



Composite scores for concurrent behaviours constructed using canonical discriminant analysis

Matthew Kramer^{a,*}, Paul J. Weldon^{b,1}, John F. Carroll^{c,2}

^a Biometrical Consulting Service, Beltsville Agricultural Research Service, U. S. Department of Agriculture

^b Conservation and Research Center, Smithsonian Institution

^c Animal Parasitic Diseases Laboratory, Beltsville Agricultural Research Service, U. S. Department of Agriculture

ARTICLE INFO

Article history:

Received 18 July 2008

Initial acceptance 15 September 2008

Final acceptance 8 December 2008

Published online 21 January 2009

MS. number: AS-08-00464

Keywords:

Amblyomma americanum

chemoreception

composite score

discriminant analysis

lone star tick

statistics

An animal's response to a given stimulus may be assessed by measuring more than one type of behaviour. For example, in studies of the chemosensory-based preferences of squamate reptiles for prey organisms, both the frequency of tongue flicking, a behaviour linked to vomerolfaction, and the attempted ingestion of prey-scented objects have been measured. Burghardt (1969; cf. Cooper & Burghardt 1990) proposed combining both tongue-flick counts and biting attacks on scent-laden swabs into a composite index, called the tongue flick–attack score, by which overall preferences for prey-derived chemicals could be compared. In Burghardt's scheme, an attack, which presumably reflects greater interest in prey-derived cues, was arbitrarily accorded a value equal in weight to the highest number of tongue-flicks displayed by any nonattacking subject to any condition, plus the number of seconds in a trial minus the attack latency. If scores were log-transformed, a value of one was added prior to transformation.

With only two behavioural components to include in a composite measure, as with the tongue flick–attack score, a workable ad hoc system of weighting is not difficult to devise. However, as the number of behavioural variables increases, a composite scoring system using ad hoc methods becomes more difficult to construct and justify. The key goal for devising such a weighting system is to reduce the dimensionality of the measures such that the axis of measurement (e.g. interest in chemical stimuli) involves only one or two values per subject. In addition, the weightings of the measured components should reflect their biological significance; they should match what is known or believed about how the overt behaviours lie along the axis of measurement. We propose the use of canonical discriminant analysis to construct a weighting system for concurrently measured behavioural variables. These variables can be quantitative or qualitative, although the latter would need to be converted into dummy variables.

Discriminant analysis, which was advanced more than 70 years ago (Fisher 1936), is a standard statistical technique for developing criteria useful for separating individuals or samples into a priori known groups. In a linear discriminant analysis, the criteria are constructed by finding the weight or loading that, following multiplication by the values of the corresponding variable (standardized to mean 0, variance 1) and summation of these products, maximizes the 'distance' between groups. One can then assess how

* Correspondence: M. Kramer, USDA, ARS, Biometrical Consulting Service, BARC-West, Bldg 005, Beltsville, MD 20705, U.S.A.

E-mail address: matt.kramer@ars.usda.gov (M. Kramer).

¹ P. J. Weldon, Conservation and Research Center, Smithsonian Institution, 1500 Remount Rd, Front Royal, VA 22630, U.S.A.

² J. F. Carroll, USDA, ARS, Animal Parasitic Diseases Laboratory, BARC-East, Beltsville, MD 20705, U.S.A.

well groups separate on one or more discriminant axes. Our proposed use of discriminant analysis takes advantage of this algebra to formalize a long-sought weighting system for creating composite scores. An 'optimal' weighting (within the constraints of discriminant analyses) will provide weights to best differentiate and, on each dimension, order the stimuli based upon the animals' measured behaviours. Thus, a single event, such as an attack on a prey-scented swab, can be more informative than behaviours performed repeatedly, such as the number of tongue-flicks, and the discriminant function will give greater weights to behaviours that better separate the responses to different stimuli. A discriminant analysis also may be used to determine the dimensionality of the score; that is, should it contain one or two values? This feature becomes important when large numbers of behavioural measures are scored and when more than one axis is involved (e.g. food preference and sexual attraction).

Linear discriminant analysis and canonical discriminant analysis originally were developed to address different problems. However, they are mathematically equivalent and produce the same weights (although the signs may be reversed, depending on the software). Discriminant analysis in the ethological literature (and elsewhere) typically has been used to determine whether the discriminant functions can correctly classify individuals, or as a dimension reduction technique (e.g. Adams et al. 1987; Ehlinger & Wilson 1988; Coleman & Wilson 1998; Zuk et al. 1998; Guillermo et al. 2004; Morisaka et al. 2005; Tinker et al. 2008). Composite scores created using discriminant analysis have been used in the plant sciences (see Vaylay & van Santen 2002). Our use of discriminant analysis creates a composite score to ask whether different stimuli are perceived to be different by animal subjects, as inferred by their behaviour.

