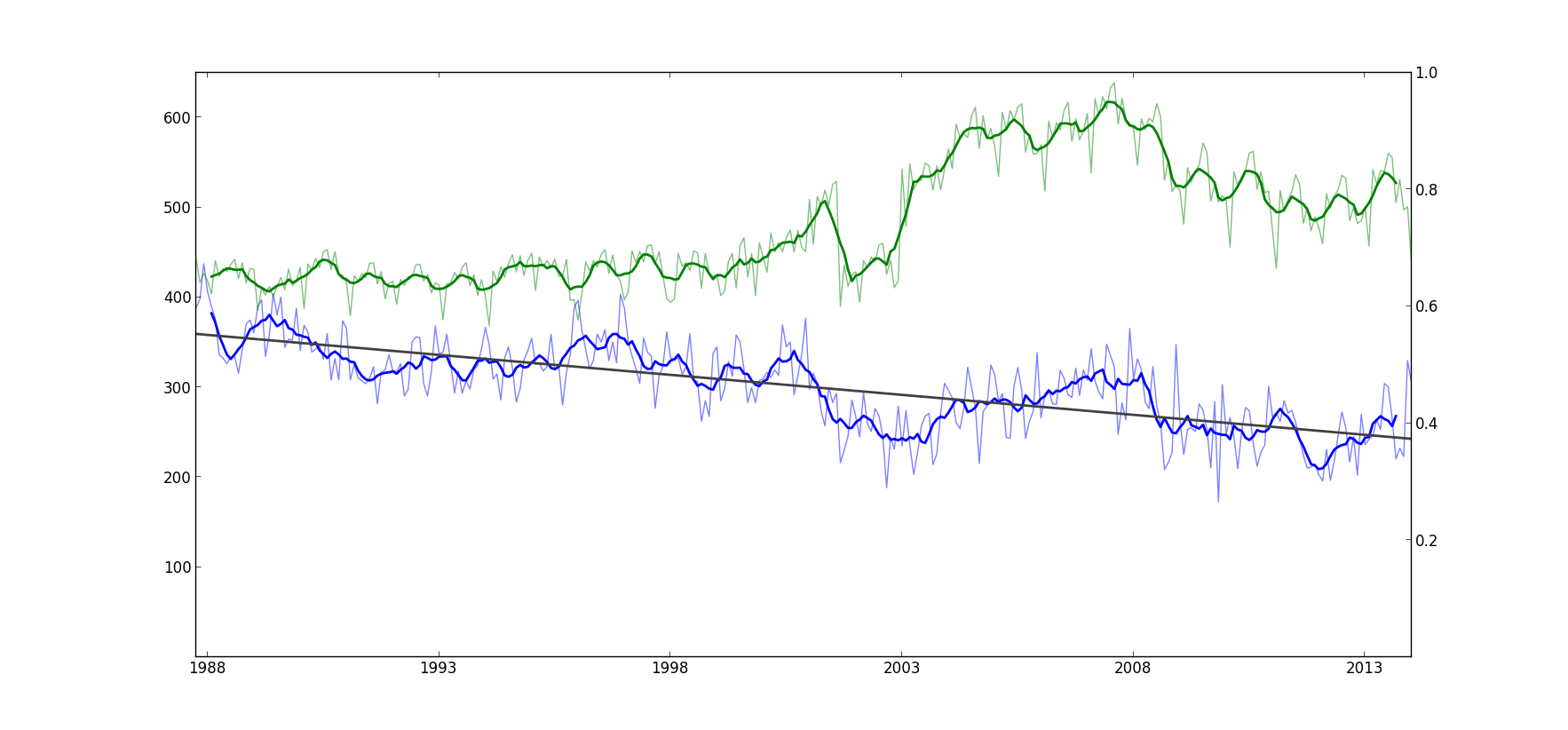
Flight Performance in the United States during Holidays

Jesse Green

INTRODUCTION

This documents looks at passenger airline travel and flight performance statistics provided by the Bureau of Transportation Statistics between October 1987 and January 2014. The focus of the paper will be a quantitative analysis of flight performance near federal holidays in the United States, as an attempt to improve upon the qualitative notion that exists for expecting delays and avoiding flights during holiday travel. To avoid seasonal variations that occur within flight volume and weather, specific holidays will also be evaluated individually.

