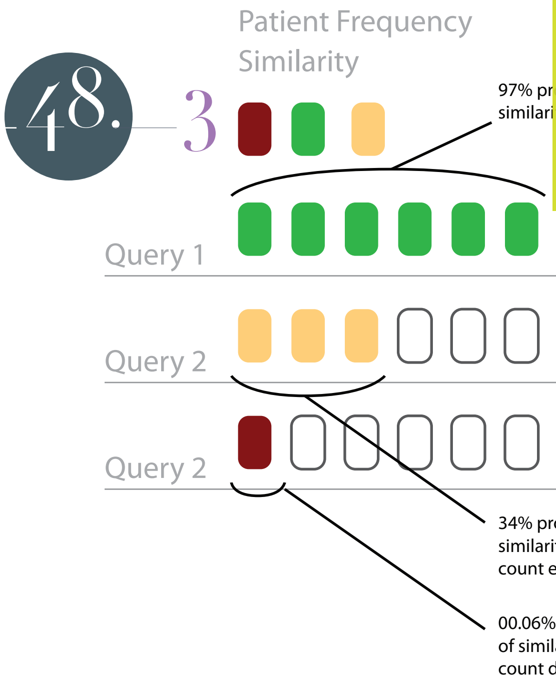
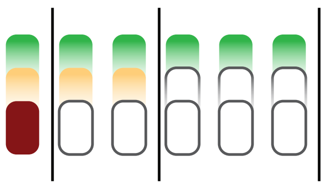


visualization logic



six boxes filled in proportion to show inter-hospital similarities
six boxes need to show probability over/under 50%
six boxes can be rapidly enumerated (in about 1 second)

A Statistical Approach for Visualizing the Quality of Multi-Hospital Data

Connolly et. al.

ISSN 0022-2224

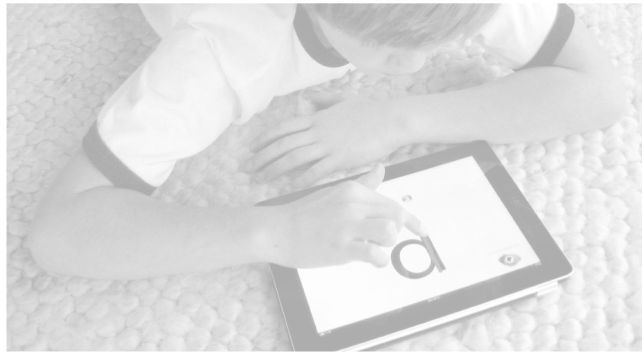
Visible Language 48.3

the journal of visual communication research

November 2014

Visible Language

the journal of visual communication research



Linking Design Principles with Educational Research Theories to Teach Sound to Symbol Reading Correspondence with Multisensory Typea

Seward et. al.

Visible Language

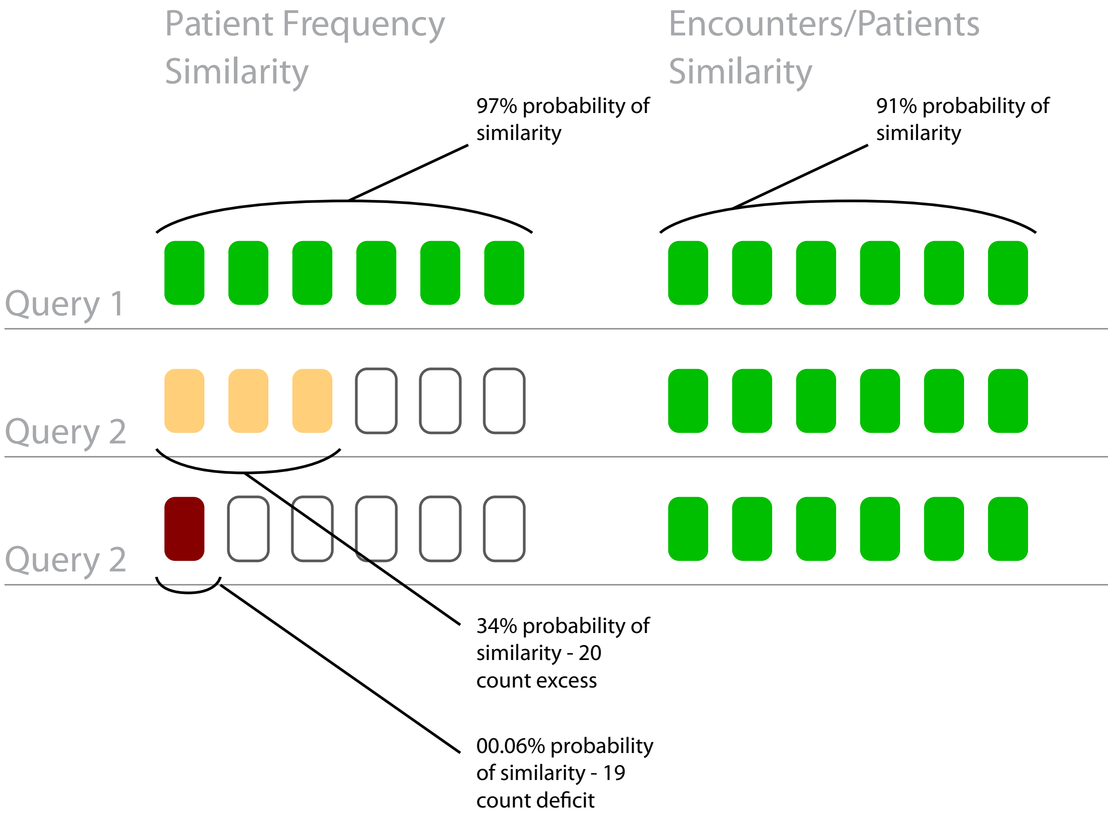
the journal of visual communication research

48 — 3

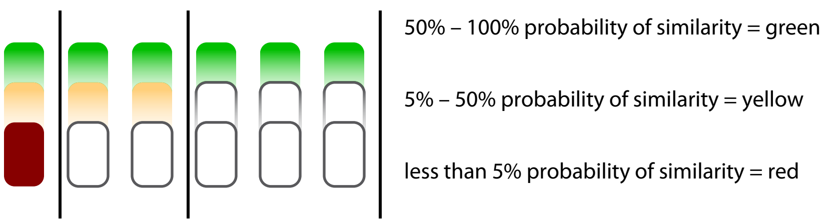
November 2014

48–3

Visible Language



visualization logic



six boxes filled in proportion to show inter-hospital similarities
six boxes need to show probability over/under 50%
six boxes can be rapidly enumerated (in about 1 second)

see *figure 1a*
for details

A Statistical Approach for Visualizing the Quality of Multi-Hospital Data

3

Brian Connolly, Robert Faist , Constance West, Katherine D. Holland, Pawel Matykiewicz ,
Tracy A. Glauser, and John Pestian[1]

Cincinnati Children's Hospital Medical Center, Cincinnati OH USA

[1] corresponding author: 3333 Burnet Ave., MLC 7024, Cincinnati, OH 45229-3039,
phone: 513-636-1627, fax: 513-636-2056, john.pestian@cchmc.org

ABSTRACT

The age of Big Data and the associated proliferation of large data sets have necessitated the development of methods that allow for an easy interpretation of data analysis results. Such methods are usually the product of a symbiotic relationship between the fields of data visualization, infographics, and statistics. In this work we explore the interplay between data visualization and the mathematical framework used to analyze inter-hospital differences in database queries. Such differences can reflect disparities in the quality of care or more fundamental disparities in data quality. As the volume of queries is large and increasing, it is important to develop an incisive way of visualizing these differences. Specifically, we demonstrate the importance of choosing a mathematical framework that calculates the statistics necessary to visualize the results in a maximally concise and intuitive way. We derive symbolic statistical representations of inter-hospital query differences using a Bayesian probabilistic formalism to indicate statistically significant discrepancies. These statistical representations serve the need for visual representation of differences and their meaning apart from statistical expertise. The calculations were performed with a publically-available package, DQM, available at http://sourceforge.net/projects/databasequalitymanagement.

KEY WORDS

visualization, statistical methods, Bayesian analysis, data display, information display

1 INTRODUCTION

1. BACKGROUND

Analyzing large data sets presents a unique set of challenges, one of which is presenting results in a concise, easy-to-interpret way. In this work, we explore the impact of selecting the right mathematical framework that can be used "behind the scenes" to fully unveil the information potential of data visualizations.

The methodology described in this work is demonstrated within the context of healthcare. The healthcare system is under pressure from private citizens, businesses, and government to reduce costs while simultaneously improving patient safety and outcomes. Healthcare organizations have the opportunity to leverage Big Data technology to utilize the wealth of information generated during patient encounters with providers, pharmacies, hospitals, and social agencies. However, much of the clinically relevant information, including physician and provider notes, laboratory and imaging results, correspondence, and insurance claims data, is unstructured and thus quite complex to analyze. Theoretically, the aggregated data may be used to understand disease and to manage both individual and population behaviors. It is reasonable to expect that information from that data will drive efficiency and economy in healthcare consumption and delivery.

One of the difficulties in aggregating data from multiple organizations or populations is understanding the similarities and differences in the prevalence, severity, and presentation of a particular disease or condition in the populations of interest. A cursory evaluation suggests that there is considerable variation between providers in the treatment of even the most common adult and pediatric conditions. For example, children in Lebanon, New Hampshire are more than twice as likely to undergo a tonsillectomy as children less than 200 miles away in Bangor, Maine (The Dartmouth Atlas Working Group, 2013). Similar variations in adult surgical care have been reported by Birkmeyer et al. (Birkmeyer et al., 2013).

In this work, we compare database queries requesting data on epilepsy patients from three hospitals: Cincinnati Children's Hospital Medical Center (CCHMC), Children's Hospital of Philadelphia (CHOP), and Children's Hospital Colorado (CHCO). It is important to identify inter-hospital variations in such queries, as they can not only reveal differences in patient care but also more fundamental differences in data quality. Such differences can make it difficult to perform meaningful inter-hospital comparisons of patient care. To facilitate understanding one might wish to develop a concise and intuitive visualization scheme that identifies and qualifies the differences.

The queries themselves take the form of searches for "facts" within the database. For instance, a search for the fact "age" returns the number of patients for whom their age was reported. Although the query results do not give specifics as to what was reported, the search for a fact itself can reveal important information. For example, the search for "Diagnostic testing / Neuroimaging / MRI" that returns non-null results indicates that MRI neuroimaging was performed.

For each hospital, a query generally returns two numbers: the patient frequency (PF), and the total number of encounters over all patients for a given time period. The PF is the number of patients for whom a certain fact was reported. The number of total encounters accounts for all of the total documented patient contacts. A "patient contact" refers to every possible encounter, including outpatient visits, surgery, telephone calls to/from providers, and prescription refills.

The PFs are, of course, different among hospitals, as hospitals treat different numbers of patients. If the hospitals operate similarly, then a query should extract, within statistical fluctuations, the same percentage of patients from each hospital. Therefore, determining whether or not the PFs from a query are "different" among hospitals is to determine whether or not there is a statistically significant difference between the distribution of total patients over the hospitals and the distribution of patients in the queries. Using an example to illustrate, suppose three hospitals see 1,000, 2,000 and 3,000 patients in a month, respectively. Queries are run, and it is found that 100, 200 and 300 patients received MRI's from the three hospitals, respectively. Comparing the ratio or distribution of total patients over the hospitals, 1,000:2,000:3,000, to queried patients, 100:200:300, it is trivially apparent that there is no statistically significant difference between them. Note that this procedure is the same as comparing the percentage of patients that have had MRIs in the hospitals; in practice, however, ratios of PFs are compared. Of course a given query regarding epilepsy patients is not performed over every treated patient. Rather, a query is performed on a "base pool" (BP) of epilepsy patients.