It may be helpful to provide guidelines for the kinds of behavioural data that might be analysed using this method. The two main criteria are that (1) two or more measures of behaviour are taken for each subject during a test session (the measures combined into a composite score) and (2) the scores are created to best separate a priori groupings of individuals, grouped using either some characteristic of the individual (age, sex, etc.) or imposed by the researcher (test stimuli differ for each group, rearing conditions differ for each group, etc.). For example, if several measures of aggressive behaviour are recorded for each individual and the researcher wants to rank aggressiveness, this would not be the appropriate method; (1) is satisfied, but (2) is not. However, if several measures of aggressive behaviour are recorded, there are several different rearing conditions, and the researcher wants to know whether rearing conditions affected aggressive behaviour, then (1) and (2) are satisfied and this methodology could be useful. Note that social behaviours, such as aggression, may be more difficult to analyse (using this or any other method) if the effect of conspecifics' behaviours is to be considered. We present our methodology for the simplest case; individuals are tested once independently. The extension to individuals tested more than once (e.g. each individual is tested once on each stimulus: a cross-over design) is conceptually not difficult; we refer to software at the end of the paper that could be used.

Protocol

We describe a three-step protocol. First, a stepwise procedure is used to identify which behavioural measures are useful to separate the responses to the stimuli. Second, a discriminant analysis is conducted using the subset of the measures selected to obtain the weights and create the composite scores. The third step entails formal testing to ascertain which of the various stimuli differ from others, based on the composite scores. At the outset of a study,

a researcher may not know which behaviours are relevant to evaluate a subject's response. Using this method, less discriminating behaviours are downweighted or eliminated from the analysis.

Although the three steps are not independent (using the data three times, once to decide on which measures to use, once to derive weights, and again to see whether the stimuli differ based on the composite scores), the method should yield valid results. While the nonindependence suggests that the final comparison of stimuli will be tested too liberally, this may be balanced in studies where animals are tested in unnatural or stressful situations, and by using proxies (the composite scores) to represent the underlying latent variables (that is, additional 'noise' is introduced into the data beyond the usual biological variation, and this will tend to make tests more conservative). If animal subjects are readily available, different sets of subjects could be used for each step, which would remove the correlation between the steps.

To illustrate this protocol, we present the results of our studies on responses by lone star ticks, *Amblyomma americanum* L., to four potentially deterrent chemicals; more extensive results involving additional compounds will be reported elsewhere. To illustrate different aspects of the protocol, results from two data sets are presented, one from an avoidance test and another from a toxicity test, the latter assessing ticks' climbing and righting performances following exposure to the chemicals. While the testing of the ticks and the behaviours recorded differed, the same analysis protocol was used for both data sets. Overall results were similar for the two assays we performed.

Methods

In two separate tests, ticks were examined for responses to four monoterpenes: limonene, citronellal, citronellol and citronellyl acetate (Bedoukian Research Inc., Danbury, CT, U.S.A.). These or related compounds have been indicated as tick repellents or acaricides (Cook 1992; Chungsamarnyart & Jansawan 1996; Thorsell et al. 2006). Compounds were dissolved in acetone in 0.1 M solutions. Acetone was used as a control. Controls were tested in every session, although not all compounds were (i.e. sessions were incomplete blocks). We tested for, but did not find, a session effect, and have ignored session effects in the subsequent analyses.

Nymphs of *A. americanum* were obtained from a colony maintained at the Agricultural Research Service, Knippling-Bushland Livestock Insects Research Laboratory, Kerrville, TX, U.S.A. and held at 23–24 °C, ≈97% RH on a 16:8 h light:dark cycle. Temperatures in the laboratory were 21–26 °C and the RH was 23–58%. Testing was conducted from 0900 to 1700 hours. At the time of testing, the ticks had been in the nymphal stage for 2–6 months.

Avoidance test

Climbing is a critical behaviour in host seeking for many ixodid ticks. In nature, questing ticks ascend vegetation to vantage points where they contact passing hosts (Sonenshine 1993). When encountering a vertical projection, a host-seeking tick tends to climb, particularly when host-derived cues (e.g. kairomones and vibrations) are present.

