are marked by the vertical lines in Figure 8. Note that we have chosen these cutoff values to illustrate our approach; it may be interesting to treat the selection of c by a data-driven approach. We return to this issue in the conclusion.

As shown in the figure, as the value of c decreases, the consensus map becomes more complex with more concepts and links between concepts. In the top panel (with $c = 0.25$), only three first-order links are present. The middle panel ($c = 0.20$) has 7 links; it also comprises predominantly first-order links, with the second order link between "AIDS/HIV" and "Different kinds of STDs" also present. The bottom panel (with $c = 0.15$) shows a very complicated concept map structure with a total of 19 links. Since the mean number of links across our dataset is 7.54 (see Table 3), we feel that the middle panel (with $c = 0.20$) seems to represent the best compromise; a more formal approach, as we discuss later, is to select the value of c through a data-driven approach and then generate the consensus map that corresponds to the chosen value of c.

Note that in the previous literature, the inclusion of a link in a consensus map has been typically based on the number of times a link is observed, i.e., Equation [5]. In such cases, the decision of whether to include a link in a consensus map has been based on the ranking of observed link occurrence (Table 4). Holding the total number of links to be the same as that of Figure 8, the map in the middle panel contains no second level links, and the bottom panel map contains only one second-level link (see Table 4). Thus, the importance of the second-level (and above) links tends to be de-emphasized. Although we do not claim that consensus maps generated by our procedure are "superior" (because there is no formal definition a consensus map), we want to highlight again the fact that the consensus maps we generate are based entirely on sufficient statistics, an important criterion for sound statistical procedures from the viewpoint of statistical decision theory (Berger (4)).

6   Conclusion and future research

In this paper, we build on the introduction of concept maps into marketing by John et al. (23) and developed the methodological framework for performing statistical analyses on concept maps. Specifically, we have developed a descriptive probability model of concept maps by extending the uniform graph model in two directions: (i) allowing for non-uniform probabilities of link presence, and (ii) introducing a pruning procedure. We developed useful estimation and hypothesis testing procedures, and proposed a methodology to create consensus maps using a model-based framework. We then calibrated our model on a dataset that studies college students' perceptions of STDs, an important issue in the health marketing literature that could help design effective educational programs to "market" the healthy beliefs about the safe sex and decrease the STD risk on college campuses.

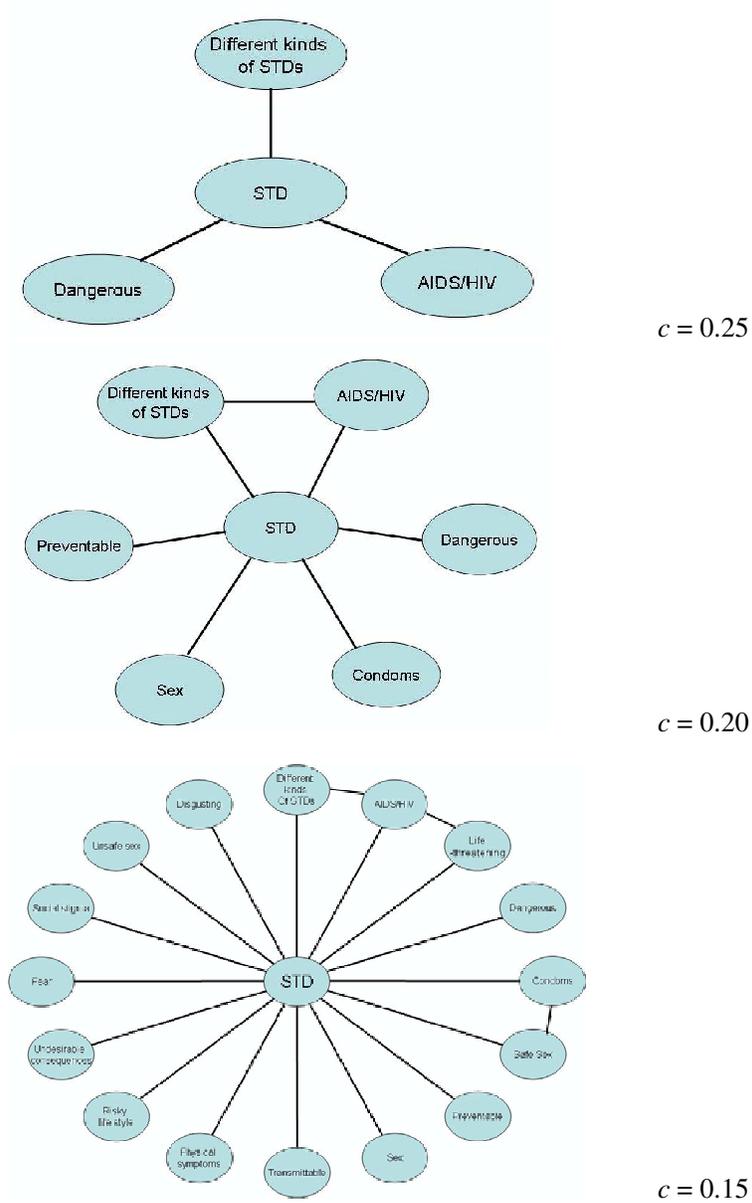Although our model is relatively parsimonious, it is able to recover considered map-