The total number of encounters can also be different among hospitals. In understanding inter-hospital differences, the number of encounters by itself is less interesting than the encounter-to-patient ratio (EPR), or the number of encounters per patient. Note that the encounters are counted over a certain time period, so it is assumed that the query is made over the same time period for all hospitals. Unlike PFs, the EPR is expected to be the same across hospitals if the hospitals operate in a similar manner with similar types of patients. Finding inter-hospital differences using the EPR is then a matter of simply comparing the ratio of total encounters to total patients in the hospitals. For instance, if three hospitals have 300, 400 and 500 patients in their BPs, and a total of 600, 800 and 900 encounters over those patients, then the encounter-to-patients ratios for the three hospitals would be 2:1, 2:1 and 4.5:1, respectively. The ratio of the third hospital is different from the other two; however, we would have to evaluate whether that difference is statistically significant using the formalism described below.

To begin to understand inter-hospital query variations, one would like to know whether the differences in the PFs or EPRs are statistically significant. It is also desirable to visually represent the meaning and magnitude of the differences found.

There are a number of approaches to quantifying the differences and similarities in a set of numbers. One of the most obvious methods is significance testing using the χ^2 statistic, which can be used to reject the hypothesis that the PFs or EPRs are similar. However, significance testing depends on the calculation of p-values, which notoriously underestimate Type I error rates (Sellke et al., 2001). This underestimation of "false positive" results stems mainly from a lack of consideration of the alternative hypothesis. Also, many of the queries discussed are plagued by low statistics; for small data sets, the χ^2 method would simply fail to reject the hypothesis that the PFs or EPRs were the same and give no information concerning the degree to which they were similar.

The maximum likelihood ratio (LR) method is a technique that avoids some of the pitfalls of the χ^2 technique, as it considers the likelihoods of both similarities and differences. However, it cannot directly quantify the likelihood that the PFs or EPRs are different because such a hypothesis does not exist for the LR test.[1]

1 For the same reason the hypothesis ``the coin is not fair" cannot be quantified within any frequentist approach.

The two tests described above are examples of frequentist approaches to statistical inference, and their failures are endemic to the frequentist approach in general. On the other hand, the Bayesian formalism, derived from logic (Cox, 1961), circumvents the flaws of the frequentist methods by using a direct calculation of the probability that, for example, the EPRs are consistent within statistical uncertainty across different hospitals. More importantly, it calculates a maximally intuitive statistic measure that can be easily interpreted with no statistics expertise. Although it is not the main purpose of this work to argue the value of the Bayesian formalism over other approaches, we demonstrate an instance where the Bayesian approach matches our aim to visualize intuitively whether PF's and EPR's are different, as well as the meaning and magnitude of any differences that exist.

This paper describes our journey from visualization need to the Bayesian mathematical formalism that supports it. Our procedure compliments the common initial step of determining the visualization technique from the represented objects and operations (Wehrend and Lewis, 1990) or through data types and the task-domain information actions that the user wishes to perform (Shneiderman, 1996). Instead, we start with a visualization need and find a statistical method to support it. The work here is most similar to Tukey's exploratory analysis (Tukey, 1977), which includes visualization techniques to extract potentially useful hypotheses from data. In this work, we describe a procedure of determining the details of the hypothesis-based data manipulations necessary to meet the visualization goals, which are anomaly identification tasks in the taxonomy of Amar et al. (Amar et al., 2005). Such a procedure would not be necessary in the instance where the visualization scheme follows directly from the information to be visualized. For instance, Farrah et al. (Farrah et al., 2009) used a mathematical framework similar to the one described here in order to quantify the similarities and differences in galactic spectra. In their case, the sheer volume of data and the nature of the comparisons begged for a certain visualization scheme (a network diagram). Likewise, Zender et al. (Zender et al., 2010) presented a Visual Language System to represent (complex) concepts in medicine; the system used icons and glyphs to ease the interpretation of raw data. In contrast, a methodology like the one described here would be followed in specific cases where, say, the design space of visualization tasks described in Schulz et al. (Schulz et al., 2013) was created with little regard for the details of the data manipulations required to ultimately produce the visualization.

While the objective of this work is to seek a mathematical structure to support the visualization of simple differences between many numbers, the calculations of their significance is, in fact, necessarily complex. That is, the complexity derives from the need for an intuitive visualization.

The article is organized as follows: Section 2 (Methods) describes the mathematical underpinnings of the visualization. Section 3 (Data) describes the data that will be simulated to demonstrate the methodology. Section 4 (Results and Discussion) presents a discussion of results. Section 5 (Conclusion) contains the conclusions.

2. METHODS

In this section, we will describe in detail how the Bayesian approach can provide the desired representations and how the query results will be presented. The calculations described below are implemented within a Perl script available at http://sourceforge.net/projects/databasequalitymanagement. Detailed documentation of the source code and mathematics can be found in the same location.

2.1 BAYESIAN FORMALISM

The Bayesian formalism allows one to calculate, intuitively and concisely, the two similarity measures previously discussed: The first quantifies the probability that, within statistical uncertainties, the fractional patient frequencies (PFs) among the hospitals are consistent with the fractional PFs in the base pools (BPs). The second is the probability that the encounter-to-patient ratios (EPRs) are similar across all hospitals. Within the Bayesian formalism, there are no statistical tests or tests of significance. The probability is simply the probability or belief that the distributions or ratios would be the same were it not for statistical fluctuations. The probability calculated requires no statistics expertise: it is as intuitive as the probability of rain showers in a weather forecast. Details of the mathematical framework follow.

2.1.1 PF Similarity

Hospital similarity in Patient Frequency (PF) is defined by the relative size of their Base Pools (BP). For instance, suppose that a query returned the PFs of 91, 81, and 220 for three hospitals, CCHMC, CHCO, and CHOP, respectively. Suppose further that the corresponding BPs were 1,000, 2,000 and 3,000 patients, respectively. We wish to quantify whether there is a statistically significant difference between the PF ratios 91:81:220 and BP ratios 1,000:2,000:3,000. That is, we want to know if we were to somehow replicate these three hospitals infinitely many times, what is the probability that the distribution of counts averaged over all the replications would be 1,000:2,000:3,000? In this context, "different" can mean either all but one hospital are consistent with the BP ratio, or all hospitals are inconsistent with the BP ratio.

Formally, the (posterior) probability that the PFs are similar is calculated using Bayes' theorem, which is the key equation in the Bayesian formalism:

$$P(\text{same}|\vec{Q}) = \frac{P(\vec{Q}|\text{same})P(\text{same})}{P(\vec{Q}|\text{same})P(\text{same}) + \sum_i P(\vec{Q}|\text{i}^{\text{th}}\text{ hosp is diff})P(\text{i}^{\text{th}}\text{ hosp is diff}) + P(\vec{Q}|\text{all diff})P(\text{all diff})}$$

where, generally, $P(\vec{Q}|X)$ is a likelihood and $P(X)$ is a prior, which are discussed in more detail below. The likelihood encapsulates the probability of obtaining the observed outcome, $\vec{Q}$, given the "same" or "different" hypothesis. For instance, the χ^2 statistic is directly proportional to the likelihood. The priors encapsulate knowledge or assumptions about the hypotheses before $\vec{Q}$ is measured. If no assumptions are made (for example, if this is the first-ever observation of anything like $\vec{Q}$), then the priors related to the "same" and "different" hypotheses are set equal to one another.

$P(\vec{Q}|\text{same})$ is the likelihood of obtaining a set of PFs, $\vec{Q}$, with the assumption they are similar. This term takes the form of a multinomial distribution where the probabilities are set according to the relative sizes of the BPs. The counts in the multinomial distribution are the PFs from the query.

$P(\vec{Q}|\text{all diff})$ is the likelihood of obtaining the PFs given the distribution of PFs is different from the distribution of PFs in the BP. Quantitatively, the hypothesis is that the PFs are not accurately determined by the BPs, and so every distribution is equally probable. Therefore, this probability also takes the form of a multinomial distribution, except the multinomial probabilities are summed over (integrated).

$P(\vec{Q}|\text{i}^{\text{th}}\text{ hosp diff})$ is similar in form to $P(\vec{Q}|\text{all diff})$, with the added constraint that all the hospitals except the ith hospital must be consistent with the BP PF distribution.

$P(\text{same})$ is the prior probability or expectation that all the hospitals are similar. Since we assume a priori that it is equally probable that the hospitals for a particular query are the same, then $P(\text{same}) = 1/2$.

$P(\text{all diff})$ and $P(\text{i}^{\text{th}}\text{ hosp diff})$ $P(\text{all diff})$ is the prior probability or expectation that all the hospitals are different. $P(\text{i}^{\text{th}}\text{ hosp diff})$ is the prior probability or expectation that all but one of the hospitals is the same. It is assumed a priori that the probability that the hospitals are different in some way is 1/2 and that it is equally probable for all the hospitals to be different or for the ith hospital to be different. Therefore, for $N_{\text{hospitals}}$ hospitals, $\text{P(all diff)} = \text{P(ith hosp is diff)} = 0.5/(N_{hospitals} + 1)$.

The Bayesian formalism empowers the formation of the limited number of conceptual categories above which in turn provide the visualization with a limited number of conceptual categories relevant to the questions of interest: are hospitals the same or different; are all hospitals different or just one; is any difference higher or lower? These are the questions the visualization seeks to answer and which the statistical framework provides.

2.1.2 EPR Similarity

Encounter-to-patient-ratio (EPR) similarity is quantified by the probability that, within statistical uncertainties, the EPRs of the three hospitals are the same. For instance, suppose a query produced EPRs 5:1, 6:1 and 8:1 for the three hospitals under consideration. The probability that they are similar is the probability that the differences are due only to statistical fluctuations. That is, if we were to somehow replicate these three hospitals many times, it is the probability that the EPRs averaged over all the replications would be the same.

To calculate the EPR similarity, we again calculate Eqn. 1 but with terms that have an entirely different meaning. The idea behind the calculation is to express the patient and encounter counts as if they were binomially distributed, with the constraint that the probability of an "encounter" is always more than that of a "patient". This constraint follows from the fact that there are always at least as many encounters as patients. The Bayesian formalism allows this assumption to be neatly folded into the calculation of the final (posterior) probability in Eqn. 1. We then determine the probability that the relative probability of a "patient" versus an "encounter" for all the hospitals is the same.

A detailed description of each term follows.

$P(\vec{Q}\,|\,\text{same})$ is the probability that, within statistical fluctuations, the EPRs of the hospitals ($\vec{Q}$) are the same. Mathematically, $P(\vec{Q}\,|\,\text{same})$ is a product of binomial distributions which are integrated under this hypothesis and under the constraint that the chance of an encounter must be larger than that of a patient, hereafter called the EPR constraint.

$P(\vec{Q}\,|\,\text{all diff})$ is the probability that, beyond statistical fluctuations, the EPRs of the hospitals are different. Again, this term is a product of binomial distributions which are integrated individually under the EPR constraint.

$P(\vec{Q}\,|\,\mathrm{i^{th}\ hosp\ diff})$ is the probability that, within statistical fluctuations, the EPRs of all but the ith hospital are the same under the EPR constraint. Again, this term is a product of binomial distributions which are integrated individually under this constraint.