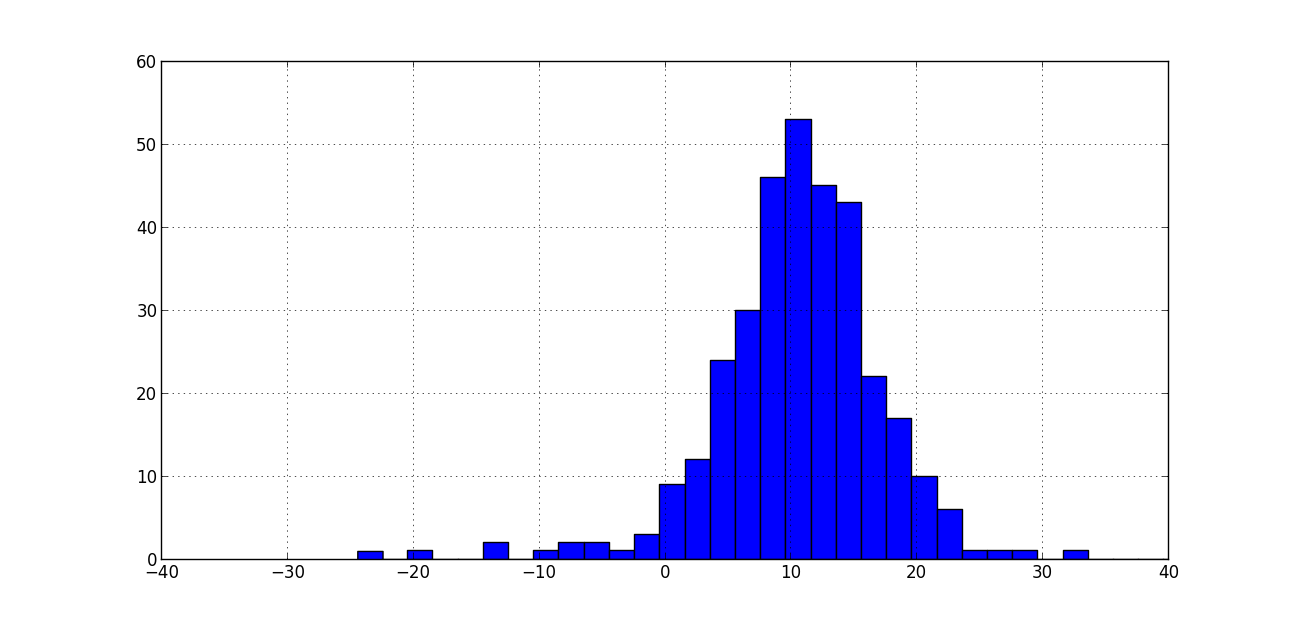
Before stepping into a detailed look at flight delays around holidays, it is worthwhile to examine flight volume and performance as it has changed over the years, and consider a few ways to define delays in a meaningful way. Figure 1 shows the total flight volume and percentage of delayed flights for all available BTS data. The figure shows monthly values overlaid with the 7-month moving average to better illustrate the overall trend. In this plot a flight is considered delayed if its actual arrival time is greater than its scheduled arrival time by any amount. It is easy to see that the overall volume of flights has increased over the time period, while the number of delays has decreased an average of 0.7 percentage points annually. The decrease in delays may be explained by an increase in airports (there are 82 new airports in 2013 versus 1988, but this only accounts for 12.5% of increased outbound flight traffic); or perhaps delays only appear to decrease because there has been an increase in CRS reservation times allocated by airlines for itineraries (see Figure 2). The following sections will look at the methods taken to understand flight delays, different ways of quantifying delays with statistics, and understanding how these statistics are affected during holiday travel.