Ticks were tested on 1 × 8 cm strips of recycled bond paper marked at 1 cm intervals to 6 cm. We evenly applied 6 µl of each 0.1 M solution or 6 µl of acetone by micropipette to each strip between the 2 cm and 6 cm marks (4 cm²). The strips were then air-dried for 1 min. A vial containing ticks was opened in a moated petri dish. Individual ticks were touched with an untreated end of the paper strip and allowed to mount it. Once a tick had mounted the strip, the paper was suspended vertically from a bulldog clip attached to a peg. The strip was positioned 20 cm from a light

(2.75 W), which shone from behind the paper strip, thus allowing the tick to be monitored on either side of the paper. A moated petri dish was placed beneath the strip to confine ticks that dropped off the strip. Ticks were scored for their latencies to cross the 2 cm (Q2), 3 cm (Q3) and 6 cm (Q6) marks on each strip, the number of times they backcrossed and recrossed the 2 cm mark (Qbc), and their latency to drop off the paper strip (Qld). 'Q' indicates the variable is quantitative (see below). Sessions were concluded when ticks passed the 6 cm mark or dropped off the paper strip, or after 10 min had elapsed. We used 138 ticks for this experiment.

Righting and climbing (R/C) test

Eight ticks were confined in a filter paper packet treated with a test solution or acetone for 1 h (Carroll et al. 2005). Immediately after an exposure packet was opened, each nymph was removed and placed on its dorsum on a clay substrate. Each tick was encircled by a vertical filter paper cylinder (~3.2 cm diameter, 3 cm high, with ~3 mm pushed into the clay). We recorded the time that elapsed until a tick righted itself and the time elapsed until the tick climbed to the rim of the cylinder (time limit of 15 min). The location of the tick at the end of the test was recorded as 0 (on bottom), 0.33 (side, lower 1 cm), 0.67 (side, upper 2 cm), or 1 (on rim or outside of cylinder).

Following an R/C test, one of us (J.F.C.) positioned the tip of his left forefinger 1–2 mm from the anterior of a tick (mouthparts or forelegs if the latter were extended) and about 45° to its right. The tick was given 5 s to climb onto the fingertip. If the tick did not climb onto the fingertip, the process was repeated with right finger 45° to the tick's left, and repeated once more with each finger if the tick failed to climb. The tick was given a final opportunity to climb with the left finger placed 1–2 mm directly in front of the tick, so that it could mount the finger without having to turn towards the finger as in the previous profferings of the finger. Ticks were motionless or crawling at the start of this test. These behaviours are referred to below as Qr (time to right), Qt (time to top), Qloc (location at end of session), Ih (whether the tick climbed onto finger), and Qth (number of trials until tick climbed onto finger); 'Q' indicates that the variable was quantitative, 'I' indicates that it was 0 or 1 (indicator). Three replicates of eight ticks each were exposed to each compound except for the acetone control, which was always included in a block (day) of trials. We used 176 ticks for this experiment (96 ticks tested against compounds, 80 controls).

Indicator variables and variable transformation

Five quantitative behavioural measures were taken in the avoidance test and four from the R/C test. For most of these, an indicator (dummy) variable was created; 1 if the behaviour was performed, 0 otherwise (we abbreviate these variables below by replacing the 'Q' with an 'I'). If the behaviour was not performed, the quantitative measure also received a 0; this value choice is arbitrary, but all individuals not performing the behaviour must receive the same value. For example, if the tick did not cross the 6 cm mark in the avoidance test, it received 0 for the indicator variable and 0 for the time variable. This resolved the problem of handling behaviours that were not performed by all individuals;

both the indicator and quantitative variables are included when creating the composite score. A similar procedure of creating an indicator variable can be used for a timed behaviour that continues beyond the observation period. For example, if the behaviour was searching and some animals were still searching at the end of the observation period, an indicator variable denoting whether the animal was searching at the session's end should be created. All time variables were square-root transformed, which typically normalizes the kinds of time variables we measured. However, analyses done without this transformation furnished similar composite scores.

Results

Avoidance test

Stepwise analysis and display of composite scores. We used a stepwise discriminant analysis (using SAS proc stepdisc, SAS Institute Inc., Cary, NC, U.S.A.) to examine how many of the variables were useful discriminators of the effects of the compounds. Five of the 10 variables were deemed useful: I2 (passed the 2 cm mark), Ibc (recrossings of the 2 cm mark), Qbc (number of recrossings of 2 cm mark), I6 (passed the 6 cm mark) and Q6 (square root of time to pass 6 cm mark) (Table 1).

We calculated the weights using both SAS (with proc candisc and proc discrim, the latter using the canonical discriminant weights) and the R statistical package (lda in the MASS library, <http://www.r-project.org/>). Identical results were obtained, except that the signs were reversed for the weights (and composite score) with R. A variety of methods is available in discriminant analysis that produces different discriminant functions. Canonical discriminant weights are the default in SAS and are appropriate for this application. We were not concerned about meeting the assumptions of a discriminant analysis because the goal was to produce a composite score for each subject; we simply used the underlying algebra to combine measurements.