$P(\mathrm{same})$, $P(\mathrm{all\ diff})$, and $P(\mathrm{i^{th}\ hosp\ diff})$ are probabilities that retain the same definitions as those in the previous section.

NOTE: both the PF and EPR similarity measures allow for the number of patients to be zero.

2.1.3 Determination of Outlying Hospitals, Excesses and Deficits

If the hospitals are different (i.e., $P(\vec{Q}\,|\mathrm{same}) < P(\vec{Q}\,|\mathrm{all\ diff})P(\mathrm{all\ diff}) \cdot + P(\vec{Q}\,|\mathrm{i^{th}\ hosp\ is\ diff})P(\mathrm{i^{th}\ hosp\ is\ diff})$) then the visualization scheme requires that we determine how they are different. Specifically, we must determine if either (1) all the hospitals are different, or (2) all the hospitals are similar except one. This determination is made by comparing the probability that all the hospitals are different, $P(\vec{Q}\,|\,\mathrm{all\ diff})$, to the probability that all the hospitals except for the ith one are similar. If it is more probable that a single hospital is different from the others (i.e., $P(\vec{Q}\,|\mathrm{i^{th}\ hosp\ is\ diff}) > P(\vec{Q}\,|\mathrm{all\ diff})$) then the outlying hospital is found by determining that i which maximizes $P(\vec{Q}\,|\,\mathrm{i^{th}\ hosp\ diff})$.

If there exists an outlying hospital, all that remains is to determine whether or not the outlier is the result of an excess or deficit. This is performed with simple subtraction. If the outlying PF is the result of an excess (deficit), its fractional PF will be larger (smaller) than the fractional PFs in the other BPs. Meanwhile, if the outlying EPR is the result of an excess (deficit), it will simply be larger (smaller) than the other EPRs. Visualization-wise, the corresponding up or down arrow colors are matched to the color in columns 2 or 3 in Figure 1a, depending on whether the outlying deviation is in PF or EPR respectively.

2.2 VISUALIZATION

The purpose of the visualization scheme is to provide a way to efficiently determine whether or not the PFs or EPRs from different hospitals differ, and, if so, how. The Bayesian formalism supplies the results in categories. To concisely and intuitively visualize these categorical results, we choose a tabulated format, shown in Figures 1a and 1b, which show two possible visualizations. Fig. 1a will be the focus of this work, while Figure 1b demonstrates an alternative representation of essentially the same information.

In Figure 1a, each row represents an independent query. The columns are discussed in turn. The far left column shows the fact to be queried, although in the simulated case presented here the queries are simply denoted Query 1, Query 2, etc. The cells in the next two columns display six boxes; the number of filled boxes is proportional to the inter-hospital similarities of the PFs and EPRs respectively. Six boxes are chosen based on the need for an even number of boxes (to discern a probability over/under 50%) while maximizing the number of boxes than can be quickly enumerated.[2] Quickly discerning differences in the probabilities across the diagrams was the main concern when determining how the probability would be represented. A horizontal row of boxes offered an intuitive way to visually measure difference quantity rather than a single continuous bar.

2 Six objects can be enumerated in about 1 second (Trick and Pylyshyn, 1994).

The visualization of the probability is reinforced by the color of the filled boxes. The boxes are colored green if the probability of similarity is between 50% and 100%, yellow if it is between 5% and 50%, and red if it is below 5%. Hovering the cursor over the boxes within these columns reveals a call-out with the precise probability that the PFs or EPRs are similar, as demonstrated in Query 4 with the black arrow cursor. The remaining columns contain the PFs and EPRs for each hospital, where each PF cell contains both the PF as well as the corresponding percentage of patients this number represents from the BP. (This percentage is useful in determining whether statistically significance differences are, in fact, important.)

In the case where all but one of the PFs or EPRs are similar, an up or down arrow indicates an excess or deficit, respectively, in the "outlying" hospital. The colors of the cells with these arrows match the color of the corresponding second or third row. For instance, row 2 shows a statistically significant excess in CHOP's PF, which is indicated with an upward-pointing arrow and a yellow cell which is matched to the yellow PF cell in column 2. In the case where all the PFs or EPRs are different from each other, bi-directional arrows fill all three cells, which are color-matched to columns 2 or 3, respectively.

It is worth emphasizing that the colors mostly serve to reinforce what is visualized in other ways. As each cell in columns 2 and 3 has at least 1 box filled, the minimum discernible probability from the visualization for a color-blind investigator would be 16%, in contrast to 5% for one who is able to differentiate red from yellow. Regardless, a precise probability can always be found by scrolling over the cell in column 2 or 3.

However, if it is less important to visualize the degree of similarity than it is to identify differences, an alternative representation like the one in Figure 1b may suffice. In this visualization, differences are simply indicated by red cells, and the degree of similarity is indicated by the number of arrows. Two arrows (on top of one another) indicate a probability of similarity less than 5%, and one arrow indicates a probability of similarity between 5% and 50%. If the PF (EPR) cell is white, the number or ratio is assumed to be consistent with the BP (other ratios). A precise probability of similarity can still be found by scrolling over any red cell, as indicated by the cell in Query 2. However, our discussion below will focus on the visualization in Figure 1a.

Regardless, the following information must be visualized for a particular query:

1. Whether or not the distribution of PFs across the hospitals is significantly different from the distribution of PFs in the BPs,
2. Whether or not differences in the EPRs are statistically significant,
3. If there are differences between the hospitals, whether all the hospitals are different from one another, or all but one of the hospitals are similar,
4. If all but one of the hospitals are similar, which one is the outlying hospital,
5. Whether an outlying hospital contains an excess or deficit,
6. A visual representation of the similarity measure for PFs and EPRs.

3. DATA

The visualization is demonstrated with simulated queries that are nearly identical to ongoing queries of epilepsy databases at CCHMC, CHOP, and CHCO. The only differences between the simulated and actual queries are the values of the PFs and EPRs; the visualizations and magnitudes of the PFs and EPRs are the same.

We "simulate" 7 database queries of epilepsy patients from three hospitals. The query results are characterized by well-defined inter-hospital variabilities, which are discussed in turn.

Query 1

The PF distribution across the hospitals is the same as the distribution of BP PFs, and the EPRs are the same for all the hospitals.

Query 2

The EPRs are the same for all hospitals. The PF distribution is the same as the BP PF distribution, with the exception of one hospital (CHOP), which has a 100% excess in PF.

Query 3

The EPRs are the same for all hospitals. The PF distribution is the same as the BP PF distribution, with the exception of one hospital (CHOP), which has a 95% deficit in PF.

Query 4

The PF distribution across hospitals is the same as the distribution of BP PFs. The EPRs are the same for all hospitals except one, which has a 100% excess in the EPR.

Query 5

The PF distribution across hospitals is the same as the distribution of BP PFs. The EPRs are the same for all hospitals except one, which has a 95% deficit in the EPR.

Query 6

The EPRs are all different, but the PF distribution across hospitals is the same as the BP PFs.

Query 7

The EPRs for all the hospitals are the same, but the PF distribution across hospitals is different from the BP PFs.

The number and ordering of the queries in Figure 1a make the diagram appear deceptively clean. The real data may contain hundreds of queries with seemingly randomly distributed probabilities. On the other hand, the diagram will be more ordered for real data from the clinical point of view, as the queries can be grouped according to type. For instance, all queries concerning anti-epileptic drugs can be grouped to help elucidate relations between (dis)similarities in patients and encounters.

Data Quality Report

Field Name	Patient Freq Similarity	Encounters/Patients Similarity	CCHMC Patient Freq (% of Population)	CCHMC Encounters/Patients	CHCO Patient Freq (% of Population)	CHCO Encounters/Patients	CHOP Patient Freq (% of Population)	CHOP Encounters/Patients
Total Population	N/A	■■■■■■	1000	3.00	3000	3.00	2000	3.00
Query 1	■■■■■■	■■■■■■	10 (1.0 %)	10.00	30 (1.0 %)	10.00	20 (1.0 %)	10.00
Query 2	■■■□□□	■■■■■■	10 (1.0 %)	10.00	30 (1.0 %)	10.00	40 (2.0 %) ↑	10.00
Query 3	■□□□□□	■■■■■□	10 (1.0 %)	10.00	30 (1.0 %)	10.00	1 (0.1 %) ↓	10.00
Query 4	■■■■■ 0.57	■■■■□□	10 (1.0 %)	10.00	30 (1.0 %)	10.00	20 (1.0 %)	20.00
Query 5	■■■■■■	■□□□□□	10 (1.0 %)	10.00	30 (1.0 %)	10.00	20 (1.0 %)	1.00 ↓
Query 6	■■■■■■	■□□□□□	10 (1.0 %)	10.00 ↕	30 (1.0 %)	23.33 ↕	20 (1.0 %)	1.00 ↕
Query 7	■□□□□□	■■■■■■	30 (3.0 %) ↕	10.00	60 (2.0 %) ↕	10.00	1 (0.1 %) ↕	10.00

Showing 1 to 8 of 8 entries

FIGURE 1A:

This visualization is the focus of this work. Each row represents an independent query. The far left column shows the fact to be queried, although in the simulated case presented here the queries are simply denoted Query 1, Query 2, etc. The cells in the next two columns display a series of boxes; the number of filled boxes is proportional to the inter-hospital similarities of the PFs and EPRs respectively. The visualization of the probability is reinforced by a common-sense color scheme for the filled boxes. The boxes are colored green if the probability of similarity is between 50% and 100%, yellow if it is between 5% and 50%, and red if it is below 5%. Hovering the cursor over the boxes within these columns reveals a call-out with the precise probability that the PFs or EPRs are similar, as demonstrated in Query 4 with the black arrow cursor. The remaining columns contain the PFs and EPRs for each hospital, where each PF cell contains both the PF as well as the corresponding percentage of patients this number represents from the BP. In the case where all but one of the PFs or EPRs are similar, an up or down arrow indicates an excess or deficit respectively in the "outlying" hospital. The colors of the cells with these arrows match the color of the corresponding second or third column.

Data Quality Demo Report

Field Name	CCHMC Patients (% of Population)	CCHMC Encounters/Patients	CHCO Patients (% of Population)	CHCO Encounters/Patients	CHOP Patients (% of Population)	CHOP Encounters/Patients
Query 1	10	10.00	30	10.00	20	10.00
Query 2	10	10.00	30	10.00	40	10.00
Query 3	10	10.00	30	10.00	1	10.00
Query 4	10	10.00	30	10.00	20	20.00
Query 5	10	10.00	30	10.00	20	1.00
Query 6	10	10.00	30	23.33	20	1.00
Query 7	30	10.00	60	10.00	1	10.00

Showing 1 to 7 of 7 entries

FIGURE 1B: This visualization shows an alternative representation where differences are simply indicated by red cells. The degree of similarity is indicated by the number of arrows. Two arrows (on top of one another) indicate a probability of similarity less than 5%, and one arrow indicates a probability of similarity between 5% and 50%. If the PF (EPR) cell is white, the number or ratio is assumed to be consistent with the BP (other ratios). A precise prob¬ability of similarity can still be found by scrolling over any red cell, as indicated by the cell in Query 2.

4. RESULTS AND DISCUSSION

We demonstrate how the Bayesian formalism allows us to meet the visualization purpose described in Section 2.2, including the requirement for an intuitive representation of the degree of inter-hospital similarity. We also demonstrate that the representation of probabilities makes the results of the seven queries described in Section 3 quite clear.

We first discuss the relation between the probability visualization and variability of the PFs and EPRs for these queries, as well as the unqueried "Total Population" in Figure 1b, in turn.