Figure 1 – Flight Volume and Percent Delayed in the United States

METHODS

Data Collection and Storage

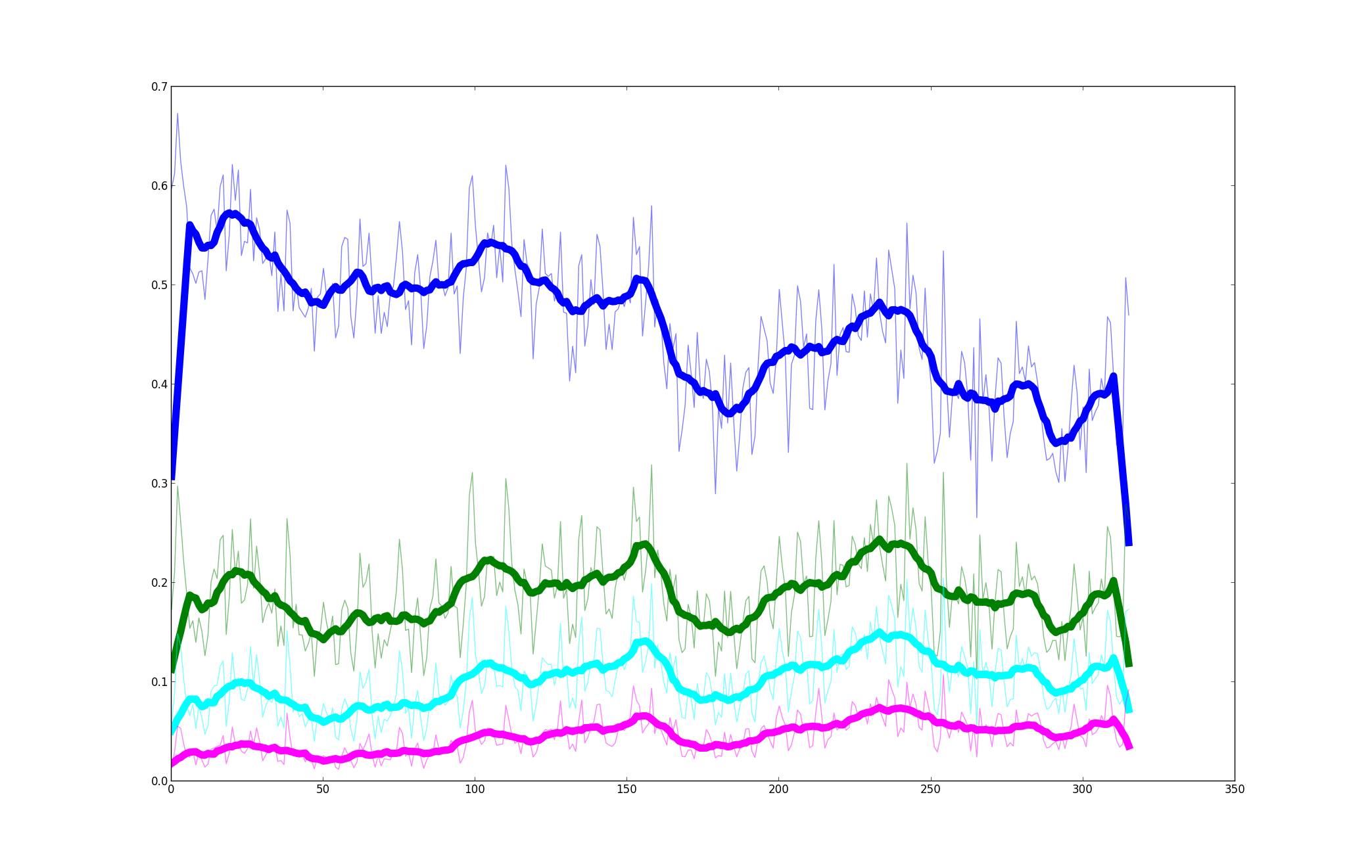
There is over 20GB of flight performance data available for download through the BTS website application. Given the quantity of data and the desire to query across the entire timeframe and various dimensions, it is most convenient to import the dataset into a database where it can be efficiently queried. Using a script the data was downloaded, imported into MySQL, indexed and then used for the analysis.