Figure 9: Consensus maps generated with *c*=0.25 (top panel), *c*=0.20 (middle panel), *c*=0.15 (bottom panel).

level, concept-level, and link-level summary statistics. We then explored the kinds of analyses that can be done using our model parameters, tested for gender difference in concept map formation, and generated consensus maps (for varying c) corresponding to our dataset.

Our model represents a first step towards a unified model-based analysis of concept maps in marketing. It can be applied in many different settings where data in the form of concept maps are collected. In addition, our parsimonious model can be extended in many directions when more information (both from the concept map and individual-level covariates) is available. We explore some of these extensions and future research directions below:

(i) *Automatic selection of c*: As we mentioned earlier, rather than choosing the value of the cutoff c subjectively, one could also use a data-driven selection of c. This could be done, for instance, using tools developed in Bayesian variable selection (e.g., George and McCulloch (18)). For example, instead of the Beta prior that we used in Equation [7], we could specify that the θ_{jk}'s are drawn from a mixture distribution of a "high" Beta distribution with parameters (a_H, b_H) and a "low" Beta distribution with parameters (a_L, b_L), with the assumption that the links from the "high" distribution should be incorporated in the consensus map, while those from the "low" distribution should be excluded. We restrict the "high" distribution to have a larger mean than the "low" distribution. Formally, let ζ_{jk} be an indicator variable that takes the value 1 if θ_{jk} comes from the high distribution (and hence should appear in the consensus map), and 0 otherwise; we can then modify Equation [7] to the specification below:

$$\theta_{jk}|\zeta_{jk} = 1 \sim Beta(a_H, b_H)$$
$$\theta_{jk}|\zeta_{jk} = 0 \sim Beta(a_L, b_L)$$
$$P(\zeta_{jk} = 1) = \lambda$$

This setup is similar in spirit to the framework in George and McCulloch (17), who perform variable selection in the context of multiple linear regression. We sample the posterior of the indicator variables ζ_{jk} , and chose the θ_{jk}'s with $E(\zeta_{jk}|\text{data}) > 0.5$ to include into the consensus map. The above method, however, involves more computational effort than our current methodology, where closed-form solutions can be obtained.

(ii) *Model-based clustering*: Specifying a latent class structure for our model parameters allows researchers to cluster a dataset of concept maps into different groups. These cluster memberships can then be linked to particular outcome variables, e.g., attitude towards safe sex, brand preference. These relationships may have important implications on how intervention/advertising efforts should be allocated in the effective marketing campaign.