This method can produce r discriminant functions, where r is the smaller of the number of behavioural variables and the number of conditions minus one to be discriminated. Each function explains a portion of the remaining total canonical correlation between conditions and variables not accounted for by previous functions, and each linear combination (of weighted variables) is uncorrelated with the others. Ideally, only the first discriminant function suffices so that a subject's score reduces to just one number. However, if stimuli belong to different classes (e.g. food, sexual attractants), one would expect that a one-dimensional composite score would be insufficient to separate conditions (i.e. food stimuli might separate on one axis, sexual attractants on another). In these experiments, with all compounds presumably falling along a single axis, we expected that a single canonical axis would capture the majority of the total canonical correlation (this latter quantity is conceptually the maximum correlation between the compound groupings and the measured behaviours explainable for these data by this class of model). For the avoidance test, we found exactly that: 91% of the total canonical correlation was explained by the first of the four discriminant functions. We give the calculated weights in Table 1.

Table 1

Canonical discriminant function (CDF) weights (loadings) for avoidance test and righting/climbing test

Behaviour*, avoidance test	Weight (1st CDF)	Behaviour, righting/climbing test	Weight (1st CDF)	Weight (2nd CDF)
I2 (passed the 2 cm mark)	-3.10	I1 (climbed to rim)	-2.30	2.90
Ibc (recrossed 2 cm mark)	2.16	Qt (sqrt of time to climb to rim)	0.06	-0.20
Qbc (number of recrossings of 2 cm mark)	0.28	Ih (tried to climb on host)	-1.79	-1.07
I6 (passed the 6 cm mark)	-4.05	Qth (number of host-acquisition attempts)	0.53	0.06
Q6 (sqrt of time to pass 6 cm mark)	0.14			

* Indicator variables begin with 'I', quantitative variables with 'Q'.

If each of the original variables had been normally distributed, a standard plot of the data on the first two discriminant axes would depict how the groups separated. For these data, with indicator variables, the standard plot produced patterns that were difficult to interpret (Fig. 1). Instead, we used an empirical distribution function (EDF) plot (Cleveland 1985) to visualize how the compounds separated along the first discriminant function (tested formally below) (Fig. 2). This plot also may identify problems with the composite scores (e.g. gaps and outliers).

To construct values for the Y axis for this plot, the composite scores (one per subject) for each compound were ranked separately, and the ranks were divided by the number of individuals tested with each compound, plus one. The X axis gives the composite score corresponding to each rank; the score values are joined by line segments separately for each compound. Individual scores are represented by points where the line segments join. The interpretation of the EDF plot for the avoidance test is straightforward (Fig. 2). Completely separable compounds have lines that do not overlap with any other lines (on the X axis; there are none in Fig. 2), and poorly separable compounds have lines that overlap or are close to each other (e.g. acetone and limonene). A line with a steep slope represents a compound for which subjects had similar composite scores (e.g. limonene). A line with a gentle slope represents a compound for which there was substantial variation among individuals (e.g. citronellyl acetate). A line with a long stretch having a low slope in the middle represents a gap in the scores (e.g. citronellal), which may result when individuals of different classes respond differently (e.g. different age or sex classes); however, we suspend interpretation of these specific results. Outliers were readily identified. One outlying high score was evident for acetone; one with a low score occurred for citronellyl acetate. The median composite score for each compound (on the X axis) occurred at $Y = 0.5$.

Compound differences. Differences between responses to the compounds can be tested formally using the composite score. If the scores meet the assumptions of ANOVA, a parametric test can be used. Otherwise, the corresponding nonparametric test should be used. Note that if two or more composite scores are necessary (i.e. the dimensionality is >1), each set of scores can be tested separately (they are orthogonal; a MANOVA approach assumes the

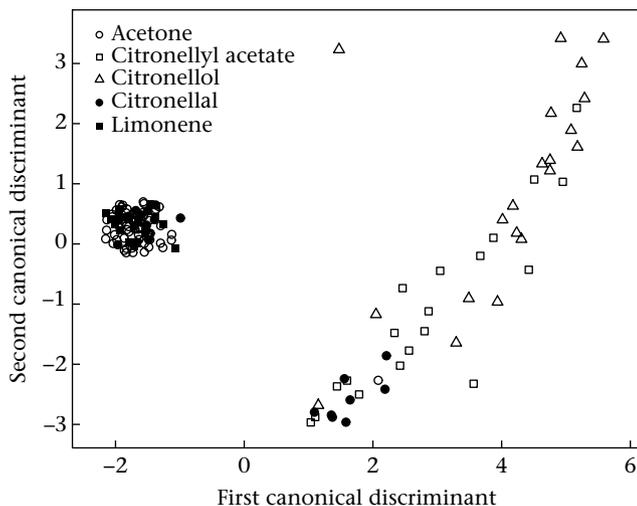


Figure 1. Scatter plot of composite scores on the first and second canonical discriminant axes. Points have been jittered (a small amount of random noise added to the values) to better separate them. In particular, points in the cluster at $(-1.5, 0.2)$ had similar Y axis values.