Query 1

The green color codes in Columns 2 and 3 are exactly what we expect from a query that renders PF proportions that reflect the PF proportions in the BP. The probabilities of both the PFs and EPRs appear as green, with probabilities of similarity 97% and 91%, respectively.

Queries 2 and 3

The EPRs are the same while CHOP shows an excess and deficit in PFs relative to the BPs for the two queries. As expected, the EPR probabilities are color-rendered green for both queries. Also as expected, the PF probabilities for the two queries show low probabilities of similarity indicated by yellow and red, respectively. The former probability turns out to be substantially larger of the two (34% versus 0.06%). Although the magnitude of the deviations from maximum similarity in the CHOP queries are similar (a 20 count excess and a 19 count deficit), the differences in their standard deviations allow for disparate statistical fluctuations in the EPRs. For instance, the standard deviation in PF for Query 2 is ~6 counts, while the standard deviation in PF for Query 3 is only 1.

Queries 4 and 5

These queries show an excess and deficit in EPRs while fixing the proportion of PFs to those in the BPs. Query 5 gives the expected red box in column 3, while it is green for Query 4 with a 57% probability. However, the high probability for Query 4 arises from two factors that suggest that such a statistical fluctuation is possible. First, a statistical fluctuation in the numbers of patients and encounters that is smaller than one standard deviation brings the EPRs into agreement. Second, the number of hospitals is folded into the probabilities. For instance, the chance of finding a one-standard-deviation fluctuation is greater in three hospitals than in two. By contrast, the same agreement does not exist in Query 5.

Queries 6 and 7

These queries render different PFs and EPRs, and give the expected representations of similarity in columns 2 and 3.

Total Population

This row contains the PFs and EPRs for the BPs; it is the unqueried data set. Quantifying the similarity of patients for the BPs is uninteresting as their PFs define the gold standard for similarity when comparing inter-hospital PFs. An "N/A" is therefore inserted into the second column of this row. On the other hand, quantifying the similarity of the EPRs is crucial as differences in their values may reflect actual differences in data quality or care. Here the EPRs for the BPs are the same, which is supported by the high probability of similarity indicated by the green box in column 3. A method of comparing BP and query results to isolate inter-hospital differences is discussed in Section 5.

Note that although the probability visualizations are reinforced with a red-yellow-green color scheme, the colors are not critical for interpretation of results. This design decision for both Figures 1a and 1b arose from the fact that about 8 percent of males, and 0.5% of females, have some degree of congenital color deficiency (Deeb and Motulsky, 2012), which requires a careful consideration of color palettes when designing visualization schemes for representation of meaning, causality, and importance of query comparison.

5. CONCLUSIONS

We showed that the challenges of visualizing inter-hospital similarities and differences in patient database queries can be met by choosing the appropriate formalism to quantify them. In so doing, we demonstrated the potentially symbiotic relationship between statistics and visualization. The chosen (Bayesian) formalism allowed us to compute concise and intuitive similarity measures that fit well into the desired visualization scheme. The ability of our formalism to include many factors, such as statistical fluctuations and prior information, not only suggests the need for a direct calculation of the probability of similarity, but also hints at how the probability measure compliments the visualization procedure. That is, a visual/mathematical formalism shown here goes beyond simply flagging numbers that fall within categories with arbitrarily set limits. Also, the measures of similarity used within the visualization scheme described here do not require any statistics expertise for interpreting them. Further, the formalism is extremely flexible and even allows the computation of similarity measures when no patients are found in one or all the hospitals. These measures are implemented within publically available software package. The software along with detailed documentation of the source code and mathematical framework can be found at http://sourceforge.net/projects/databasequalitymanagement.

However, the formalism does contain a number of subtle approximations. For instance, the binomial and multinomial distributions technically allow for the number of patients to be greater than the number of encounters. One can certainly introduce additional constraints to prevent this unphysical situation, but the mathematics then becomes non-analytic.

It must also be argued that some of the test queries yielded rather non-intuitive results (seemingly questionable representations); however, they were shown to be reasonable when it was revealed that statistical fluctuations and the number of inter-hospital comparisons were factored into the measure of similarity.

Sophisticated analyses to isolate the sources of inter-hospital differences becomes possible with the presented formalism. For instance, if the EPRs for the BPs are different but the PFs and EPRs for a query are the same, then the query's "fact" is unlikely to be the source of the EPR differences in the BPs. Isolating differences is further facilitated by the ability to point to outlying hospitals that cause differences in PFs and EPRs.

Visualization allows for these and similar analyses to be performed on a large scale over many queries and many similarity measures. With this in mind, there remain many possible extensions to the query comparisons presented in this work. For instance, comparing EPRs was important to understanding possible differences in care or reporting of care. However, various factors may contribute to inter-hospital variability in the number of clinical notes for a given procedure. These inter-hospital similarities and differences can be explored by comparing "encounters":"clinical notes":"patient frequencies" ratios and by changing the binomial distributions in Section 2.1.2 into tri-nomial distributions. Adding similarity measures, as well as increasing the sample size via bringing in additional hospitals, will result in a larger diagram. Such a diagram will be more challenging to design and display in an easily readable manner.

Analysis could also be facilitated by ordering the queries. For instance, it can be imagined that ordering by degree of inter-hospital dissimilarity would facilitate the identification of "problematic" queries. However, differences in data quality and care are easier to identify if the queries are grouped by, say, types of drug refills or diagnostic tests.

Better, the queries could be ordered by relevance using a survey strategy similar to the one described in (Pestian et al., 2013) where cluster analysis was combined with expert opinion to identify subject matter "groups" and weight those groups by importance. Regardless, tracing the sources of differences could be facilitated by providing a button that would enable one to bounce between various ordering schemes.

The identification of important differences will be facilitated through clinical expertise; while certain differences in values may be statistically significant, they may not be important. The Bayesian formalism, in fact, allows one to fold in such allowable ranges by imposing a range of allowable proportions. This range is folded into the probability (specifically, $P(\vec{Q} \mid \mathrm{same})$) by integrating over that range.

The number of hospitals can also be varied by changing the order of the multinomial distribution in the PF comparisons, as well as the number of binomial distributions in the EPR comparisons. Further, PFs and EPRs can be allowed to differ beyond statistical fluctuations

The more profound problem of the so-called "look elsewhere effect" should also be folded into the calculations/visualizations. This problem arises when, over many queries, the inevitable occasional large statistical fluctuation is misinterpreted as a significant deviation. This problem is solved in the classical (frequentist) approach by imposing corrections on p-values (for example, Aickin and Gensler, 1996). However, the Bayesian approach allows one to naturally normalize the probabilities to account for multiple queries. Regardless of the statistical approach, folding this effect within the similarity measures serves to compliment the visualization scheme by preventing chance fluctuations in the data from appearing as statistically significant differences and thereby preventing false findings.

ACKNOWLEDGEMENTS

The authors would like to thank the three anonymous reviewers for their valuable comments and suggestions, which improved the quality of the paper.

REFERENCES

Aickin, M. and Gensler, H. (1996). Adjusting for multiple testing when reporting research results: the Bonferroni vs Holm methods. *American Journal of Public Health*, 86(5):726–728.

Amar, R., Eagan, J., and Stasko, J. (2005). Low-level components of analytic activity in information visualization. In *IEEE Symposium on Information Visualization*, 2005., pages 111–117. IEEE.

Birkmeyer, J. D., Reames, B. N., McCulloch, P., Carr, A. J., Campbell, W. B., and Wennberg, J. E. (2013). "Understanding of regional variation in the use of surgery." *The Lancet*, 382(9898):1121–1129.

Cox, R. T. (1961). *The algebra of probable inference*. Johns Hopkins Press Baltimore.

Deeb, S. S. and Motulsky, A. G. (2012). *Disorders of Color Vision in Genetic Diseases of the Eye*. Oxford University Press, New York, second edition.

Farrah, D., Connolly, B., Connolly, N., Spoon, H., Oliver, S., Prosper, H., Armus, L., Houck, J., Liddle, A., and Desai, V. (2009). "An evolutionary paradigm for dusty active galaxies at low redshift." *The Astrophysical Journal,* 700(1):395.

Pestian, J., Matykiewicz, P., Holland-Bouley, K., Standridge, S., Spencer, M., and Glauser, T. (2013). "Selecting anti-epileptic drugs: a pediatric epileptologist's view, a computer's view." *Acta neurologica Scandinavica*, 127(3):208–215.

Schulz, H.-J., Nocke, T., Heitzler, M., and Schumann, H. (2013). "A design space of visualization tasks." *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2366–2375.

Sellke, T., Bayarri, M., and Berger, J. O. (2001). "Calibration of ρ values for testing precise null hypotheses." *The American Statistician*, 55(1):62–71.

Shneiderman, B. (1996). "The eyes have it: A task by data type taxonomy for information visualizations." In *Visual Languages, 1996. Proceedings., IEEE Symposium on,* pages 336–343. IEEE.

The Dartmouth Atlas Working Group (2013). *The Dartmouth Atlas of Children's Health Care in Northern New England*. Hanover, NH.

Trick, L. M. and Pylyshyn, Z. W. (1994). "Why are small and large numbers enumerated differently? a limited-capacity preattentive stage in vision." *Psychological Review*, 101(1):80.

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley, Reading, Mass.

Wehrend, S. and Lewis, C. (1990). "A problem-oriented classification of visualization techniques." In *Proceedings of the 1st Conference on Visualization'90*, pages 139–143. IEEE Computer Society Press.

Zender, M., Pestian, J., and Glauser, T. (2010). "Visual language system for representing medical concepts." *Information design journal + document design*, 18(3):184–197.

Shaikh, A. D. (2007). *Psychology of on-screen type: investigations regarding typeface personality, appropriateness, and impact on document perception*. Department of Psychology, Wichita State University. PhD.

Waller, R. (2012). "Graphic literacies for a digital age: the survival of layout." *The Information Society: an international journal* 28(4): 236–252.

Wendt, D. (1968). "Semantic differentials of typefaces as a method of congeniality research." *The Journal of Typographic Research* II(1): 3–25.

Book Review:

Design for information, an introduction to the histories, theories, and best practices behind effective information visualizations.

by

Isabel Meirelles

Beverly, MA: Rockport Publishers, 2013.

Jorge Frascara

Design for information is a thorough representation of both the field of information visualization and the research interests of the author, whose focus is on "the theoretical and experimental examination of the fundamentals underlying how information is structured, represented and communicated in different media."

Beginning with the "big picture," the book includes an amazing collection of examples, the most thorough I have seen to date in a volume. The author organizes the content according to several categories represented by the titles of the chapters: 1. Hierarchical structures: trees; 2. Relational structures: networks; 3. Temporal structures: timelines and flows; 4. Spatial structures: maps; 5. Spatio-temporal structures; and 6) Textual structures. An appendix, notes, a bibliography, a contributors list, and an index complete the content of the book.

Design for information is an extensive taxonomy of data visualization types and is "a must" for anybody interested in the work done in the area. Each one of the hundreds of examples is explained and discussed, forming a kind of encyclopedia on the subject. It seems that nothing escaped the exemplary collection that Meirelles assembled. The discussions and explanations normally focus on what information is represented and how it is represented.

It is interesting to see, as well, how many different professional fields today use diagrams to organize and represent information: basic science, applied science, education, engineering, medicine, and technologies, etc. The value of the book is centered on the inclusion of examples of how many different problems are now being addressed through data visualizations, how many historical efforts preceded whatsoever is done today, and how the advent of the computers have allowed the field to explode by handling large data sets as well as dynamic representations.