(iii) *Individual-level covariates*: If individual-level variables of interest are available for each participant, we can incorporate them into our model parameters to study the effect of individual-level covariates. Formally, let $\Theta^i_{N \times N}$ be the matrix of parameters for the

i-th individual, $\vec{x}_i$ be a vector of individual-level covariates for the i-th individual and let ϕ^i be the vector of parameters by writing out all the parameters in the matrix $\Theta^i_{N \times N}$. We can specify:

$$\text{logit}(\phi^i) = \bar{\phi} + \Psi \vec{x}_i + \vec{\epsilon} \tag{20}$$

where $\bar{\phi}$ is a vector of intercepts, and Ψ is a matrix of coefficients that describes how model parameters can related to the individual-level characteristics in $\vec{x}_i$, and $\vec{\epsilon}$ are the errors terms.

(iv) *Model extension*: Finally, as we briefly mentioned earlier, some other forms of concept maps record more information than the undirected graph version that we focus on in this paper. In particular, some concept maps record not only the presence of links between two concepts, but also the *directionality* of those links. This allows us to capture potential "asymmetries" of the relationship between concept, e.g., the probability of generating concept A given concept B may not be the same as the probability of generation concept B given concept A.[11] For such maps, it may be more appropriate to construct a dynamic model that captures the order in which concepts are generated, instead of the "static" approach taken in this paper. More generally, we believe that it would be a challenge and opportunity for us and other researchers to extend our model to the specific situation when different forms of concept maps are collected. We hope that this paper can serve as a useful guideline for future analyses of concept map data.

7 Appendix

In the following discussion, we outline a proof which shows that the estimator in Equation [6] is the MLE estimator, and that the estimator in Equation [5] is a biased estimator of our model parameters. For illustrative purpose, we consider the following case which resembles the "pruning" feature of our model. Let i be a person index that goes from $1, 2, \ldots, I$. Let u_{i1} and u_{i2} be latent Bernoulli random variables with parameters θ_1 and θ_2, respectively. Thus:

$$P(u_{i1} = 1) = \theta_1; \quad P(u_{i2} = 1) = \theta_2 \tag{21}$$

In order to replicate the "pruning" feature in our model, we let observed random variables v_{i1} to be always equal to u_{i1}, and let v_{i2} to be equal to u_{i2} if $u_{i1} = 1$, and $v_{i2} = 0$ otherwise. Formally,

$$v_{i1} = u_{i1}; \quad v_{i2} = u_{i2} 1_{(u_{i1}=1)} \tag{22}$$

Note again that only the variables v_{i1} and v_{i2}, but not u_{i1} and u_{i2}, are observed. Thus, the structure of Equation [21] and Equation [22] corresponds to the pruning feature of

---

[11]We thank an anonymous referee for this suggestion

our model. Under our model of concept map, even when a link between two concepts may exist (i.e., in our analogy here, $u_{i2} = 1$), that link may not appear in the observed map (i.e., $v_{i2} = 0$ in our analogy) if neither of the concepts is connected to the core concept (i.e., the event $v_{i2} = 0$ in this analogy).

We will now show that the MLE estimator for θ_2 is

$$\widehat{\theta}_2^{MLE} = \frac{\sum_{i=1}^{I} v_{i1} v_{i2}}{\sum_{i=1}^{I} v_{i1}} \tag{23}$$

which corresponds to the estimator we proposed in Equation [6]. We begin by writing down the probability function of (v_{i1}, v_{i2}) by considering the three possible cases as follows:

$$P(v_{i1} = 0, v_{i2} = 0) = P(u_{i1} = 0, u_{i2} = 0) + P(u_{i1} = 0, u_{i2} = 1) \tag{24}$$

$$= (1 - \theta_1)(1 - \theta_2) + (1 - \theta_1)\theta_2 = 1 - \theta_1 \tag{25}$$

$$P(v_{i1} = 1, v_{i2} = 0) = P(u_{i1} = 1, u_{i2} = 0) = \theta_1(1 - \theta_2) \tag{26}$$

$$P(v_{i1} = 1, v_{i2} = 1) = P(u_{i1} = 1, u_{i2} = 1) = \theta_1 \theta_2 \tag{27}$$