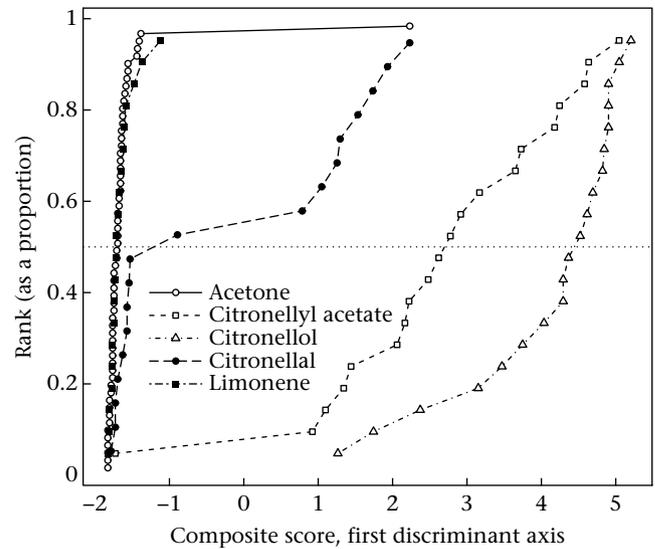


Figure 2. Empirical distribution function (EDF) for composite scores on the first discriminant axis for avoidance test. The Y axis gives the within-testing compound rank, divided by the number of ticks tested, plus one. Points represent individual tick composite scores.

scores are correlated). For these composite scores, the Anderson–Darling test for normality was met for two of the compounds (citronellol: $P = 0.17$; citronellyl acetate: $P = 0.31$) and was rejected for the others (including acetone, even with the outlier removed: all $P < 0.01$). Of more importance, variances of the composite scores among the groups differed greatly, from 0.052 for limonene to 2.734 for citronellal, and were strongly rejected by a Bartlett test of homogeneity of variances ($P \ll 0.01$). Thus, we used a nonparametric Kruskal–Wallis rank-sum test to determine whether compounds differed in their effects ($P \ll 0.01$, $df = 4$), followed by a test of pairwise comparisons (the R function `kruskalmc` in the `pgirmess` library) (Table 2).

Righting and Climbing Test

Composite scores. Analysis of the R/C test yielded results similar to those of the avoidance test. The stepwise discriminant procedure determined that four of the seven candidate variables, It (climbed to rim), Qt (time to top), lh (climbed onto finger) and Qth (number of trials until tick climbed onto finger), were useful to create composite scores. The first discriminant function explained 78% of the total canonical correlation, the second explained 16%. While ordinarily one might ignore the second discriminant function, we include it here to illustrate an analysis yielding two scores per individual.

Compound differences. On the EDF plot (Fig. 3), the first discriminant function grouped the distribution of scores for acetone and

Table 2

Means (standard deviations) of the canonical discriminant scores for avoidance test and righting/climbing (R/C) test

Compound*	Avoidance test, 1st CDF	R/C test, 1st CDF	R/C test, 2nd CDF
Acetone	-1.71 (0.54) a	-0.56 (0.97) a	0.27 (1.07) b
Limonene	-1.73 (0.23) ab	-0.37 (1.04) ab	-0.49 (1.24) ab
Citronellal	-0.20 (1.65) bc	0.31 (1.08) b	-0.55 (1.07) a
Citronellyl acetate	2.76 (1.55) cd	0.28 (1.24) b	-0.31 (0.91) ab
Citronellol	4.30 (1.09) d	1.71 (0.69) c	0.35 (0.37) b

* Composite scores are given for the first canonical discriminant function (CDF) for the avoidance test, and for the first and second canonical discriminant function for the righting/climbing test. For each column, compounds not found to be significantly different using a posteriori comparison based on Siegel & Castellan (1988, pp. 213–214) are followed by the same letter.

limonene together and those for citronellal and citronellyl acetate together, and it separated citronellol, as in the avoidance test results. Tests for normality were met only by citronellal ($P = 0.24$), were borderline for citronellol ($P = 0.07$) and citronellyl acetate ($P = 0.05$), and were rejected for acetone ($P < 0.01$) and limonene ($P < 0.01$). A Bartlett test of homogeneity of variances was not rejected ($P = 0.10$). A Kruskal–Wallis rank-sum test found that at least one compound differed from the others ($P < 0.01$), with similar results to those from the avoidance test for pairwise comparisons (Table 2).

The second discriminant function grouped the compounds differently. Composite subject scores for citronellol were similar to each other (note the steep slope in Fig. 3), and acetone had somewhat higher scores than the remaining compounds. Not surprisingly, because of the inclusion of the citronellol data, the Bartlett test of homogeneity of variances failed ($P < 0.01$). The Kruskal–Wallis test indicated that at least one compound differed from the others ($P < 0.01$). The results of the a posteriori multiple comparison test are given in Table 2, where only two pairs of compounds, acetone–citronellal and citronellol–citronellal, differed from each other.