At the end of the examination of the 224-page volume I became curious as to how these diagrams might have performed with the users they were intended for in terms of ease of comprehension; what conclusions I could arrive at from an evaluation of the examples from a perceptual and cognitive human factors perspective; or how a complementary book could contribute to the development of best practices. I would not expect that one volume could be so extensive as this one and also cover the field critically. However, I have to wonder how the super-complex visualizations permitted by computer programs today would perform regarding comprehension, memorization, and use of the information presented. The discussion on perception and cognition is very brief, and it might leave some readers wondering about the assertions made: they are proposed as principles without them being discussed. This topic, as well as Gestalt theory, are not considered during the description of examples. The size of some reproductions is too small to assess their quality as data visualizations. They appear as examples of problems addressed but not as information in themselves. To compensate for this, the book includes valuable URLs for people interested in seeing in better detail many of the diagrams shown.

While the above issues could be perceived as weaknesses, the strength of the book is its truly amazing array of examples and the rare historical diagrams it offers. It also displays an uncommon erudition and includes an extensive and useful bibliography. I do not know how long Meirelles took to complete the manuscript, but it feels like a lifetime project. These assets, coupled by excellent production, make this an indispensable publication for anyone interested in information visualization.

Book Review:

Isotype: Design and contexts 1925–1971

by

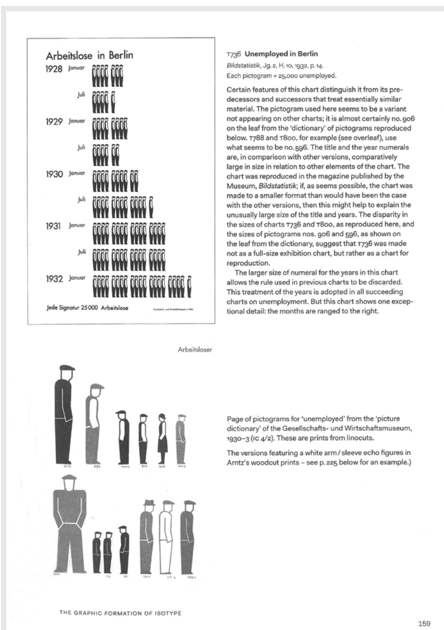
Burke, Christoher; Kindel, Eric; Walker, Sue (Eds.)

Hyphen Press, London, 2013

Per Molleup

Isotype, why not?

The term Isotype, an acronym for International System Of TYpographic Picture Education, is a technique of data visualization introduced by sociologist, economist, and philosopher Otto Neurath. Originally called "The Vienna School of Pictorial Statistics" and developed and practiced at the Gesellschafts- und Wirtschafts-museum in Wien (Social and Economic Museum of Vienna), 1925-34, Isotype's purpose was to communicate societal information to visitors. In 1935, Neurath's technique was renamed, and Isotype began its own life and was used for other purposes in other places.



T736 **Unemployed in Berlin**

Bildstatistik, Jg. 2, H. 10, 1932, p. 14.

Each pictogram = 25,000 unemployed.

Certain features of this chart distinguish it from its predecessors and successors that treat essentially similar material. The pictogram used here seems to be a variant not appearing on other charts; it is almost certainly no. 906 on the leaf from the 'dictionary' of pictograms reproduced below. T788 and T800, for example (see overleaf), use what seems to be no. 596. The title and the year numerals are, in comparison with other versions, comparatively large in size in relation to other elements of the chart. The chart was reproduced in the magazine published by the Museum, *Bildstatistik*; if, as seems possible, the chart was made to a smaller format than would have been the case with the other versions, then this might help to explain the unusually large size of the title and years. The disparity in the sizes of charts T736 and T800, as reproduced here, and the sizes of pictograms nos. 906 and 596, as shown on the leaf from the dictionary, suggest that T736 was made not as a full-size exhibition chart, but rather as a chart for reproduction.

The larger size of numeral for the years in this chart allows the rule used in previous charts to be discarded. This treatment of the years is adopted in all succeeding charts on unemployment. But this chart shows one exceptional detail: the months are ranged to the right.

Page of pictograms for 'unemployed' from the 'picture dictionary' of the Gesellschafts- und Wirtschaftsmuseum, 1930–3 (IC 4/2). These are prints from linocuts.

The versions featuring a white arm / sleeve echo figures in Arntz's woodcut prints – see p. 225 below for an example.)

THE GRAPHIC FORMATION OF ISOTYPE

159

Page 158
Males with flat cap and hands in pocket stereotype unemployed persons. Striking workers have crossed arms.

Isotype builds on the idea expressed in Neurath's often-repeated adage, "To remember simplified pictures is better than to forget accurate numbers" (p. 85). Therefore, Isotype is best known for Picture tables—graphic displays with rows of repeated pictograms each standing for a number of real world units. The picture tables embody the proposition that it is easier to compare quantities by comparing numbers of well-presented symbols, than to compare symbols of different size. Pictograms in the Isotype picture tables are scaled: in a display showing unemployment, each pictogram would stand for 1,000; 100,000; or 1 million – or another round number of unemployed persons. In picture tables, the reader must count the pictograms in different groups and multiply with the scaling factor to get the total amounts. The number of the repeated pictograms in a picture table is most often rounded off. Some Isotype picture tables feature half, quarter, or smaller fractions of pictograms. Even then, Isotype displays are typically not as precise as the numbers they represent.

Isotype: Design and contexts 1925-1971 describes Isotype in a period delimited by 1925 when the Gesellschafts-und Wirtschaftsmuseum in Wien was founded and 1971 when the Isotype Institute in London closed. The book comprises 12 chapters dealing with the genesis and further development and design of Isotype. The book includes two kinds of information; it describes Isotype design principles, and it describes the process by which Isotype was developed and disseminated. To this reviewer the former part is the most interesting, while the latter part serves as a useful historical backdrop.

TEACHING MUSEUM

In the first of two central chapters, Christopher Burke covers the ten-year lifetime of The Gesellschafts-und Wirtschaftsmuseum in Wien including the formative years of Isotype. The idea behind Isotype predated the museum. Otto Neurath, sociologist, economist, and philosopher, had already applied charts with pictorial descriptions of quantities for the Museum für Siedlung und Städtebau, Museum for Settlement and Town Planning. Eager to educate by means of these didactic tools, Otto Neurath suggested a new museum to expand the ordinary population's understanding of national and world relations. The social democrat regime in Vienna understood the importance of education and provided the necessary financial support.

Otto Neurath invented Isotype, but more than that, he promoted it. Philosophically and organisationally trained in addition to being well connected academically and politically he spread the word and established the connections that were vital to the incubation of a new idea like Isotype. Neurath partnered his strong interest in education of ordinary people with his equally strong social commitment resulting in his belief that progress depends on knowledge, and knowledge should be delivered in ways that are both attractive and memorable – essential qualities of Isotype.

The Gesellschafts und Wirtshaftsmuseum in Wien was not a museum in the traditional sense of that word; and therefore consistent with Neurath's view that "The modern museum should be a teaching museum, a means of education, a schoolbook on a grand scale…" (p. 30). The Gesellschafts and Wirtshaftsmuseum in Wien consisted primarily of graphic charts explaining societal matters, first and foremost quantities. The museum introduced a number of advanced ideas to meet its audience. To accommodate prospective visitors the museum was open evenings and Sunday mornings. Also, the museum branched out at several places where visitors would be. At a certain time, the museum would exhibit at several different locations in Vienna including the town hall. A central corner shop museum open in the lunch hours had as many as two thousand visitors a day. At the corner shop special interactive knowledge machines were placed where visitors could test their knowledge – anticipating a distant, digital future. Exhibition material was reused and exchanged between permanent and time limited exhibitions in several places. Along with its own exhibitions, the museum took part in fairs and exhibitions in Austria and abroad.

The museum also published books, pamphlets, and journals to reach its audience in time and space. *Gesellschaft und Wirtschaft*, Society and Economy, was a collection of 100 Isotype charts. *Fernunterricht*, distant teaching/learning, later renamed *Bildstatistik*, Pictorial Statistics, prefigured modern distance learning.

Three persons led the development of Isotype: Otto Neurath, a sociologist, economist, and philosopher; Marie Reidemeister, a German academic; Georg Arntz, a German graphic designer.

Austrian Otto Neurath's past career included his initiative to, and directorship of, the German museum of war economy in Leipzig during WWI. After the war, his presidency of the Central Office of Economics, in the Bavarian Soviet Republic, was followed by a conviction of assisting high treason and an eighteen-month, later suspended, prison sentence. In 1920, Neurath was back in Vienna to become the director of the Forschungsinstitut für Gemeindewirtschaft, Research Institute for Co-operative Economy. In this capacity Neurath initiated a Museum for Settlement and Town Planning, which within a year – also on Neuraths initiative – was replaced by The Gesellschafts-und Wirtschaftsmuseum in Wien.

Marie Reidemeister (after 1940 Marie Neurath) met Otto Neurath before the start of the Gesellschafts- und Wirtschafts Museum in Wien, became his right hand, and continued working with Isotype after Otto Neurath's death in 1945. Most importantly, Marie Reidemeister played and developed the role of 'transformer'. Otto Neurath and Marie Reidemeister considered the 'transformation' of a message into a principle for a graphic chart the crucial part of the work with Isotype. Transformation was the link between science and graphic design. According to Marie Reidemeister: "We think out which is the point that has to be brought home, and then we try to do so in such a way that everybody will grasp it. We avoid distracting the attention from the more important issues." (p. 15). Also according to Marie Reidemeister, other designers impressed by Isotype would emulate the form but hardly master the transformation (p. 14). Today, the term 'transforming' is not used, but the substance is a natural part of the work of information designers engaged in data visualisation in news media and elsewhere.

Georg Arntz was a German artist working with woodcuts in precise graphic shapes, which caught the attention of Otto Neurath. Georg Arntz began working for the museum in 1928 and in 1929 moved to Vienna where he developed the schematic graphic form that became a signature quality of Isotype. In the process he also changed the technical production from paper cuts to printing from linocuts.

Three conditions for launching Isotype were present. First, a strong-minded initiator with a firm social and educational commitment who was well connected politically; second, highly qualified principal collaborators; and third, a friendly political market.

Partly inspired by the political winds and the following possible need for a foothold outside Austria, Otto Neurath established in 1932 the affiliate Mundaneum to take care of international relations. In 1932 and 1933, Mundaneum established branches in Amsterdam and London respectively. In 1934, the International Foundation for the Promotion of the Vienna Method of Visual Education was established in The Hague. Later in 1934, when the political situation in Austria and Vienna as envisioned by Neurath became dangerous, he, his wife Marie Reidemeister, Georg Arntz, and two other collaborators moved to the Netherlands. The Gesellshafts- und Wirtschaftsmuseum was closed. Part of its material was already transferred to Mundaneum. The rest was seized by the new regime, not the first time a design initiative has been subject to political change. In 1940 the Neuraths moved on to England.

THE NETHERLANDS

Two chapters of the book deal with the continuous work in the Netherlands and England. In Vienna, Isotype had been a means to inform the visitors of the Gesellschafts- und Wirtschaftsmuseum. In the Netherlands the Neurath team had to earn their way from projects. Otto Neurath wrote two books in Basic English: *International Picture Language* and *Basic by Isotype*. Other jobs included production of a children's theatre puppet show and an art exhibition, *Rondon Rembrant*. Also, commissions resulted from Otto Neurath's frequent travels to the USA.