Delay Statistics

An important step in this analysis is to determine the statistics by which to measure delayed flights. The usefulness of a statistic may vary depending on the audience, and this analysis will assume the audience to be a typical airline passenger interested in traveling for the holidays. Assuming a delay can also be negative for flights arriving early, using mean and median to summarize delay provide little insight. For example, considering all flights on Mondays in 2013 scheduled to depart between 4 p.m. and 5 p.m. are on average delayed 11.5 minutes. The median delay for this subset is -1 minute. These two statistics reveal something commonly seen in the dataset: this subset has a positively skewed probability distribution.

This type of information has little practical value for an everyday airline passenger looking for some insight. What passengers are really looking to know is whether or not there is a good chance their travel plans will be interrupted by a significant amount of time spent in an airport, or even worse on the tarmac. Instead of considering the number of flights that were delayed overall, let us consider only the delays occurring for more than a significant amount of time. Figure 3 shows the percentage of flights delayed more than *n* minutes over the dataset time period. The number of overall delays is decreasing (i.e. decreasing median), but the average delay remained constant. This means the flights that are delayed recently are delayed for longer amounts of time, as seen by the slightly upward trend in the 30+ and 60+ minute delays.



Additional Datasets

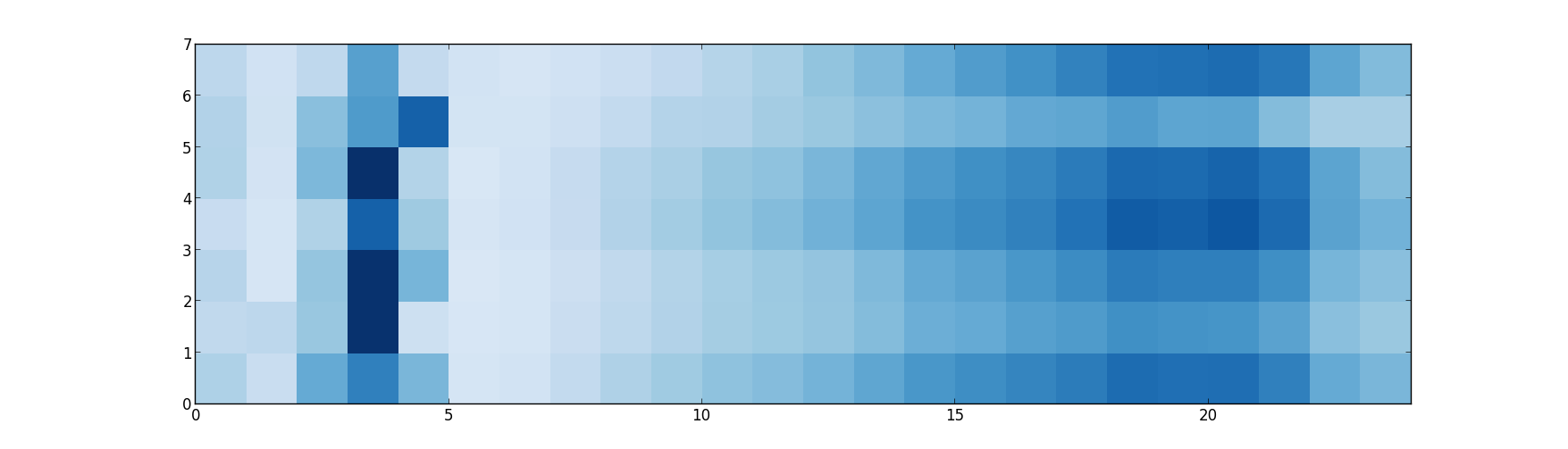
Two additional datasets were used in this analysis. The website timeanddate.com provided the dates for actual and observed holidays in the United States during the period of interest. For the interactive website, GPS data is provided by openflights.org.