Visualizing the contribution of original variables to composite score. After creating composite scores, the relative importance of a score's components to the discriminant function for each variable–group combination may be visualized. This differs from examining the original variables, one at a time, because the latter ignores the effect of the weighting and the different scales of the original variables. One way to visualize a decomposition of the composite scores is to create a 'heat map' (here in shades of grey using the R image function), with each rectangle representing the relative contribution of each variable–compound combination (Fig. 4). The values were obtained by first standardizing each variable, then multiplying the mean of each standardized variable for each variable–group combination by the weight from the discriminant function.

For the first discriminant function, the two most variable columns (behaviours with a larger contribution to separating the compounds) are Ih and Ih; this axis may represent a host-acquisition axis. For the second discriminant function, most of the differences

appeared to derive from contributions of the Ih and Qth variables. For this data set, this second function would ordinarily have been ignored because it explained little of the total canonical correlation. Note that the signs (=shade of grey) for Ih and Qth were almost the reverse of the first discriminant function.

Extensions

The main problem with using an ad hoc method to create a composite score is justifying a comparison among subjects belonging to different classes (species, age, sex, etc.), since the ad hoc weights may differ between classes based largely on subjective judgments. For example, had we also tested a less active tick than *A. americanum* (Waladde & Rice 1982), the largest source of variation in the behaviour measured might be due to species differences, and not to different compound efficacies. If composite scores had been calculated separately using discriminant functions for each group, then one could reasonably compare the two species, since each species would have scores optimal for distinguishing among their responses to the compounds. One could then ask whether two compounds that are readily distinguishable by one species are also distinguishable by the other. It would not be necessary for the measured behaviours to be the same for both species, as long as the behaviours used to develop the composite scores for both species captured the underlying deterrence axis to a similar degree.

Another extension of this method occurs when the same subject is tested on more than one stimulus (i.e. a cross-over design is used), where one must account for the additional within-subject correlation in the discriminant analysis. This topic has been discussed elsewhere (e.g. Tomasko et al. 1999; Marshall & Baron 2000). Although we know of no off-the-shelf software specifically designed to estimate coefficients for this model, macros are available for S-PLUS (Insightful Corp., Seattle, WA, U.S.A.) and SAS from Werneck et al. (2004). This approach is based on the mixed model framework, where individuals are treated as a random block effect (allowing individual intercept adjustments) or the within-subject time series correlation is modelled; both may be necessary. A mixed models approach could also be used to create composite scores to determine whether individuals differ, if one collects data

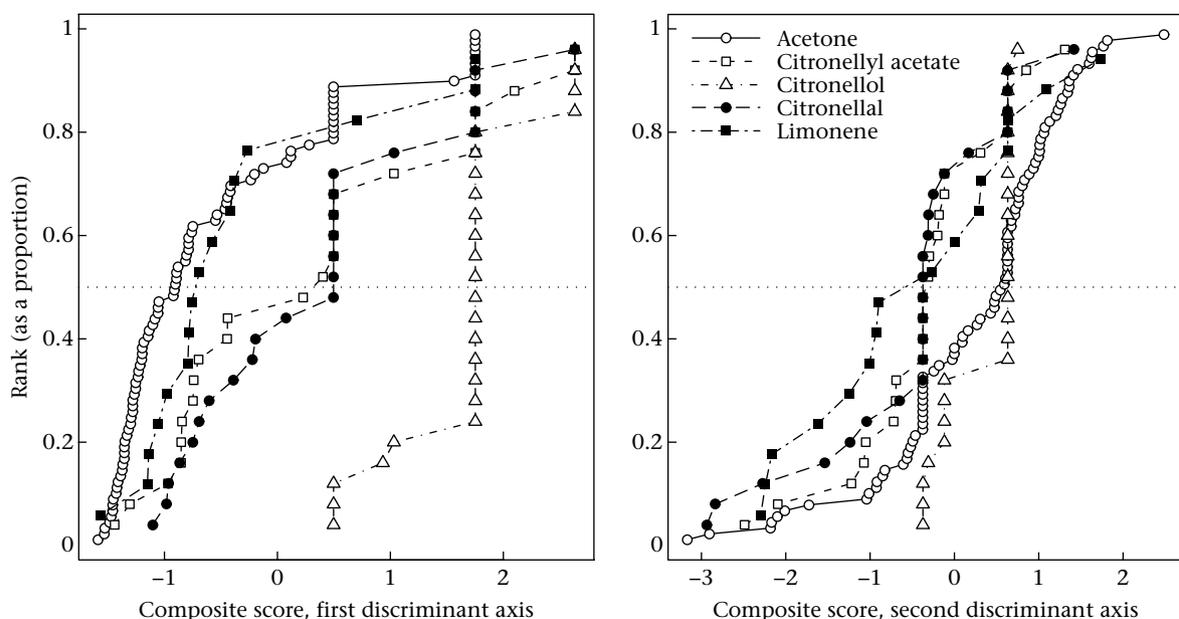


Figure 3. Empirical distribution function (EDF) for composite scores on the first and second discriminant axes for righting/climbing (R/C) test. The Y axes give the within-testing compound rank, divided by the number of ticks tested, plus a value of one. Points represent the individual tick composite scores.