ENGLAND

When Germany occupied the Netherlands the Neuraths moved on to England, where Otto Neurath had been promised a teaching position at Oxford. The Isotype Institute was then established in 1942. The Isotype work in England followed two lines. The Neuraths wrote and designed a number of books for Adprint, a book packager who also published, and they worked on informative films together with British film producer John Rotha. The books dealt with the war effort and social policy. Apart from a couple of booklets this work included a three book series: *America and Britain*, *The Soviets and ourselves*, and *New Democracy*. Two chapters in *Isotype: Design and contexts 1925-1971* deal with film work and children's books respectively.

FILM

Documentary filmmaker John Rotha approached the Neuraths soon after their arrival in England to initiate a collaboration concerning films for the Ministry of Information. The first film, *A few ounces a day* about saving waste to be used in the war effort, was based exclusively on animated Isotype graphics. The Neuraths acted as de facto directors and Maria Neurath made a complete storyboard as well as the graphics to be animated. Later followed several films, where Isotype animations were combined with live action. A series of films that had significant results included, *Worker and* *warfront*, which was shown for workers in factory canteens, *World of Plenty* and *Land of promise* which dealt with food and with planning respectively.

In 1945 Rotha established a special company, Unifilm, with himself and Otto Neurath as directors. After Otto Neurath's death Marie Neurath would continue the cooperation with John Rotha until 1948, when Unifilm closed down. In 1954 Marie Neurath contributed to a TV series, *The World is ours*, and in 1965 to a film *The physics and chemistry of water*.

The film work was not without problems. Some critics found that serious matters should not be treated through the genre of animation. The Neuraths complained when they did not have full control of the work, and Paul Rotha did not always find the necessary support for Isotype work from the Ministry. Professional designers recognise these kinds of problems. Otto Neurath also had some didactic reservations. Isotype on paper lets the viewer see and compare several displays concurrently in space, while film – working in time – doesn't provide that possibility. Also paper media, in contrast to film, gives viewers as much time as they want. Today video technology has solved this problem.

CHILDREN'S BOOKS

In her chapter about children's books Sue Walker rightly states that "The children's books produced by the Isotype Institute provide an excellent insight into Marie Neurath's work as a transformer and show that she had a particular skill in making charts and illustrations that were accessible to children of all ages." (p. 391). This chapter reaches beyond children's books: The account of Marie Neurath's approach is relevant to all designers concerned with data visualisation.

The children's book production took place from the 1940s into the 1970s. Otto Neurath took the initiative, but after his death Marie Neurath edited, wrote, and designed a large number of children's educational books, some of which were schoolbooks. Children's book series included *If you could see inside*, *I'll show you how it happens*, *Visual history of mankind*, *Visual science*, *Wonders of the Modern world*, and *Wonder world of nature*. The Isotype institute delivered both the text and design for these books. Illustrations would include pictograms and all kinds of explanatory drawings. In another series, *They lived like this*, the majority of the illustrations were drawings of contemporary artefacts. This series was co-written by external artists.

Marie Neurath's thoughts about the work with children's books are interesting to everyone working with data visualisation:

> I had to ask myself: what are the essential things we want to show, how can we use comparison, direct the attention, through the arrangement and use of colour, to bring out the most important things at the first glance, and additional features on closer scrutiny. Details had to be meaningful, everything in the picture had to be useful for information. (p.395)

From a note addressed to the readers of the second book in the Visual history series:

> Everything which would not help you understand the meaning, or which would confuse you, is left out. Colours are used only to help make the meaning clearer, never simply as decoration. This means that every line and every colour in these pictures has something to tell you. (p. 403)

U S A

Three factors obstructed Isotype's introduction into the USA. The timing was not good. It was the middle of the depression, there were several imitators (just competitors?) around, and there was Rudolf Modley, a former part-time employee in the administration of the Gesellschafts- und Wirtschaftsmuseum in Wien. Rudolf Modley would cooperate and compete with the Isotype team in Den Hague and Oxford, and challenge Otto Neurath's views.

A group of influential supporters worked together to get Otto Neurath and Isotype to the USA. When in 1930 there was an opportunity to use Isotype at the Museum of Science and Industry in Chicago, Otto Neurath recommended the employment of Rudolf Modley. Here and later Modley acted more independently than envisaged by Otto Neurath.

In 1934 the supporters founded the Organizing Committee for the Institute for Visual Education "to establish in the United States an organization which can develop and promote the graphic methods of presenting social and economic facts, which have been characterised by the Vienna Method as exemplified in the work of the Gesellschafts- und Wirtschaftsmuseum in Wien under the direction of Dr. Otto Neurath" (p. 307). When the organisation did not follow Modley's advice, he created his own company, Pictorial Statistics Inc. Otto Neurath and Rudolf Modley held differing views. Neurath wanted standardised pictograms designed centrally while Modely had a more relaxed view. Neurath explained:

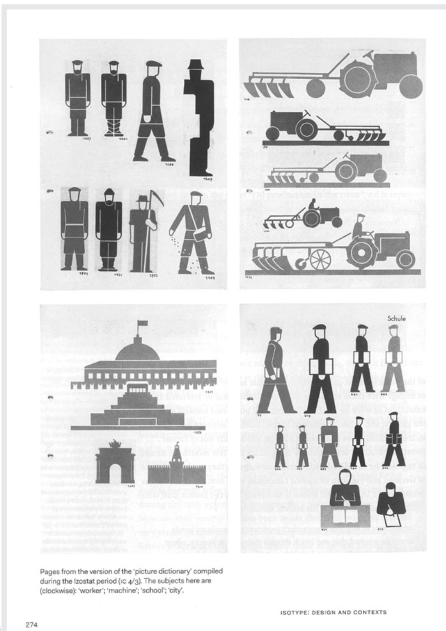
> That is to say, for our picture language one general list of a limited number of signs is needed for international use, and this has to be worked out by or under the control of one chief organization (This organization is now the ISOTYPE work-room at the Hague). (p. 332).

Also, Modley saw the pictograms as elements that could have their own life while Neurath saw pictograms as parts of visual arguments enabled by transformation. Modley was not interested in transformation. In line with this view he published symbols sheets with pictograms to be used by everyone and a book entitled, *1000 Pictorial Symbols* (1942).

Otto Neurath travelled to the USA several times and secured important commissions, primarily in the health sector. Isotype also delivered illustrations to *Compton's pictored encyclopedia* (1939) and Otto Neurath wrote *Modern man in the making* (1939) for Knopff publishers. After Neurath's death Marie Neurath wrote an essay on Isotype for Henry Dryfuss's *Symbol Source Book* (1972).

RUSSIA

The Isotype team's efforts in Russia took place from 1931-1934. Russia did not want to commission Isotype work from Vienna. Instead, they wanted Isotype staff to help establish a Soviet institute. A special organisation named Izostat was established with Otto Neurath as one of two directors, and several Isotype staff would join them for shorter or longer periods. The total staff at times would be as high as 75. A number of problems hindered cooperation. The work primarily dealt with visualising the established success of the first five-year plan 1928-1932 and the predicted success of second five-year plan 1933-1937. While The Vienna Method as practiced in Vienna was based on empirical facts, the Russians wanted forecasts to play an essential role. The estimates were often exaggerated. Another problem was that the Russians wanted naturalistic pictograms aligned with wanted Soviet realism. Also, the Russians wanted more, sometimes propagandistic, illustration. The cooperation resulted in some publications with more or less Viennese influence. Georg Arntz made a series of charts for Izvestia, charts for exhibitions, and a number of publications more or less influenced by the Isotype team. One Isotype idea was used with a new meaning. In the Vienna Method black was sometimes used to illustrate worse, while red would stand for better. In Russian charts red would stand for Russia while black would stand for other nations.



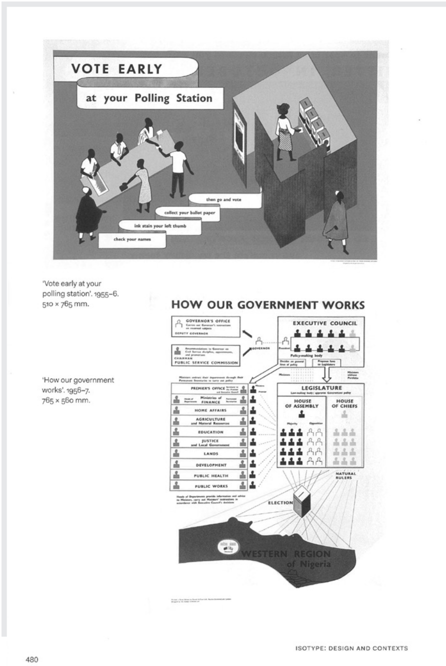
Page 174
Isotype work for Izostat in Moscow characterised by Russian realism.
(original in color)

In 1934 the Russians wrote to Otto Neurath that the contract did not comply with Russian law and the amount due at the end of the contract would not be paid. The latter incident was a major blow to the Isotype organisation, which in Den Hague depended on paid work. Izostat continued without Isotype help until 1940.

AFRICA

Some Isotype work in Africa took place from 1952-1958. Otto Neurath reportedly said that Isotype was not for the Viennese, but for the Africans (p. 449). In 1943 he worked on a proposal for an exhibition for the British Colonial Office entitled, *Human life in Africa*. This project did not materialize.

In 1953 a partnership between Buffalo Books, a subsidiary of Adprint, the Isotype Institute, and Purnell and Sons, a printing firm, planned and published the magazine *Forward* addressing the three British colonies Gold Coast (Ghana), Sierra Leone, and the Western Region of Nigeria soon to become independent. A trial issue and an issue number 1, dealing with culture, adventure, sports, and practical advice were published before the magazine was determined to be economically impossible in 1953.



Page 480
Isotype work with limited abstraction for Western Region, Nigeria, promoted participation in an upcoming election. *(original in color)*

In 1954 Marie Neurath wrote a memorandum sketching what Isotype could do in the Western Region of Nigeria. It included a visual explanation of the aims of the government and the establishment of a local workshop for Isotype run by trained Nigerians. Marie Neurath visited Africa three times and a series of booklets and poster-leaflets were produced, while the workshop remained a plan.

For the Western Region a series of 16-page White Paper Booklets were published including *Education for all in the Western region*, *Better farming for better living in the Western Region*, *Health for all in the Western Region*, and *Paying for Progress in the Western Region*. Also a series of poster-leaflets dealing with health issues was produced. As indicated by the name, the poster-leaflets worked both as wall charts and as folded leaflets to take away. A following visit to the Western Region resulted in commissions for a new series of booklets.

Compared with the Western Region in Nigeria the Gold Coast and Sierra Leone had little work for Isotype. In 1956 Isotype made one leaflet for The Gold Coast, *The Volta River Project: what it means to you*, and in 1957 one for Sierra Leone *Voting, general election*.

In the late fifties the Isotype work in Africa came to an end. The chapter author, Eric Kindel, offers a number of probable explanations, one being the distance, another being the failure to establish a local workshop with trained local people. To the Isotype Institute the African experience meant realising the need for locally adapted symbols for man, woman, house, and more. The pictograms should 'speak'. In the same vein, charts were given familiar backgrounds. According to Marie Neurath: "[A]dherence to the method cannot go as far as imposing an alien background on those unable to share one's experience of it." (p. 495). This did not, according to Marie Neurath "imply that the system had to be radically changed" (p.495).