Using the set of equations in [24-27], the likelihood function for (θ_1, θ_2) can be written as:

$$l(\theta_1, \theta_2) = (1 - \theta_1)^{\sum_i 1_{(v_{i1}=0)}} \left(\theta_1(1 - \theta_2)\right)^{\sum_i 1_{(v_{i1}=1 \ and \ v_{i2}=0)}} \left(\theta_1 \theta_2\right)^{\sum_i 1_{(v_{i1}=1 \ and \ v_{i2}=1)}} \tag{28}$$

Taking logs,

$$L(\theta_1, \theta_2) = \sum_i 1_{(v_{i1}=0)}(\log(1 - \theta_1)) + \sum_i 1_{(v_{i1}=1 \ and \ v_{i2}=0)}(\log(\theta_1) + \log(1 - \theta_2))$$
$$+ \sum_i 1_{(v_{i1}=1 \ and \ v_{i2}=1)}(\log(\theta_1) + \log(\theta_2)) \tag{29}$$

Finally, we maximize Equation [29] with respect to θ_2 and consider the first-order condition:

$$\frac{\sum_i 1_{(v_{i1}=0 \ and \ v_{i2}=1)}}{\theta_2} - \frac{\sum_i 1_{(v_{i1}=1 \ and \ v_{i2}=0)}}{1 - \theta_2} = 0 \tag{30}$$

$$\Rightarrow \widehat{\theta}_2^{MLE} = \frac{\sum_i 1_{(v_{i1}=1 \ and \ v_{i2}=1)}}{\sum_i 1_{(v_{i1}=1)}} = \frac{\sum_{i=1}^{I} v_{i1}v_{i2}}{\sum_{i=1}^{I} v_{i1}} \tag{31}$$

Since the estimator in Equation [31] is the MLE, it is efficient and (asymptotically) unbiased (Casella and Berger (9)). It is also sufficient for θ_2. In contrast, the estimator in Equation [5] is a biased because $\frac{\sum_{i=1}^{I} v_{i1}v_{i2}}{I} \leq \frac{\sum_{i=1}^{I} v_{i1}v_{i2}}{\sum_{i=1}^{I} v_{i1}}$ (equality holds if and only if $v_{i1} = 1 \ \forall i$).

References

[1] Albert, R. and Barabasi, A.-L. (2002). "Statistical Mechanic of Complex Networks." *Review of Modern Physics*, 74: 47–97. 484

[2] Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard Press. 479

[3] Anderson, J. R. and Bower, G. H. (1973). *Human Associative Memory*. Washington, D.C.: Winston. 479

[4] Berger, J. O. (1993). *Statistical Decision Theory and Bayesian Analysis, 2nd Edition*. NY: Springer. 488, 504

[5] Block, L. G. and Keller, P. A. (1995). "When to Accentuate the Negative: The Effects of Perceived Efficacy and Message Framing on Intentions to Perform a Health-Related Behavior." *Journal of Marketing Research*, 32: 192–203. 493

[6] — (1997). "Effects of Self-Efficacy and Vividness on the Persuasiveness of Health Communications." *Journal of Consumer Psychology*, 6(1): 31–54. 493

[7] Bollobas, B. (2001). *Random Graphs*. NY: Cambridge University Press, 2nd edition. 483, 484

[8] Carlin, B. P. and Louis, T. A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*. NY: Chapman and Hall, 2nd edition. 482, 489

[9] Casella, G. and Berger, R. L. (2002). *Statistical Inference*. CA: Duxbury, 2nd edition. 509

[10] Chandran, S. and Menon, G. (2004). "When a Day Means More than a Year: Effects of Temporal Framing on Judgments of Health Risk." *Journal of Consumer Research*, 31: 375–389. 493

[11] Chang, K.-E., Sung, Y.-T., Chang, R.-B., and Lin, S.-C. (2005). "A New Assessment for Computer-based Concept Mapping." *Educational Technology and Society*, 8(3): 138–148. 481