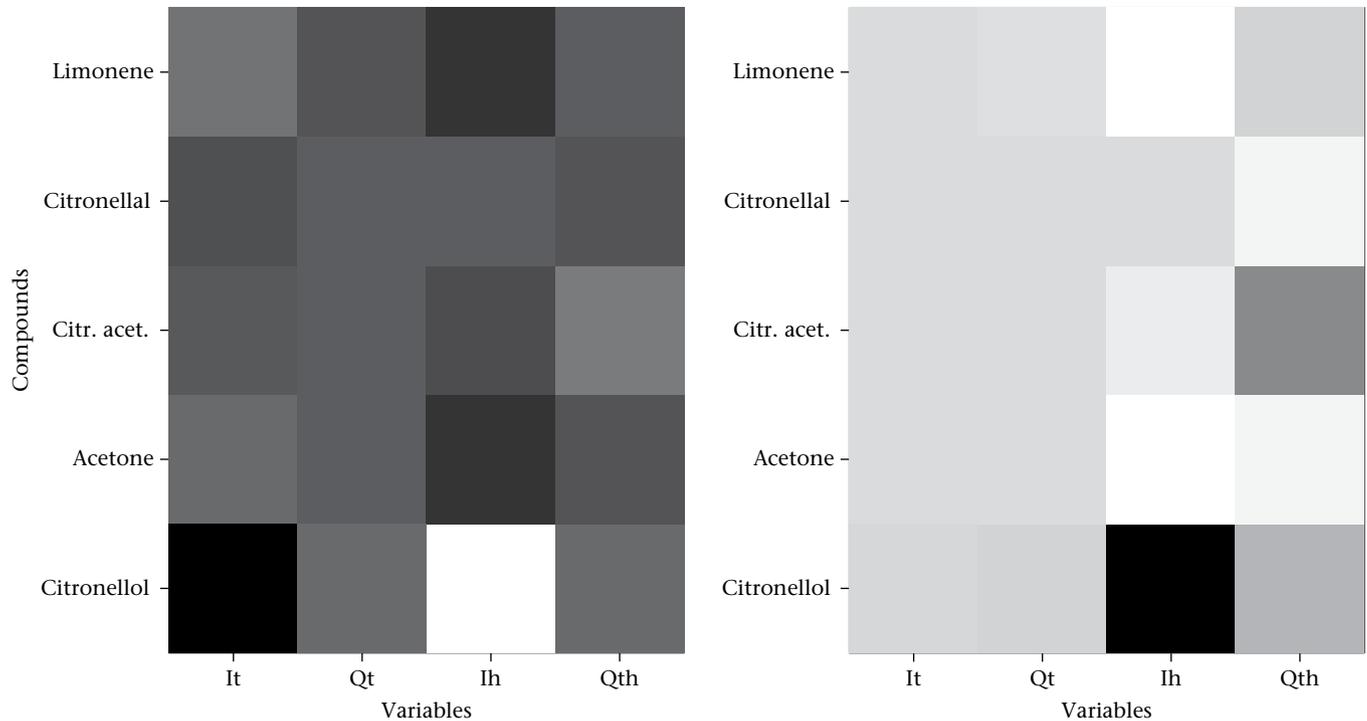


Figure 4. Heat map representation of the mean value of the contribution of each behaviour to the composite score for the righting/climbing test. The panels give results for the first and second discriminant axes, respectively. Citronellyl acetate is abbreviated as 'citr. acet.'.

repeatedly on individuals (so the repeated measures serves the same purpose as replicated individuals in the methodology we described). The model would probably need to include within-subject time series correlation, a random time of test (e.g. day) effect, and perhaps a trend to account for changes over time, such as habituation.

The magnitude of the loadings themselves may be of interest in phylogenetic studies, where responses to the same stimuli can differ both quantitatively and qualitatively between species. One may be able to use the loadings as behavioural characters for a species when constructing cladograms, as the loadings will grow or shrink in absolute value as the behaviour becomes more or less useful for separating responses to the test stimuli.

