A Talk given at Sensometrcs, 2014.

1. Thanks and good morning every one. Thanks for inviting me to give a talk at this meeting; it is a big honor. And thanks for coming to my talk. I am a quantitative psychologist, who retired from McGill a few years ago, and is now an adjunct professor at University of Victoria, BC.
2. My mission today, I understand, is to overview a technique called Constrained PCA (CPCA) and its applications to sensometric research. In CPCA we generally consider the situation, in which we have a data matrix Z, but also matrices G and/or H of external information about the rows and columns, respectively, of the main data matrix. The rows of Z often represent subjects, and the columns represent variables, in which case G may be the matrix of demographic information about the subjects (e.g., gender, age, level of education), and H may be a stimulus design matrix.
3. CPCA consists of 2 major phases, External and Internal Analyses. EA applies regression analysis with G and H as predictor variables, and the rows and columns of Z as criterion variables. Regression analyses decompose Z into several additive constituents. IA applies PCA to each decomposed matrix (obtained by EA) to further explore structures within the component. Since each decomposed component has specific meaning, PCA results may be more interpretable.
4. Here are some possible decompositions of Z in EA. Z is decomposed into two when only G or H is available, and into four when both G and H are used. Here, P_G is defined by G(G'G)^{-1}G', is the orthogonal projector onto Sp(G), and Q_G = I – P_G, its orthogonal complement, is the projector onto the null space of G'. These two are symmetric, idempotent, and mutually orthogonal. P_H and Q_H are similar, but they apply to the rows of Z. P_GZ represents the portion of Z that can be explained by G, and Q_GZ the portion of Z that cannot be explained by G. Similarly, P_GZP_H: the portion of Z that can be explained by both G and H, Q_GZP_H: the portion that can be explained by H but not by G, P_GZQ_H: the portion that can be explained by G, but not by H, Q_GZQ_H: the portion that can be explained by neither G nor H.
5. IA, on the other hand, amounts to Singular Value Decomposition (SVD). It looks for, within each term of the decompositions, the subspace that captures the largest variations in the space, which is nothing but applying PCA to each term (and extracts a few principal components from each term).
6. Example No. 1. In this data set, one hundred subjects made pairwise preference judgments on 9 stimuli; stimuli were presented in pairs to the subjects, who were asked to rate the degree to which they prefer one over the other on 25-point rating scales. The 9 stimuli were celebrities in three distinct groups, 3 politicians (Mulroney, Reagan, Thatcher), 3 athletes (Gareau, Gretzky, Podborsky), and 3 entertainers (Anka, Hunter, Murray). Z is a 100 subjects by 36 stimulus pairs matrix.
7. A vector preference model was fitted that represents stimuli as points and subjects as vectors in a MD space, and subjects' preferences are predicted by the projections of the stimulus points onto the subject vectors. This is a special case of CPCA (The 2nd decomposition), which uses only H. Here, H is a design matrix for pair comparisons. For 4 stimuli, this matrix looks like this. (The rows of this matrix may be permuted and/or reflected.) And no G is used.
8. Here's the derived stimulus configuration in 2 dimensions. The 9 stimuli are labeled by integers from 1 to 9. Subject vectors are indicated by pointed arrows emanating from origin (only tips of the arrows are indicated), but only for the first 10 subjects (out of 100 in total). The stimulus points are surrounded by ellipses which indicate 95% confidence regions obtained by the bootstrap procedure. Here, we have 3 politicians; here, we have 2 athletes; here, we have 3

entertainers plus 1 athlete (JG), clustered together. This reflects the so-called similarity effect in pair comparison judgments, in which similar stimuli are easier to compare than dissimilar stimuli (e.g., a choice between RR and MT is easier to make than that between RR and WG) due to higher covariances, tending to cluster together. JG (number 4, who is an athlete) is an exception, but it may be because many subjects thought she was an entertainer. Vectors labelled by E and N (which are not very much different from each other) indicate mean preference vectors of anglophones and non-anglophones, respectively, and V indicates the grand mean preference vector. (This kind of information could be used as G and a full four-term decomposition of Z could have been possible. This was not done because what could be explained by such G would have been small.)

9. When G and/or H consist of more than one subset of variables, finer decompositions of Z are possible that correspond with finer decompositions of projectors. Here are some examples: (1) This holds when X and Y are orthogonal. (2) This holds when P_X and P_Y are commutative (or X and Y are orthogonal except for their intersection space). This decomposition is used for two-way ANOVA without interaction effects. (3) This holds unconditionally and is useful when we first fit X or Y, and then fit the other to the residuals. (4) This is useful when we have a constraint like B'C = 0 (or AC* = C) on the matrix of regression coefficients C. The terms on the right are mutually orthogonal. Example No. 2 will use the second decomposition.
10. One hundred rectangles were generated by combining 10 levels of height and 10 levels of width. Four groups (1$^{st}$, 3$^{rd}$, 5$^{th}$, and 7$^{th}$ graders) of approximately 40 subjects in each group made two-category judgments (either large or small) on the rectangles, and the frequencies with which the rectangles were judged large were counted and used as ordinal measures of the subjective area of the rectangles. Here Z is a 100 rectangle by 4 age groups matrix. The ordinal measure means that we look for the “best” monotonic transformations of the columns of the data matrix.
11. G consists of two submatrices, G_H and G_W, which are the matrices of dummy variables indicating levels of height and width, respectively. The projectors corresponding to these matrices are commutative, and the second decomposition in the previous slide applies. There is a common space between X and Y, which turns out to be the space of a constant vector. Eliminating the effect of this vector (pertaining to the grand means) also eliminates the third term in (2). So we have a two-way ANOVA-like situation without interaction effects. The model fitted may be formally called the WAM, which may be written as where u_H and u_W quantify levels of height and width (which are assumed common across the four groups) , respectively, and v_H and v_W are the weights attached to the height and width dimensions by the four age groups.
12. Here are the estimated u vectors. They are both fairly linear.
13. Here are the estimated weights. The weight attached to the height tends to decrease while that to the width tends to increase, confirming the empirical evidence that younger children put more emphasis on height (than width) when they make area judgments of rectangles. (The data were initially collected to verify this phenomenon.)
14. Here are the estimated “optimal” monotonic transformations of the data for the four groups (mild ceiling effects).
15. Further extensions: A nonnegative definite matrix K such that rank(KG) = rank (G) may be used as a non-identity metric matrix. This is useful in the WLS estimation if the rows of Z are correlated and/or have hetero-schedastic variances.

16. Other possible extensions include, among others, oblique projectors (for IV estimation), ridge operators (for ridge regularization), components restricted to lie in the column space of Z (we simply set H = Z'G instead of directly using G), the case of non-orthogonal decompositions (iterative procedures, DCDD and GSCA), etc.
17. Here is some information regarding software for CPCA and others.
18. Here are a list of relevant literature. Three books, two of which have been published already, and the third, soon to come out. If you enter Promo Code EZL20 when you order the second book, you may get a 20% discount.
19. This is the front cover of the second book listed.
20. The same for the third book.
21. Thank you for your attention.