[12] Chattopadhyay, A. and Alba, J. W. (1988). "The Situational Importance of Recall and Inference in Consumer Decision Making." *Journal of Consumer Research*, 15: 1–12. 495

[13] Dansereau, D., McDonald, B., Collins, K., Garland, J., Holley, C., Diekhoff, G., and Evans, S. (1979). "Evaluation of A Learning Strategy System." *Cognitive and affective learning strategies (H.F. O'Neil Jr. and C. D. Spielberger (Eds)), New York: Academic Press*, 3–43. 479

[14] Erdos, P. and Renyi, A. (1959). "On Random Graphs." *Publicationes Mathematicae*, 6: 290–297. 481, 482, 484

[15] Fisher, R. A. (1922). "On the Mathematical Foundations of Theoretical Statistics." *Philosophical Transactions of the Royal Society of London A*, 222: 309–368. 488

[16] Gelman, A., Meng, X.-L., and Stern, H. (1996). "Posterior Predictive Assessment of Model Fitness Via Realized Discrepancies." *Statistica Sinica*, 6: 733–807. 498

[17] George, E. I. and McCulloch, R. E. (1993). "Variable Selection Via Gibbs Sampling." *Journal of the American Statistical Association*, 88: 881–889. 506

[18] — (1997). "Approaches for Bayesian Variable Selection." *Statistica Sinica*, 7: 339–373. 506

[19] Goldsmith, T. E., Johnson, P. J., and Acton, W. H. (1991). "Assessing Structural Knowledge." *Journal of Educational Psychology*, 83(1): 88–96. 481

[20] Hill, R. P. (1988). "An Exploration of the Relationship between AIDS-Related Anxiety and the Evaluation of Condom Advertisement." *Journal of Advertising*, 17(4): 35–42. 479, 481

[21] Huang, Y., Hui, S. K., and Kahn, B. E. (2007). "Health Communication Effectiveness: Using Underlying Processes to Understand the Relationship between Risk Attitudes and Behavioral Intentions." *University of Pennsylvania Marketing Department Working Paper*. 479, 493

[22] Jeffreys, H. (1961). *Theory of Probability*. Oxford: Clarendon Press. 491

[23] John, D. R., Loken, B., Kim, K., and Monga, A. B. (2006). "Brand Concept Maps: A Methodology for Identifying Brand Association Networks." *Journal of Marketing Research*, 43: 549–563. 479, 480, 481, 482, 492, 493, 495, 496, 504

[24] Joiner, C. (1998). "Concept Mapping in Marketing: A Research Tool for Uncovering Consumers' Knowledge Structure Associations." *Advances in Consumer Research*, 25: 311–317. 479, 480, 481, 482, 488, 492, 493, 496

[25] Keller, P. A. (2006). "Regulatory Focus and Efficacy of Health Messages." *Journal of Consumer Research*, 33: 109–114. 493

[26] Keller, P. A., Lipkus, I. M., and Rimer, B. K. (2002). "Depressive Realism and Health Risk Accuracy: The Negative Consequences of Positive Mood." *Journal of Consumer Research*, 29: 57–69. 493

[27] Klein, J. P. and Moeschberger, M. L. (2005). *Survival Analysis: Techniques for Censored and Truncated Data.* NY: Springer, 2nd edition. 488, 489

[28] LaTour, M. S. and Pitts, R. (1989). "Using Fear Appeals in Advertising for AIDS Prevention in the College-Age Population." *Journal of Health Care Marketing*, 9: 5–14. 479, 481, 493

[29] Levy, E., Louchard, G., and Petit, J. (2005). "A Distributed Algorithm to Find Hamiltonian Cycles in Random Graphs." *Lecture Notes in Computer Science*, 3405: 63–74. 484

[30] Lindsay, P. H. and Norman, D. A. (1972). *Human Information Processing.* San Diego: Academic Press. 479

[31] Liu, C.-C., Don, P.-H., and Tsai, C.-M. (2005). "Assessment Based on Linkage Patterns in Concept Maps." *Journal of Information Science and Engineering*, 21: 873–890. 481