results

This section will analyse domestic flights for the last 10 year (2004 – 2013) and will use the percentage of delays longer than 60 minutes as the delay statistic. This allows the reader to interpret the results in a meaningful way and avoids delays as a result of variations in scheduled flight times as seen in Figure 2. However these analyses still represent an aggregate performance independent of airport size, flight volume, weather, etc. For a much more focused investigation, an interactive application will then be presented that allows a user to review performance of specific route and holiday combinations.

A Typical Week of Flight Delays

The likelihood of delays in a given week look as one might expect, with almost no chance of long delays on early flights, and the frequency of delays increasing throughout the day and peaking in the evening. The heat-map in Figure 4 visualizes the frequency of delays for a typical week in the last 10 years, allowing for a quick comparison across day or time. Tuesday and Saturday are the best days to fly for avoiding a long delay. Although flights between 3 a.m. and 4 a.m. are rare, they have extended delays more often than any other time block. Figure 5 shows the same information as Figure 4, but is organized linearly to reveal shape of delays stacking up each day. It also contains the total number of flights each day in the US on average.



Three-Day Weekends and Memorial Day

Your paper’s title should be in Time New Roman 18-point bold and centred. Authors’ names should be in Times New Roman 12-point and centred, and affiliations in Times New Roman 11-point italic and centred. Use capitals for the author’s surname. Use superscripted alphabetic characters for connecting authors with their affiliations. In case of multiple affiliations, commas separate the characters.

Thanksgiving

Every submission should begin with an structured abstract of no more than 150 words. The abstract should be a concise statement of the problem, approach, and conclusions of the work described. It should clearly state the paper's contribution to the field. If possible, use an IMRAD structure.

The abstract is followed by a list of up to 6 keywords separated by semi-columns. The first keyword must be chosen among the areas (Table 2), and the second among the fields. Try to adhere to these lists, but in case your domain is not listed in the fields, you may create a new one.

Memorial Day

Please use a 10-point Times New Roman font or, if it is unavailable, another proportional font with serifs, as close as possible in appearance to Times New Roman 10-point. Please use sans serif or non-proportional fonts only for special purposes, such as headings or source code text.

DISCUSSION

It is important that you write for the general audience. It is also important that your work is presented in a professional fashion, which is what this guideline is intended to help you with. By adhering to the guideline, you also help the conference organizers tremendously in reducing our workload and ensuring impressive presentation of your conference paper. We thank you very much for your cooperation and look forward to receiving your nice looking, camera-ready version!

REFERENCES

Agarwal, R. & Karahanna, E. (2000) Time Flies when You're having Fun: Cognitive Absorption and Beliefs about Information Technology Usage, *MIS Quarterly*, 24, 4, 665-694.

Ajzen, I. (1988) Attitudes, personality, and behavior, The Dorsey Press, Chicago.

Ajzen, I. (1991) The theory of planned behavior, *Organizational Behavior & Human Decision Processes*, 50, 2, 179-211.

Ghani, J. A., Supnick, R. & Rooney, P. (1991) The experience of flow in computer-mediated and in face-to-face groups, Proceedings of the Twelfth International Conference on Information Systems, New York, NY.

Shneiderman, B. (1998) Designing the User Interface - Strategies for Effective Human-Computer Interaction, Addison-Wesley.

Tractinsky, N. (1997) Aesthetics and Apparent Usability: Empirically Assessing Cultural and Methodological Issues, *Proceedings of the CHI 97*, Atlanta, GA.

Zhang, P., Benbasat, I., Carey, J., Davis, F., Galletta, D. & Strong, D. (2002) Human-Computer Interaction Research in the MIS Discipline, *Communications of the AIS*, 9, 20, 334-355.