DESIGN

In the second central chapter of the book, Robin Kinross presents the design of Isotype. This chapter was originally a part of Kinross's MPhil thesis from 1979. (Robin Kinross is the owner of Hyphen Press, the publisher of this book as well as Otto Neurath's 'visual biography': *From hieroglyphics to Isotype* (2010) and a string of well presented books on typography.) The provenance of the chapter may explain the order in which it deals with the subject. The description begins with the formats of the wall charts, and moves through colour, male and female, qualifying symbols, and guide pictures, down to pictograms and configuration. In the latter part Kinross codifies six types of displays dealt with by Isotype. This would have been a natural start of the chapter to be followed by pictograms. Apart from this peculiar arrangement the chapter gives a robust description of the elements used in Isotype. Kinross calls Isotype "a coherent approach to ordering material in graphic form" (p.107). It covers what we today call 'data visualisation'. In contrast to the remaining part of the book, Kinross offers a few critical remarks on Isotype.

Considering Neurath's interest in education it is remarkable that there exists no manual, no single, document explaining the Isotype design principles thoroughly. One reason could be that Isotype remained a work in progress. Another reason could be that Neurath did not want everyone to design visual displays, but rather to commission the displays from the initiators. So-called notes, single sheets each describing a design subject, were descriptive rather than prescriptive. They described current practice more than recommending what should be done. Also, Neurath's publication, *International Picture Language*, 1936, written in Basic English doesn't serve as a manual either.

Kinross's description of Isotype design gives a clear impression of Isotype being a work in progress. Pictograms, qualification, grouping of pictograms, use of colour, use of typography, and configuration would change considerably between 1925 and 1934, especially after Georg Arntz joined the team. However, this development did not always follow a straight line. Different principles were sometimes used concurrently; old design features were sometimes used after new design features were introduced. The development involved standardization, modularization, and simplification. Pictures would be reused and be combined; the use of colours would be restricted.

Kinross refers to the common misunderstanding that "quantified rows and columns" "might be typical of the work as a whole" (p. 142). Well, these picture tables and their pictograms are what most of us think about first when we think about Isotype. The picture tables and their pictograms are featured on the covers of publications and wherever Isotype is discussed. Kinross shows the width of Isotype by the following classification (p. 139).

Charts showing quantified material:

1. rows and columns [picture tables],
2. division of a whole (usually a checker-board),
3. geographically ordered pictograms and more diagrammatic charts,
4. quantities related to area (usually showing densities),
5. flows.

Charts showing non-quantified material:

Neurath broke down the picture table category into six sub-categories (p. 140):

1.1	comparison of total quantities,
1.2	where sizes of constituent parts are interesting, as well as total quantities,
1.3	where relative sizes of constituent parts are most important,
1.4	to make a shift particularly clear; alignment left and right to form an axis,
1.5	where the sizes of parts and of wholes are equally important; one compares both horizontally and vertically,
1.6	to allow comparison of parts and wholes, and to make a shift clear; especially important in showing changes over time.

A schematic drawing and an Isotype picture table illustrate each of these sub-categories.

While the fact that a large part of the text of this chapter is devoted to picture tables and pictograms may support the idea that Isotype first and foremost is picture tables, the book's numerous illustrations establish some balance. Isotype is both picture tables and a general approach to data visualisation.

In a chapter about pictograms, Christopher Burke confirms that a direct line from the Isotype pictograms to the pictograms used in transport and communication today hardly exists. However, qualities such as standardization, modularity, and schematization are parts of the Isotype heritage. Isotype pictograms worked in lines in picture tables to compare something, while modern pictograms in transport and communication simply point to the existence or condition of something. Otto Neurath, however, suggested that the Isotype pictograms could possibly also be used for public information signs. This application was not realised.

WHY NOT?

Isotype: Design and contexts 1925–1971 is a comprehensive introduction to the Isotype idea. The book's 12 chapters written by nine authors are well planned with a minimum of overlaps. While the main text goes into considerable historical detail, the illustrations present the elements and charts by which Isotype should be known and appreciated. The numerous illustrations – more than 400 – and their elaborate captions turn the book into a portable archive, which for everybody unable to access the Isotype collection at University of Reading will remain the most important Isotype resource.

Implicitly the book relays a well-known phenomenon: how a design idea born to solve one problem if successful becomes a solution that looks for other problems. From informing the Viennese citizens the problem changed into finding potential outlets for the newfound method.

In the beginning of *Isotype: Design and contexts 1925–1971* Christopher Burke states, "The best way to bring these [the qualities of Isotype] to the fore is to examine it as a historical phenomenon in all the complexities of its contexts." (p. 14). This is questionable. To compare is the basic function of Isotype. Isotype should itself be compared with competing data visualizing formats. How can we evaluate the virtues of airships without comparing airships with other airborne vessels?

We still need a balanced discussion of the qualities of Isotype. What exactly is the Isotype approach? How does it survive today? How does Isotype compare with the display formats that news media and others prefer today? What are the pros and the cons of Isotype compared with other more frequently used data visualising formats such as bar charts, bubble charts, and line charts? Understandability, accuracy, attraction, and memorability are factors that should be discussed. The discussion should also include the intended target groups of Isotype and the contemporary audiences of news media and professional literature. Is Isotype only for uneducated people?

One probable finding is that most contemporary audiences prefer exact information to the visual explanation offered by Isotype picture tables. Today bar charts, pie charts, and bubble charts, which in principle present *visual* messages, are as a rule supplied with exact figures. Such displays are hybrids. They are half visual display, half table. The visual part lets the reader get a fast idea, while the figures deliver accuracy. In a short period Isotype's picture tables were also supplied with exact figures. In later displays the figures were abandoned. The big, undisputable advantage of isotype displays is that they are attention grabbing and attractive to look at. The visual attraction may be accompanied by good memorability.

P S

Design and contexts 1925-1971 is well crafted with a pleasing design. However, there are two minor flaws. First, the key to source abbreviations is located on page 18, while readers expect to find it in the beginning of the book where it would be easy to retrieve for later reference. Second, tiny, alphanumeric caption designations of up to four characters are written with lowercase numerals (6 and 8 with ascenders and 3, 4, 5, 7, and 9 with descenders), which are difficult to read, especially when presented in the very small type used for marginalia.

Per Mollerup
14 August 2014

THE MOST IMPORTANT DESIGN BOOKS MOST DESIGNERS HAVE NEVER READ:*

The Case for Mental Imagery

by

Stephen M. Kosslyn, William L. Thompson, and Giorgio Ganis

Mike Zender

date: 2035 CE
location: a design office

The designer looked at the screen and watched the child's memory of being sick. "I'm sorry to ask you this, but think of diarrhea again please," she said. The toilet shimmered into view briefly followed by transparent wavy lines. The designer noted the lines, then replayed the other children's memory and noted that 67% of them included shimmering, wavy lines to represent smell. "Thank you, that's all I needed. You've really helped me design this icon," she said.

Paul Rand once said that communication design is about "saying the commonplace in an uncommonplace way." (Rand, 1970, p. 36) This suggests that effective communication is essentially enhancing the familiar. For visual communication design, this means creating unique images that will connect in predictable ways with the images people already hold in their minds. From this perspective, the whole user-centered design movement is a cultivation of means for designers knowing, not just assuming, the mental images people have. Stephen Kosslyn, William L. Thompson, and Giorgio Ganis' book *The Case for Mental Imagery* (Kosslyn, Thompson, & Ganis, 2006, p. 4) gives designers an accurate glimpse into how mental images work.

MENTAL IMAGES

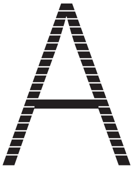
The plausibility of the fictional design office above hinges on the answer to a debate that has raged for at least decades, perhaps centuries: do we see mental images or not? According to Kosslyn, Thompson , and Ganis, "A mental image occurs when a representation of the type created during the initial stages of perception is present but the stimulus is not actually being perceived." Mental imagery is seeing what is not there, not an illusion or a mirage, but seeing in our mind something familiar and then perhaps using that mental image to think with or solve a problem. We might experience this by answering this question: how many windows face the street in your house or apartment? Given this task most people gaze blankly for a second or two as they push into memory an image of their house and then briefly count the windows in the image. Kosslyn et. al. cite similar questions such as "Do you know which is darker green, a frozen pea or a pine tree? Or the hand in which the Statue of Liberty holds the torch" as examples where people use mental imagery.

Belief that this phenomena exists are not new. The authors briefly note that thinkers from the classic Greeks to Einstein claimed to use mental images "in memory, problem solving, creativity, emotion, and language comprehension." However, introspective experiences are notoriously difficult to study, easy to refute, and thus ripe territory for endless debate. Kosslyn, Thompson, and Ganis use Chapters 2 and 3 to detail the debate, Chapter 4 to marshal empirical findings from a broad range of cognitive psychologists and neuroscientists in order to settle the debate, and Chapter 5 to articulate a well-founded theory of mental imagery.

The theory articulated in the book is based on the process of visual perception which it describes. The eye is just the start of a process that occupies much of the brain. In fact, approximately 50% of the cerebral cortex is devoted to visual processing. The brain is not like a general-purpose computer with generic processing capacity to which are downloaded different problems for analysis. Rather, the brain is like a special purpose device with different neurons in different regions hard-wired to accomplish specific tasks. Vision is one of the brain's largest tasks. In visual perception a huge volume of sensations are processed and reduced to simpler more organized forms. It's as if individual camera pixels are processed for simple features then structured into units that correspond to distinct objects and key properties that define and distinguish those objects from each other. Kosslyn and his colleagues propose that we can reverse this process and push the abstracted memory of a visual object backward onto the brain's early visual processing areas and there mentally re-construct a representation of something. Representation is a key idea here. The authors point out that stored depictive representations are not like photographic pictures. They are simplified forms that can be re-presented and then examined and to which detail can be added. It may help communication designers to think of these abstracted representations as "brain icons" because they, like drawn icons, are simplified and focus on key features of an object or idea rather than inessential details. We can use these mental images to reason about problems, like whether a jar could squeeze into a crowded pantry shelf, or to communicate with people by creating images that connect with their visual imagery. It is important to note that understood this way, there is a deep and complex relationship between seeing and thinking that deserves attention.

COMPETING THEORIES

Kosslyn et. al. identify the core of the debate as two competing conceptions of the format we use to store internal visual representations: depictive



Depictive format above

FIGURE 1

FIGURE 2

Propositional format below

"two diagonal lines that meet at the top, joined halfway down by a short horizontal segment."

(Kosslyn et al., 2006, p. 12)

or propositional. The depictive approach suggests that our brain encodes images in a visual format using points, similar to the way a computer screen uses pixels. The blocks forming the triangle in Figure 1 illustrate a depictive format. The propositional approach suggests that we format images using abstract concepts like words in language or computer software. The words "two diagonal lines that meet at the top, joined halfway down by a short horizontal segment." in Figure 2 illustrate a propositional format for the same image.

The format used may seem like an academic debate, but it matters because the format of representation makes possible, or at least preferences, different kinds of thinking and from this the creation of different knowledge.

To settle the debate the authors call upon findings that add significant detail to the outline of the perceptual process noted above. Very early in this process the image from our retina is topographically mapped point-for-point on our brain. Objects close to each other on our retina are also close to each other on the cortical area called V1. There are, in fact, several topographically organized layers in V1 with each layer providing different kinds of processing. Cutting down through layers are columns that distinguish different line orientation, curve, value, and hue. These topographic layers are part of what Kosslyn labels the "visual buffer." The "visual buffer" then "reports" the results to other areas of the brain where patterns and shapes are assembled, where objects are formed and subsequently identified. Kosslyn asserts that through these successive stages a "population code" is assigned containing in abstracted form the key visual features that define an object. Kosslyn posits a "hybrid representation" that combines information for each point about its role in the depiction of the object, as well as additional abstracted information. "In spite of their coding nondepictive information, these hybrid representations cannot be reduced to propositional representations. Crucially, they use space (literally, on the cortex) to represent space in the world. The fact that each point codes additional information does not obviate its role in depicting the shape."