[32] Lord, C. G., Desforges, D. M., Fein, S., Pugh, M. A., and Lepper, M. (1994). "Typicality Effects in Attitudes Toward Social Policies: A Concept-Mapping Approach." *Journal of Personal and Social Psychology*, 66(4): 658–673. 479, 480, 495

[33] MacDonald, N. E., Wells, G. A., Fisher, W. A., Warren, W. K., King, M. A., Doherty, J. A., and Bowie, W. R. (1990). "High-Risk STD/HIV Behavior among College Students." *Journal of American Medical Association*, 263: 3155–3159. 481

[34] Mahoney, C., Tombs, D., and Ford, O. (1995). "Health Belief and Self-Efficacy Models: Their Utility in Explaining College Student Condom Use." *AIDS Education and Prevention*, 7: 32–49. 481

[35] Marshall, B., Chen, H., and Madhusudan, T. (2006). "Matching Knowledge Elements in Concept Maps using a Similarity Flooding Algorithm." *Decision Support Systems*, 42: 1290–1306. 481

[36] Martin, J. (1985). "Measuring Clients' Cognitive Competence in Research on Counseling." *Journal of Counseling and Development*, 63: 556–560. 479, 480

[37] Menon, G., Block, L. G., and Ramanathan, S. (2002). "We're at as Much Risk as We are Led to Believe: Effects of Message Cues on Judgments of Health Risk." *Journal of Consumer Research*, 28: 533–549. 493

[38] Moore, D. and Readence, J. (1984). "A Quantitative and Qualitative Review of Graphic Organizer Research." *Journal of Educational Research*, 78: 11–17. 479

[39] Novak, J. D. and Gowin, D. B. (1984). *Learning How to Learn.* Cambridge, England: Cambridge University Press. 479, 480

[40] Raghubir, P. and Menon, G. (1998). "AIDS and Me, Never the Twain Shall Meet: The Effects of Information Accessibility on Judgments of Risk and Advertising Effectiveness." *Journal of Consumer Research*, 25: 52–63. 479, 481, 493

[41] Rumelhart, D. E., Lindsay, P. H., and Norman, D. A. (1972). "A Process Model for Long-Term Memory." *Organization of Memory (Endel Tulving and Wayne Donald (Eds)), New York: Academic Press.* 479

[42] Shavitt, S. and Brock, T. C. (1986). "Self-Relevant Responses in Commercial Persuasion: Field and Experimental Test." *Advertising and Consumer Psychology (Jerry Olson and Keith Sentis (Eds)), New York: Praeger Publishers.* 495

[43] Stuart, H. A. (1985). "Should Concept Maps be Scored Numerically?" *European Journal of Science Education*, 7(1): 73–81. 479

[44] Takeya, M., Sasaki, H., Nagaoka, K., and Yonezawa, N. (2004). "A Performance Scoring Method based on Quantitative Comparison of Concept Maps by a Teacher and Students." *Concept Maps: Theory, Methodology, Technology- Proceedings of the First International Conference on Concept Mapping (A.J. Canas, J.D. Novak, and F.M Gonzalez (Eds)), Spain: Direccion de Publicaciones de la Universidad Publica de Navarra.* 481

[45] Takeya, R. N. and Young, S. S. (1986). "The Incorporation of Historial Control Information in Tests of Proportions: Simulation Study of Tarone's Procedure." *Biometrics*, 42: 343–349. 489

[46] Treise, D. and Weigold, M. (2001). "AIDS Public Service Announcements: Effects of Fear and Repetition on Predictors of Condom Use." *Health Marketing Quarterly*, 18: 39–61. 479, 481, 493

[47] Wainer, H., Bradlow, E. T., and Wang, X. (2007). *Testlet Response Theory and Its Applications.* NY: Cambridge University Press. 488

[48] Wasserman, S. and Faust, K. L. (1994). *Social Network Analysis: Methods and Appliations.* NY: Cambridge University Press. 500

Acknowledgments