Summary

We have demonstrated a method of creating a composite score of observed behaviours useful in tests of a priori groupings of subjects (say, grouped by stimuli to which animals are exposed) using canonical discriminant functions as a base. This method combines the various behaviours measured in a way that is optimal for discriminating among the groups. In addition, it indicates the appropriate dimensionality of the scores. That is, should individual scores have multiple components, reflecting the number of latent axes represented by the set of stimuli presented? We also demonstrate some graphical tools useful to display these kinds of results.

Robert H. Bedoukian (Bedoukian Research Inc., Danbury, CT, U.S.A.) kindly supplied pure compounds.

References

Adams, R. A., Lengas, B. J. & Bekoff, M. 1987. Variations in avoidance responses to humans by black-tailed prairie dogs (*Cynomys ludovicianus*). *Journal of Mammalogy*, **68**, 686–689.

Burghardt, G. M. 1969. Comparative prey-attack studies in newborn snakes of the genus *Thamnophis*. *Behaviour*, **33**, 77–113.

Carroll, J. F., Kramer, M., Weldon, P. J. & Robbins, R. G. 2005. Anointing chemicals and ectoparasites: effects of benzoquinones from millipedes on the lone star tick, *Amblyomma americanum*. *Journal of Chemical Ecology*, **31**, 63–75.

Chungsamarnyart, N. & Jansawan, W. 1996. Acaricidal activity of peel oil of *Citrus* spp. on *Boophilus microplus*. *Kasetsart Journal (Natural Science)*, **30**, 112–117.

Cleveland, W. S. 1985. *The Elements of Graphing Data*. Summit, New Jersey: Hobart Press.

Coleman, K. & Wilson, D. S. 1998. Shyness and boldness in pumpkinseed sunfish: individual differences are context-specific. *Animal Behaviour*, **56**, 927–936.

Cook, S. P. 1992. Influence of monoterpene vapors on spruce spider mite, *Oligonychus ununguis*, adult females. *Journal of Chemical Ecology*, **18**, 1497–1504.

Cooper Jr, W. E. & Burghardt, G. M. 1990. A comparative analysis of scoring methods for chemical discrimination of prey by squamate reptiles. *Journal of Chemical Ecology*, **16**, 45–65.

Ehlinger, T. J. & Wilson, D. S. 1988. Complex foraging polymorphism in bluegill sunfish. *Proceedings of the National Academy of Sciences, U.S.A.*, **85**, 1878–1882.

Fisher, R. A. 1936. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, **7**, 179–188.

Guillermo, P. C., Bond, A. B., Kamil, A. C. & Balda, R. P. 2004. Pinyon jays use transitive inference to predict social dominance. *Nature*, **430**, 778–781.

Marshall, G. & Baron, A. E. 2000. Linear discriminant models for unbalanced longitudinal data. *Statistics in Medicine*, **19**, 1969–1981.

Morisaka, T., Shinohara, M., Nakahara, F. & Akamatsu, T. 2005. Geographic variations in the whistles among three Indo-Pacific bottlenose dolphin *Tursiops aduncus* populations in Japan. *Fisheries Science*, **71**, 568–576.

Siegel, S. & Castellan, N. J. 1988. *Nonparametric Statistics for the Behavioral Sciences*, 2nd edn. New York: McGraw-Hill.

Sonenshine, D. E. 1993. *Biology of Ticks*. Vol. II. New York: Oxford University Press.

Thorsell, W., Mikiver, A. & Tunón, H. 2006. Repelling properties of some plant materials on the tick *Ixodes ricinus* L. *Phytomedicine*, **13**, 132–134.

Tinker, M. T., Bentall, G. & Estes, J. A. 2008. Food limitation leads to behavioral diversification and dietary specialization in sea otters. *Proceedings of the National Academy of Sciences, U.S.A.*, **105**, 560–565.

Tomasko, L., Helms, R. W. & Snapinn, S. M. 1999. A discriminant analysis extension to mixed models. *Statistics in Medicine*, **18**, 1249–1260.

Vaylay, R. & van Santen, E. 2002. Application of canonical discriminant analysis for the assessment of genetic variation in tall fescue. *Crop Science*, **42**, 534–539.

Waladde, S. M. & Rice, M. J. 1982. The sensory basis of tick feeding behavior. In: *Physiology of Ticks* (Ed. by F. D. Obenchain & R. Galun), pp. 71–118. New York: Pergamon.

Werneck, K. D., Kalb, G., Schink, T. & Wegner, B. 2004. A mixed model approach to discriminant analysis with longitudinal data. *Biometrical Journal*, **46**, 246–254.

Zuk, M., Rotenberry, J. T. & Simmons, L. W. 1998. Calling songs of field crickets (*Teleogryllus oceanicus*) with and without phonotactic parasitoid infection. *Evolution*, **52**, 166–171.