The highlight of Kosslyn's argument is that these encoded representations can be recalled and when they are, an image is reconstructed from memory using the same topographic neural space in the "visual buffer" that was used to "see" the initial image from the eye. In Kosslyn's words, the "stored shape representation is primed so strongly that activation is propagated backwards, including a representation of a part or characteristic in the visual buffer (which corresponds to the depictive image itself)." Kosslyn theorizes that we literally re-construct the object from memory and create a representation of it. These are mental images.

VALIDATION

When Kosslyn and his colleagues wrote this book some years ago the viability of their theory was still open to debate. But much has happen since then to support its basic premises.

In 1988 Tootell et. al. provided a foundation for how we see when they showed a topographically represented visual image mapped on the surface of a primate brain. Over the years various scientists have developed techniques that enable them to dye a primate brain and see there on the cortex - in real time - the images from the eye. A 2012 NIH presentation by Dr. Eyal Seidemann is one example of video showing this. ("Decision Related Activity and Top-down Attentional Modulations in Primate VI" http://videocast.nih.gov/summary.asp?Live=11769&bhcp=1) More recently researchers used fMRI to produce an image of a person's recalled memory (mental image) of a simple object. In 2014 Dr's. Cowen, Chun and Kuhl presented findings that through observing brain activity they were able to reconstruct recognizable individual faces from people's mental images of faces they were seeing ("Neural portraits of perception: Reconstructing face images from evoked brain activity" in *Neuroimage).* The title of the March 28, 2014 Fox news article reporting this paper was *"We know what you're thinking: Scientists find a way to read minds."* by Maxim Lott. While the face reconstruction study may be as much about the inventiveness of the computer processes employed as it is about the biological ones, its findings dramatically support the foundation of Kosslyn's hypothesis: mental images are seen reconstructed in the visual buffer. These studies can "see" them.

SO WHAT?

The Case for Mental Imagery may sound interesting to some readers, but to others the question "So what?" may have been lingering for some time now.

So what? How is the information in this book relevant to the designer? One answer is "A theoretical foundation for communication design." Communication design has entertained competing theories to guide practice. Some, such as semiotics, are based in linguistics. Theories of visual perception such as the one articulated by Kosslyn, Thompson, and Ganis may help provide more appropriate visual ground for a theory of visual communication. Knowing how people process, store, and use images is at the heart of visual communication. It's true that communication designers create objects that use both textual and visual forms to communicate and much has been written recently about the role of designer as author and the need for more writing in design education. Without dismissing the positive role designers can play in crafting the written content of the communication they create (designer as author), or diminishing the role writing can play for organizing and expressing thought in design classes (writing in design education), Kosslyn's theory suggests that there is a good reason that college "communication" programs focus on writing while "visual communication" programs focus on image making. Visual images are the essence of visual communication. Communication designers employ forms of communication that largely bypass language. Kosslyn reminds us that people think with images. One benefit of

Kosslyn's theory as it applies to design is that it is founded on hard-wired neurobiological perceptual processes common to all people, suggesting that design principles founded on this approach may be universal, applying to people of every age, language, literacy level, and culture. With limited research resources to invest, design might do well to focus on universal visual processes that can apply to nearly everyone before building upon theories focused on individual differences.

In addition to providing a theoretical foundation for visual communication, "Visual thinking" is another defense of the book's relevance. Kosslyn's theory, based as it is on depictive representations, means that visual designers use a visual language that is inherently more flexible and less inhibited by arbitrary encoding structures than language. Depictive reasoning can be more ambiguous and less structured than propositional reasoning. Images are more direct, less categorical, less overtly defined, and thus better suited for creative thinking and problem solving than language-based propositional representations that seem plodding by comparison. "I see" is a common visual-based metaphor for sudden understanding and an apt metaphor for visually-empowered design thinking.

"User centered" is yet another response. Kosslyn's theory means that communication designers now and in the future can reliably identify the mental images that people have, thus gaining direct insight into how to communicate with them more accurately. Designers who know their subjects' mental images can more reliably produce images that evoke the correct meaning. Knowing people's mental images moves user-centeredness into the user's head, literally. The point of view for designers might be transformative.

"Evidence-based" is another reply. Kosslyn's theory doesn't just apply to the front end of design creation, but also to the back end of design evaluation. Using people's mental images to evaluate communication objects could give not only very accurate measurement of communication but insight to corrective action. A loop of creation and evaluation based on reliable measurement of mental image may provide communication design with some solid principles for practice.

Another Paul Rand quote suggests a final answer when he states, "…the designer must steer clear of visual clichés by some unexpected interpretation of the commonplace." "Innovation" is the final answer. Design has been said to be the process of converting existing states to preferred ones. Designers don't just create what already exists; they create something new. So how can knowing the images people already have in their heads help create something new? To a designer the question is the answer. Knowing what people think enables us to take liberties, to explore novel variations and "unexpected interpretations", to both connect and expand what is in people's minds. Apart from knowing the people's mental images designers innovate in the dark, ignorant of whether their novel approaches support or hinder communication.

Kosslyn, Thompson, and Ganis' theory means several things to communication design. It means that seeing and thinking are complementary, helping to explain the effectiveness of information visualization. It means that visual communication designers who create images are directly connecting to the one of the most significant means people have for processing information, for thinking. It means that emphasis on visual thinking is one key to why design is good at creative problem solving.

If design theories should be founded on research findings in visual perception and cognition, then they will in some measure be founded on work by Kosslyn and his colleagues. It's a book that most designers should read.

Cowen, Chun, & Kuhl. (2014). "Neural portraits of perception: Reconstructing face images from evoked brain activity." Neuroimage

Kosslyn, Stephen M., Thompson, William L., & Ganis, Giorgio. (2006). *The Case for Mental Imagery*. New York, NY: Oxford.

Rand, Paul. (1970). *Thoughts on Design* (Third ed.). New York: Wittenborn Schultz.

Advisory Board

WELCOME

Visible Language wishes to welcome new advisory Board member Keith Crutcher. Keith reflects our interest in connecting to disciplines whose research is well-advanced and whose knowledge is related to visual communication.

Keith Crutcher

Keith A. Crutcher, Ph.D., has over 30 years of experience in biomedical research and technology including prior tenured faculty appointments at the University of Utah (7 years) and the University of Cincinnati (22 years), a founding role in an early stage drug discovery company (ApoLogic, Inc.), and four years serving as a Scientific Review Officer at the Center for Scientific Review at the National Institutes for Health. In the latter role, he managed the peer review of hundreds of grant applications. His academic research program, funded by the NIH, NSF, and other agencies, included studies of brain injury and Alzheimer's disease resulting in over 100 peer-reviewed publications, two issued patents, and numerous presentations. He has also served as an ad hoc reviewer for several federal agencies and private foundations as well as serving on the editorial boards of several journals including *Experimental Neurology*, *Aging Cell*, and *Neurobiology of Aging*, where he participated in the peer review of numerous manuscript submissions. He currently does consulting work for applicants seeking research funding and provides assistance in preparing proposals and navigating the peer review system at various federal agencies.

JOURNAL INFORMATION

Visible Language is an academic journal focused on research in visual communication. We invite articles from all disciplines that concern visual communication and would be of interest to designers.

READERSHIP

Visible Language, an academic journal, seeks to advance research and scholarship for two types of readers: academics and professionals. The academic is motivated to consume knowledge in order to advance knowledge thorough research and teaching. The professional is motivated to consume and apply knowledge to improve practice. *Visible Language* seeks to be highly academic without being inaccessible. To the extent possible given your topic, *Visible Language* seeks articles written to be accessible to both our reader types. Anyone interested may request a copy of our editorial guidelines for authors.

EDITORIAL CORRESPONDENCE

Article concepts, manuscripts, inquiries about research and other contributions to the journal should be addressed to the editor. We encourage article concepts written as an extended abstract of 1 to 2 pages single-spaced. We will offer prompt feedback on article concepts with our initial opinion on their suitability for the journal. Manuscripts accepted for peer review will receive a summary response of questions or comments within three weeks. Letters to the editor are welcome. Your response — and the author's reply — will not be published without your permission and your approval of any editing. If you are interested in submitting an article to the journal and would like a copy of our Notes on the Preparation of a Manuscript, please obtain it from the journal's website at http://visiblelanguagejournal.com

Editorial correspondence should be addressed to:

Mike Zender
Editor, *Visible Language*
College of Design, Architecture, Art, and Planning
School of Design
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
email: mike.zender@uc.edu

If you are interested in serving as guest editor for a special issue devoted to your specific research interest, write to the editor, outlining the general ideas you have in mind and listing a half dozen or so topics and possible authors. If you would rather discuss the idea first, call the editor at: phone number

BUSINESS CORRESPONDENCE

Subscriptions, advertising and related matters should be addressed to:

Visible Language
Sheri Cottingim
Office of Business Affairs
College of Design, Architecture, Art, and Planning
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
telephone 513 556-4377
email: sheri.cottingim@uc.edu

48 — 3

Before there was reading there was seeing. *Visible Language* has been concerned with ideas that help define the unique role and properties of visual communication. A basic premise of the journal has been that created visual form is an autonomous system of expression that must be defined and explored on its own terms. Today more than ever people navigate the world and probe life's meaning through visual language. This journal is devoted to enhancing people's experience through the advancement of research and practice of visual communication.

If you are involved in creating or understanding visual communication in any field, we invite your participation in *Visible Language*. While our scope is broad, our disciplinary application is primarily design. Because sensory experience is foundational in design, research in design is often research in the experience of visual form: how it is made, why it is beautiful, how it functions to help people form meaning. Research from many disciplines sheds light on this experience: neuroscience, cognition, perception, psychology, education, communication, informatics, computer science, library science, linguistics. We welcome articles from these disciplines and more.

Published continuously since 1967, *Visible Language* maintains its policy of having no formal editorial affiliation with any professional organization — this requires the continuing, active cooperation of key investigators and practitioners in all of the disciplines that impinge on the journal's mission as stated above.

Website
http://visiblelanguagejournal.com

Postmaster:
send address changes to:
circulation manager name
Office of Business Affairs
College of Design, Architecture, Art, and Planning
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016

Published tri-annually in January, May and October

SUBSCRIPTION RATES

United States	Individual	Institutional
1 year	$35.00	$65.00
2 year	$65.00	$124.00
3 year	$90.00	$183.00

Canadian*	Individual	Institutional
1 year	$44.00	$ 74.00
2 year	$83.00	$142.00
3 year	$117.00	$210.00

Foreign**	Individual	Institutional
1 year	$ 56.00	$ 86.00
2 year	$107.00	$166.00
3 year	$153.00	$246.00

Prepayment is required. Make checks payable to University of Cincinnati *Visible Language* in U.S. currency only, foreign banks need a U.S. correspondent bank.

* Canadian subscriptions include additional postage ($9.00 per year).
**Foreign subscriptions include additional postage ($21.00 per year).

ISSN 0022-2224
Published continuously since 1967.
Index included in last issue of volume year.

BACK COPIES

A limited number of nearly all back numbers is available. The journal website at http://visiblelanguagejournal.com is searchable and lists all issues, contents and abstracts.

COPYRIGHT INFORMATION